UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE DES SCIENCES CHIMIQUES

Habilitation à Diriger des Recherches

présentée par

Yannis-Nicolas FRANCOIS

Soutenue le :

30 novembre 2016

pour obtenir le grade de : **Maître de Conférences HDR**

Discipline/ Spécialité : chimie analytique et spectrométrie de masse

Développement du couplage électrophorèse capillaire – spectrométrie de masse: applications à la caractérisation fine de protéines

Garant HDR :

Dr. Emmanuelle LEIZE-WAGNER	Directrice de recherche, Université de Strasbourg

RAPPORTEURS :

Dr. Valérie GABELICA	Rapporteur externe, IECB Bordeaux
Pr. Myriam TAVERNA	Rapporteur externe, Institut Galien Paris-Sud
Dr. Sarah CIANFERANI	Rapporteur interne, Université de Strasbourg

---

AUTRES MEMBRES DU JURY :

Pr. Maurice MILLET	Président du jury, Université de Strasbourg
Dr. Alain BECK	Examinateur, Centre d'immunologie Pierre Fabre
Dr. Davy GUILLARME	Examinateur, Université de Genève, Université de Lausanne

« A garanty job for life only encourage the faculty to become complaisant »

Dr. Sheldon Cooper

Remerciements

Ahhh!!!!! Cette fameuse partie des remerciements ! Ce paragraphe si important dans un manuscrit de doctorat qu'il en est parfois le plus long à écrire. En effet, il ne faut oublier personne : de la rencontre d'un jour à l'amitié de toujours, de la famille et des amis « d'avant » aux collègues si précieux des trois dernières années. Surtout chaque anecdote symbolysant trois années de labeur, de rencontre et d'amitié doit être sanctifiée dans ces quelques pages post-introduction. Je me souviens le plaisir que m'a procuré la rédaction des remerciements de ma thèse, et l'émotion que j'ai eu en lisant ceux de mes deux premiers thésards et amis Michael et Rabah. Inoubliable !!

Mais alors qu'en est-il de l'habilitation à diriger des recherches. Considérée parfois comme une seconde thèse, l'HDR est censée être l'âge de la maturité (censée seulement !!!). Alors, comme le 2 est toujours moins bon que le premier opus (cet adage n'est évidemment vrai que pour les remerciements !!), les remerciements pour mon HDR seront plus concis tout en conservant une extrême sincérité.

Ce travail a été réalisé au sein de l'UMR Chimie de la Matière Complexe (CNRS-UMR 7140) dirigé par le Professeur Mir Wais HOSSEINI. Je lui suis très reconnaissant pour son accueil au sein de l'unité.

Je tiens à remercier le Dr. Valérie GABELICA, le Pr. Myriam TAVERNA et le Dr. Sarah CIANFERANI pour avoir accepté d'évaluer les travaux présentés lors de mon HDR. Je remercie également le Dr. Davy GUILLARME d'avoir accepté de faire parti de mon jury d'HDR.

Je souhaite également remercier le Pr. Maurice MILLET qui a accepté de présider le jury de mon HDR. Merci surtout pour la grande collaboration que nous avons en enseignement au sein de la faculté de Chimie.

Je tiens particulièrement à remercier le Docteur Emmanuelle LEIZE-WAGNER pour s'être portée garant pour mon HDR. Je tiens surtout à te remercier pour le soutien que tu as porté pour ce projet et pour le développement du couplage CE-MS. Les nombreuses discussions scientifiques que nous avons régulièrement m'ont permis d'apprendre énormément et d'acquérir une expertise en spectrométrie de masse. J'espère que cette collaboration continuera longtemps avec la même réussite.

Le Laboratoire occupe une place centrale dans la vie d'un chercheur. Le mien a eu différents noms (LDSM2, LSMIS) et différents endroits (ISIS, Tour de Chimie, Institut Lebel), mais malgré

des conditions souvent précaires, on y est bien. Je voulais par conséquent remercier l'ensemble des personnes permanentes composant le LSMIS.

Que serait le service sans le labo et le labo sans le service ? La question reste posée, malheureusement la croisée des chemins est arrivée. En tout cas, merci au Dr. Hélène NIERENGARTEN, mais aussi à Romain CARRIERE, Mélanie LEBRETON et Stéphanie KOUAHO d'avoir été là toutes ces années.

Deux collaborations académiques ont été au centre de ces recherches et ont permis la bonne tenue de cette HDR.

- Tout d'abord, je tiens à remercier la plateforme protéomique de l'IBMC dirigée par Philippe HAMMANN et composée de Lauriane KUHN, Johana CHICHER et Philippe WOLF. En plus d'être des personnes exeptionnelles humainement parlant, vous êtes mes références en protéomiques et ARNomiques. J'espère continuer à travailler encore longtemps avec vous.
- Je tiens à remercier l'équipe du Dr. Alain BECK au Centre d'Immunologie Pierre Fabre. Alain, il n'y a rien à rajouter tant cette collaboration académique a bouleversé le sens de mes recherches. Et en plus, il reste tant de choses à faire… Vivement la suite !!

Deux collaborations industrielles m'ont aussi permis de réaliser ces recherches.

- Je tiens à remercier la société Sciex (anciennement Beckman Coulter) pour m'avoir fait confiance en me prêtant le couplage CESI-MS. Merci particulièrement à Michel ANSELME pour qui j'ai une très grande amitié.
- Je tiens à remercier la société Bruker Daltonics et plus particulièrement Pierre-Olivier SCHMIT et Yann HEBERT pour leur soutien tout au long de ces années.

Il ne faut jamais oublier ceux qui permettent le bon fonctionnement des systèmes et qui ont une très grande connaissance des machines. Merci à Dominique MELLON, Jean-Michel BILLMANN, Claude TROESCH, Philippe KORDEK et Patrick SCHMIDT pour leur patience, leur humour et leur gentillesse.

Enfin, l'HDR ne pourrait pas être soutenu sans ceux qui font la science, ceux qui ont la fraîcheur et l'insouciance de la jeunesse : « les étudiants ».

Sans eux, rien n'est possible. Parfois simple, parfois compliqué, un encadrement est toujours riche en enseignement. Sur ce point, j'ai eu la chance de « tomber » sur des étudiants exceptionnels.

Je dois commencer par Michael et Rabah. Sans eux deux, rien de tout cela n'aurait été possible. Tout ce qui est dans ce manuscrit vous appartient aussi, donc merci pour tout. Je

copie dans ces remerciements le commentaire que j'ai fait à vos thèses respectives. Je n'en changerai pas une ligne.

- **Michael BIACCHI** :

Alors évidemment tu t'en doute, je ne vais pas poser de questions sur ton travail Michael. Je vais juste essayer de faire un retour et un petit bilan sur ce projet un peu fou qu'a été ta thèse.

5 ans aujourd'hui que tu nous as dit « OK » à Emmanuelle et à Moi à la question « Mais pourquoi tu ne ferais pas une thèse chez nous ??? » « Franchement, tu en es capable… » « En plus tu vas voir, on a un super sujet CE-MS à te proposer. Bon y a pas de CE au labo, niveau MS, faut voir avec le service pour avoir du temps machine, pour le financement, faut voir avec ta formation continu et en gros si tout va bien, tu viendras une semaine tous les deux mois pendant quatre ans… »… « Mais ce serai génial quand même… » Ce projet Michael, je crois qu'aujourd'hui on peut te dire la vérité avec Mannue, c'était quand même de la folie pure… Mais on s'est basée sur une intuition,

Ta capacité hors du commun à chercher, à creuser, à créer, à fouiner bref ta très grande curiosité. Et il ne nous a pas fallu longtemps pour le savoir puisqu'on s'est uniquement basé tes trois semaines de présence au laboratoire pour ton Master. Donc c'était un pari.

Et aujourd'hui on peut dire que le pari a été gagnant et que c'est une grande réussite.

Alors je ne vais pas détailler toute ta thèse, l'organisation totalement anarchique, la formation de la CE que j'ai été faire à Thionville dans ton laboratoire (où j'ai toujours été très bien reçu d'ailleurs), le développement instrumental en partant de zéro c'est-à-dire de machine commerciale, les heures passées au labo devant ta CE quasiment à lui parlée. Parce que si, il y a une chose que je peux attester, c'est le sérieux avec lequel tu as menés ton projet. Tout au long de ta thèse, lorsque tu étais au labo, tu n'as jamais compté tes heures jusqu'à des soirées très tardives et musicales avec Rabah et des instants chocolats avec Marianne. Et tu as même été jusqu'à prendre toutes tes vacances au laboratoire. Bref, il m'arrivait de t'appeler à minuit au labo pour faire un point manip et évidemment tu y étais. Et ben, aujourd'hui tu peux être fier du résultat.

Ah oui j'oubliai, parce que tu aimes te faciliter la tâche (ou bien la thèse), tu as décidé de construire ta maison de tes mains, d'organiser ton très beau mariage et même comme dernier challenge d'obliger ta fille à naitre 5 jours après la fin du l'écriture du manuscrit de ta thèse.

Bref, je fini juste par te dire que ta thèse est évidemment exceptionnelle dans son déroulement et comme tu l'as dit « un thésard comme toi c'est unique et j'en aurais plu », mais je pense surtout que ta thèse est exceptionnelle par la qualité de tes résultats et par l'excellence des recherches que tu as menées au laboratoire. Et c'est très important que tu t'en convainques.

Donc, je pense réellement qu'aujourd'hui tu allies parfaitement la bonne équation entre la technique (qui est l'un de tes atouts majeurs, quoi que tu en penses) et la création scientifique. J'espère que tes expériences avenir te permettront de t'épanouir car je crois que maintenant que tu as ouvert la boite de Pandore, tu ne pourras pas t'arrêter là.

Plus personnellement, tu as été mon premier encadrement de thèse et franchement je n'aurais pas pu rêver mieux. J'ai appris énormément sur moi et sur mon travail en suivant ton doctorat. Ceci s'est en plus transformé en amitié et j'en suis très content.

- **Rabah GAHOUAL** :

Alors évidemment tu t'en doutes, je ne vais pas poser de questions sur ton travail Rabah. Je n'en trouverai plus au nombre d'heure que nous avons discuté de tes résultats. Je vais juste essayer de faire un petit bilan sur ton travail de thèse.

Déjà, je me dois de remercier Sacha Schifrin. Car c'est lui qui un jour m'a dit « Tu sais Yannis, j'ai un pote à l'ECPM qui cherche une thèse en Analytique. Tu veux que je lui parle de ton sujet ? » Evidemment j'ai dit oui et j'ai vraiment bien fait. Quelques jours plus tard, tu es venu dans mon bureau

et j'ai tout de suite vu qu'il y avait un truc. Alors évidemment, je t'ai vendu du rêve sur un nouveau couplage CE-MS qu'on allait avoir au labo. Que tout était à faire, le mettre en place au laboratoire et que le champ d'applications était énorme. Evidemment le système n'était pas encore présent et l'axe des applications était encore assez vague, mais j'ai assez vite senti que cela te tentait.

Deux mois après, tu obtenais la bourse du ministère et l'aventure a réellement commencée.

Et là, je crois qu'on a été tous vite bluffé par ta capacité hors du commun à travailler. Je peux attester que depuis le premier mois de ta thèse jusqu'à aujourd'hui je pense que tu ne t'es jamais arrêté de travailler, jamais arrêté de penser thèse. Week End, soirée, vacances, tu n'as jamais compté tes heures et je crois qu'aujourd'hui tu en a un résultat exceptionnel. Je note aussi ton organisation toujours nickel car le timing a souvent été serré car Emmanuelle et surtout moi nous t'avons très souvent sollicité. Juste pour illustrer cela, je me rappelle d'un coup de fil de Beckman au mois d'avril 2013 qui nous sollicitait pour présenter une conf d'une heure à l'ASMS le mois suivant sur des résultats que nous n'avions pas encore et que nous devions obtenir sur une nouvelle machine. Tu as accepté sans sourciller et avec l'aide indispensable de la plateforme protéomique de l'IBMC on a fait la conf le mois suivant. Parce que parmi tes nombreux points forts, il y en a un qui te caractérise et qui est vraiment primordiale, c'est ta capacité à ne jamais abandonner et à toujours vouloir répondre aux verrous scientifiques qui te sont proposés. Et c'est cette abnégation qui t'a poussé à rendre cette thèse exceptionnelle à mes yeux.

Alors trêve de blabla… Rabah au LSMIS en chiffre c'est :

11 publications dans des journaux à comités de lecture et 2 proceedings.

C'est 6 oraux acceptés dans des conférences internationales dont 4 que tu as données. Une vingtaine oraux invités dans les journées utilisateurs ou des clubs scientifiques.

C'est 2 oraux acceptés dans des conférences nationales.

C'est une douzaine de posters

Et pour finir, c'est une coupe au tournoi de Foot de l'Université, bien aidé par tes amis je crois.

Bref, chapeau bas. Je pense que tu peux en être fier et que tous tes proches peuvent l'être aussi.

Alors je vais finir par une partie plus personnelle. Outre le fait d'avoir tissé des liens d'amitié, ta thèse, cela a été pour moi une des périodes les plus enrichissante de ma vie. Pour moi, il y aura un avant et un après. J'ai énormément appris de ton travail et des innombrables discussions scientifiques ou pas que nous avons eues à 18h30 autour d'un chery coke. Et je peux dire qu'elles vont me manquer ces discussions. Et maintenant que ton apprentissage touche à sa fin et que tu as un bon recul sur ta spécialité, c'est des discussions comme celle que nous avons eu avec Mannue sur l'avenir de la MS que j'espère nous continuerons à avoir.

Je te souhaite le meilleur pour l'avenir que je sais assez incroyable. Tu as devant toi un boulevard de connaissance et je suis sûr que tu vas contribuer à faire avancer la science.

Maintenant l'avenir s'appelle Nassur SAID. Le prochain sur la liste de « mes » docteurs. Je tiens aussi à te remercier pour le travail que tu effectues actuellement au laboratoire. Là encore, j'ai eu la chance d'encadrer un très bon étudiant particulièrement doué dans la mise au point de méthodes analytiques et dans la prise en main des systèmes de CE-MS. Chaque jour j'apprends avec toi le rythme de la chimie analytique.

Enfin, il y a tous les stages M2, M1, L3 et L2 que j'ai eu durant toutes ces années. Merci à eux pour tout le travail produit et toutes ces belles histoires :

Bruno GUYONNET, Nina POURHASSAN, Alicia BURR, Clotilde DEVAUX, Jérémie GIORGETTI, Camille LINKER, Jordan KIRMANN, Charly RENARD, Iliana TSONEVA, Ricky BHAJUN, Heidi SOUFFARY, Marc LIBERATORE, Jasmine HERTZOG, Sarah DERAS, Anthony LECHNER, Julie

CANONGE, T. DURAND, H. SIRINDIL, L. ATI, E. BONFANTI, T. MARLEUX, M. PAQUET, C. ADOLF, C. MAZOHIE, M. METZELER, EL MASOUSI (1 mois), J. HEISER, S. PASHINA, T. GUGGENBUHL, M. COURBET, M. GUILLAUME, M. DAVRIL, M. MONNIER, C. BALHONZER

Bon, il va quand même falloir que j'avoue une chose importante. Au laboratoire, ces cinq dernières années ont été extraordinaires au niveau de l'ambiance entre étudiants. Outre « mes » thésards, Marianne IBRAHIM, Cécile PERRET, Nha-Thi NGUYEN-HUYNH et Margaux FRESNAIS ont complété la fine équipe du 12ème étage. Merci pour ces bons souvenirs qui marqueront le LSMIS

Plus personnellement, je tiens à remercier Lauriane KUHN et Jérôme PANSANEL et leur jolie petite famille pour nous avoir permis de se sentir chez nous, ici, dans cette belle région Alsacienne. Votre amitié nous est chère à Elise, Titou, Pauline et moi.

Enfin, et pour finir ces remerciements qui devaient être concis (on ne se refait pas !!!), je dois remercier ma famille et plus particulièrement mes parents, toujours présents à nous supporter et à nous aider. Merci d'être là, je vous aime. Ma sœur, Arnaud et leurs deux bacheliers qui pour des jeunes padawans commencent sévèrement à devenir Jedis !!

Et comme la fin du 2 ne doit pas être identique à la fin du premier opus, je ne vais pas finir cette fois-ci uniquement par ma femme Elise que j'aime et qui reste mon équilibre depuis la première année de mon doctorat. En effet, comme tu m'as donné deux magnifiques enfants, je vais finir cette partie par l'amour et le soutien que vous me donnez, Elise, Titouan et Pauline. Je vous aime tant.

Table des Matières

Introduction générale……11

PARTIE I : Développement, mise en place et évaluation du couplage CE-MS

I - Etude du couplage CESI-MS et application en tant que plateforme d'infusion nanoESI pour l'étude de complexes non-covalents biologiques 26

1. Introduction 26
2. Novel sheathless CE-MS interface as an original and powerful infusion platform for nanoESI study: from intact proteins to high molecular mass noncovalent complexes 29
3. Additional discussions 47
4. Conclusions 49

II - Etude comparative des performances des couplages CESI-MS et NanoLC-MS sur l'analyse protéomique d'un extrait mitochodrial de *Saccharomyces cerevisiae*. 51

1. Introduction 51
2. Improvement of mitochondria extract from *Saccharomyces cerevisiae* characterization in shotgun proteomics using sheathless capillary electrophoresis coupled to tandem mass spectrometry 54
3. Conclusion 76

III - Développement et évaluation du couplage CE-UV/MALDI-MS pour l'analyse de protéines intactes et d'applications protéomiques. 78

1. Introduction 78
2. Analysis of monoclonal antibody by a novel CE-UV/MALDI-MS interface 81
3. Conclusion 100

PARTIE II : Caractérisation de la structure primaire d'anticorps monoclonaux et évaluation de biosimilarité à l'aide du couplage CESI-MS

I - Caractérisation rapide et multi-niveaux d'anticorps monoclonaux par CESI-MS/MS 102

1. Introduction 102
2. Rapid and multi-level characterization of trastuzumab using sheathless capillary electrophoresis-tandem mass spectrometry 105

3. Conclusion 124

II - Caractérisation complète de la structure primaire et des micro-hétérogénéités d'anticorps en une seule injection par isotachophorèse et CESI-MS/MS 126

1. Introduction 126

2. Full antibody primary structure and micro-variant characterization in a single injection using transient isotachophoresis and sheathless CE-MS 129

3. Additional discussion regarding characterization of mAbs by t-ITP CESI-MS/MS 144

4. Conclusion 149

III - Evaluation de la biosimilarité d'anticorps monoclonaux par isotachophorèse transitoire et CESI-MS/MS 152

1. Introduction 152

2. Monoclonal antibodies biosimilarity assessment using transient isotachophoresis capillary zone electrophoresis-tandem mass spectrometry 156

3. Conclusion 176

PARTIE III : Caractérisation d'anticorps monoclonaux par approche middle-up à l'aide du couplage indirect CE-MS

I - Séparation et caractérisation de glycoformes de Cetuximab par Off-line CZE-UV/ESI-MS 182

1. Introduction 182

2. Glycoform Separation and Characterization of Cetuximab Variants by Middle-up Off-line CZE-UV/ESI-MS 184

3. Conclusion 203

II - Caractérisation de dimères de fragments $F_{c/2}$ de cetuximab par Off-line CZE-UV/ESI-MS ……..204

1. Introduction 204

2. Characterization of Cetuximab $F_{c/2}$ aggregates by Off-line CZE-MS 206

3. Conclusion 223

PARTIE IV : Caractérisation structural d'anticorps conjugués par CESI-MS à différents niveaux : intact, middle-up et peptide

I - Caractérisation à différents niveaux d'ADCs par la combinaison d'approches MS native, middle-up et bottom-up protéomique à l'aide du couplage CESI-MS 226

1. Introduction 226

2. Structural characterization of antibody drug conjugate by a combination of intact, middle-up and bottom-up techniques using sheathless capillary electrophoresis – tandem mass spectrometry as nanoESI infusion platform and separation method..... 229

3. Conclusion 248

Projets de Recherche……245

Liste des produtions scientifiques……248

Liste des abréviations

AA : acide aminé

ADN : acide désoxyribonucléique

ARN : acide ribonucléique

BGE : électrolyte support

CDR : Complementarity determining region

CE : électrophorèse capillaire

CEC : électrochromatographie capillaire

CE-ESI-MS : électrophorèse capillaire couplée à la spectrométrie de masse à source ESI

CID : dissociation induite par collision

cIEF : focalisation isoélectrique capillaire

CL : chromatographie liquide

CZE : électrophorèse capillaire de zone

Da : dalton

EACA : ε-aminocaproic acid

ECD : dissociation par capture d'électrons

EMA : European medicines agency

EOF : flux électroosmotique

ESI : électrospray

ESI-MS : spectrométrie de masse à source d'ionisation électrospray

ETD : Dissociation par transfert d'électrons

Fab : Fragment antigen binding

Fc : Fraction crystallizable

FDA : US food and drug administration

GCE : électrophorèse capillaire en gel

GlcNac : N-acétlyglucosamine

H : hauteur de plateau théorique

HC : chaîne lourde

HPC : hydroxypropyl cellulose

HPLC : chromatographie liquide haute performance

HPMC : Hydroxypropylmethyl cellulose

IEC : chromatographie à échanges d'ions

Ig : immunoglobuline

IgG : immunoglobuline G

ISD : dissociation en source

IT : analyseur de masse de type piège à ions

ITP : isotachophorèse

t-ITP : isotachophorèse transitoire

LC : chaîne légère

LC-ESI-MS : chromatographie liquide couplée à la spectrométrie de masse à source ESI

mAbs : anticorps monoclonaux

MALDI : matrix assisted laser desorption ionization

Man : mannose

MCP : microchannel plate

MEKC : électrochromatographie capillaire en phase micellaire

MS : spectrométrie de masse

MS/MS : spectrométrie de masse en tandem

m/z : rapport masse sur charge

N : nombre de plateaux théoriques

nanoESI : nano électrospray

nanoLC : nano chromatographie liquide

PB : polybren

PEG : polyéthylène glycol

PD : pharmacodynamique

PK : pharmacocinétique

PMF : approche par empreinte massique (peptide mass fingerprinting)

PFF : approche par empreinte massique des fragments (peptide fragment fingerprinting)

PTM : modification post-traductionnelle

Q : quadripôle

R&D : recherche et développement

RP-LC : chromatographie liquide en phase inverse

SEC : chromatographie à exclusion stérique

TIC : courant d'ion total

t-ITP : transient isotachophoresis

TOF : temps de vol

Tris : thrihydroxymethylamino methane

Uma : unité de masse atomique

UPLC : chromatographie liquide ultra-haute performance

Introduction générale

Les dernières décennies ont vu l'avènement de la spectrométrie de masse (MS) comme méthode analytique de référence. Cette technique qui a permis de repousser les limites de la chimie analytique en termes de caractérisation structurale et de sensibilité, est actuellement appliquée dans un grand nombre de champs d'applications en chimie et en biologie. Ainsi, la MS a démontré être particulièrement adaptée pour l'étude d'échantillons biologiques (peptides, protéines, complexe macromoléculaire) grâce au développement des sources d'ionisations douces telles que l'électrospray (ESI) et la désorption/ionisation laser assistée par matrice (MALDI). Le développement de ces deux sources ont conduit à l'attribution du prix Nobel de Chimie 2002 pour John B. Fenn et Koichi Tanaka, attestant la puissance de ces avancées dans le domaine de la chimie analytique et leur impact croissant dans de nombreux champs de recherche scientifique. En effet, la MS a permis l'émergence de nouveaux espaces de recherche tel que la protéomique, la métabolomique, la glycomique et plus récemment la lipidomique. L'amélioration constante des performances instrumentales en termes de précision de mesure de masse, de résolution, de gamme de masse ou de sensibilité, autorise l'investigation de problématique de plus en plus complexe (nature des échantillons, analyse de traces), et permet la caractérisation structurale fine de molécule allant jusqu'à la détermination de la formule brute. Cependant, malgré la puissance de ce système de détection, la MS est le plus généralement couplée à une méthode séparative en veine liquide telle que la chromatographie liquide à haute performance (HPLC). Le développement du couplage HPLC-MS, aussi noté LC-MS, a permis d'améliorer l'analyse de mélanges protéiques de plus en plus complexes notamment en permettant d'obtenir des informations de séquences et de structures de plus en plus précises. Techniquement, le couplage LC-MS est le plus communément utilisé car il permet d'obtenir des analyses reproductibles avec une bonne résolution. De plus, la facilité de mise en place du couplage avec un collecteur de fraction pour un couplage MS à source MALDI ou directement dans la source pour un couplage MS à source ESI est un facteur majeur du bon développement de ce système. La LC-MS est aujourd'hui la méthode de référence pour la caractérisation de mélange protéique. Cependant, une méthode ne permet pas de répondre au besoin infini d'une telle caractérisation.

L'électrophorèse capillaire (CE) est une méthode séparative développée à la fin du XX^{ème} siècle par le Prof. Hjerten à l'université d'Uppsala (Suède). Cette méthode permet la séparation d'analytes sous l'effet d'un champ électrique dans un capillaire de faible diamètre interne. La CE est une méthode miniaturisée offrant une excellente efficacité de séparation en un temps réduit. De plus, l'électrophorèse, principalement mis en place sous la forme de gel, est un

phénomène physico-chimique bien adapté pour la séparation d'analytes biologiques tels que les protéines, les peptides, l'ADN et l'ARN, et peut être utilisé dans différents modes électrophorétiques. Cependant, les propriétés intrinsèques de la CE comme des électrolytes supports (BGE) de fortes forces ioniques ainsi que la nécessité de maintenir le courant tout au long de la séparation, font du couplage CE-MS, un système compliqué à mettre en place. En effet, la MS est très sensible à la présence de sels qui peut provoquer des phénomènes de suppressions d'ions. De plus, contrairement à la LC-MS, la séparation en CE-MS doit impérativement conserver le courant électrique ce qui complique la géométrie de l'interface entre les deux technologies. Malgré ces problèmes techniques, le couplage de la CE avec la MS apparait aujourd'hui comme un système très attractif permettant de grandes avancées dans de nombreux domaines de recherche tels que la protéomique ou l'analyse de protéines entières.

La première interface permettant le couplage de la CE et de la MS a été développée à la fin des années 1980 par l'équipe du Prof. Smith. Cette interface a été conçue pour être compatible avec une source ESI. En effet, dû à son caractère direct, c'est-à-dire intégré à la source MS, le couplage CE-ESI-MS a été fortement étudié. La source ESI permet d'ioniser en phase gazeuse les composés chargés ou polaires présents dans un milieu liquide, comme le BGE. De ce fait, l'ESI est parfaitement adaptée pour le couplage avec la CE et s'impose tout naturellement comme la source d'ionisation la plus employée pour ce type d'association. Afin de permettre le maintien du courant tout au long de la séparation, trois types d'interfaces basées sur des principes différents ont été décrits : l'interface coaxiale à liquide additionnel dite « Sheath liquid », l'interface à jonction liquide dite « Junction liquid » et l'interface sans liquide additionnel dite « Sheathless ». Toutefois, la majorité des interfaces développées pour le couplage CE-ESI-MS ont tendance à souffrir d'un manque de robustesse et pour la plupart d'entre elles impliquent une réduction de la sensibilité. La baisse de sensibilité due à la géométrie de l'interface CE-ESI-MS représente la perte des performances majeures de cette méthode séparative pemettant d'obtenir de vrai nano-débits couplée à la sensibilité optimale de la MS.

Récemment, une interface innovante CE-ESI-MS a été développée dans le but d'aller au-delà des limites précédemment décrites. Ce développement représente la dernière génération d'interfaces sheathless CE-ESI-MS ce qui signifie qu'aucun liquide additionnel n'est nécessaire pour maintenir le contact électrique entre les électrodes de la CE et ainsi effectuer la séparation électrophorétique. Cette interface, notée CESI-MS dans la suite de ce manuscrit, basée sur l'idée originale du Prof. Moini (Smithsonian Institution, Suitland, MD), est aujourd'hui commercialisée par la société Sciex (anciennement Beckman Coulter). Le travail présenté

dans ce manuscrit sera en partie focalisé sur la mise en place, l'évaluation et l'application à la caractérisation de protéines thérapeutiques du couplage CESI-MS au sein du laboratoire.

Malgré les avancées technologiques importantes réalisées sur le couplage CESI-MS, l'une des limites majeures de ce couplage est que la nature du BGE (présence de sel, volatilité) peut affecter fortement la stabilité du spray et donc limiter le nombre de milieux de séparations compatibles avec l'ESI-MS. La source MALDI est une technologie MS qui présente une meilleure tolérance en présence de sels ou de détergents. Le couplage CE-MALDI-MS représente donc une alternative intéressante au couplage CE-ESI-MS. Tout comme le couplage CE-ESI-MS, une grande variété d'interfaces ont été développées depuis la fin des années 1980. Concernant les procédés permettant de maintenir le courant tout au long de la séparation, la quasi-totalité des interfaces reprennent les principes physiques du couplage de la CE avec une source ESI, à savoir l'utilisation d'un liquide additionnel coaxial, la jonction liquide ou l'utilisation d'un capillaire conducteur. En ce qui concerne la réalisation de dépôt solide, principe fondamental du MALDI, cet impératif a conduit au développement de deux types de couplages distincts : le couplage direct, avec un dépôt à l'intérieur de la source, suivi d'une étape de cristallisation intégrée et le couplage indirect, avec une étape de dépôt sur cible hors de la source et une étape de passage manuel ou automatisé de la cible dans la source. Concernant le couplage indirect, cette stratégie, plus facile à mettre en œuvre, ne fait pas intervenir de modifications instrumentales sur l'instrument MALDI-MS. De plus, le caractère indirect ouvre de nouvelles voies dans le domaine du traitement de l'échantillon. En effet, le fait de fractionner l'échantillon permet de le retraiter (enrichissement, digestion enzymatique, dérivation chimique…) indépendamment avant analyse en MALDI ou même de le déposer sur différents supports : plaque PCR, capillaire creux, membrane greffée ou cible MALDI. Cependant, malgré le fait que la quasi-totalité des interfaces CE-MALDI-MS soit efficace et robuste, aucune d'entre elles n'est automatisable et commercialisée ce qui rend ce couplage confidentiel par rapport au couplage LC-MS. Malgré cela, un grand nombre de publications ont été recensées montrant l'intérêt grandissant pour ce couplage, particulièrement pour l'étude de protéines entières ou digérées. Lors de mon installation au LSMIS, nous nous sommes intéressés à la mise en place au laboratoire d'un couplage indirect CE-MALDI-MS totalement automatisable dans le but de développer de nouvelles méthodologies analytiques permettant la caractérisation fine de protéines entières.

Donc, dans le but de compléter l'expertise du laboratoire en terme d'universalité du couplage CE-MS avec les sources ESI et MALDI, le travail présenté dans ce manuscrit sera donc focalisé sur la mise en place, l'évaluation et l'application à la caractérisation de protéines thérapeutiques des couplages CESI-MS et CE-MALDI-MS au sein du laboratoire.

Ce manuscrit sera divisé en trois parties distinctes :

- **La première partie de ce manuscrit (PARTIE I) sera consacrée à la mise en place et à l'évaluation des couplages CESI-MS et CE-MALDI-MS**.

Les travaux se sont tout d'abord concentrés sur l'étude des caractéristiques de l'ESI formé par l'intermédiaire du système CESI-MS : débits d'infusions accessibles, influence du débit d'infusion sur la sensibilité en particulier, et de comparer les performances d'ionisation avec celle d'une source ESI conventionnelle. Dans le cadre de cette comparaison, l'étude d'un complexe protéine-ligand a été réalisée en conditions natives. Les résultats obtenus ont ainsi permis de déterminer les différentes stœchiométries d'un complexe formé entre l'anneau beta et le peptide P14. La comparaison des spectres MS en infusion ESI conventionnelle et infusion CESI-MS du même échantillon a permis de mettre en évidence une intensité du signal environ 4 fois plus importante dans le cas de l'infusion CESI-MS. Enfin le système CESI-MS a été mis en œuvre afin de réaliser l'étude de complexes non-covalents de haut poids moléculaire : les hémocyanines chez les arthropodes et les mollusques. Elles forment des complexes multi-hexamèriques dont chaque sous-unité possède une masse aux alentours de 75 kDa. Les résultats obtenus ont montré la capacité du système CESI-MS à maintenir des interactions non-covalentes qui structurent ces complexes en réduisant par un facteur 400 la quantité d'échantillon injectée.

La suite de ce travail sur le couplage CESI-MS a consisté à réaliser l'évaluation d'un protocole expérimental combinant la CESI-MS/MS ainsi que la NanoLC-ESI-MS/MS afin d'explorer le protéome mitochondrial de *Saccharomyces cerevisiae*. La levure de *Saccharomyces cerevisiae* est souvent utilisée comme système modèle dans le développement et la validation de nouvelles approches instrumentales. Cette étude a été réalisée sur une faible quantité d'échantillon (100 ng) représentant environ 1000 protéines. Les résultats d'identification obtenus en CESI-MS/MS et en NanoLC-ESI-MS/MS ont été comparés. L'évaluation de chaque méthode séparative a été analysée suivant différentes propriétés physico-chimiques des peptides identifiés : le poids moléculaire, le ratio charge sur masse, le point isoélectrique, la couverture de séquence ainsi que la qualité spectrale MS/MS. Une analyse de la localisation cellulaire de toutes les protéines identifiées a également été effectuée. De plus, l'accent a été mis sur l'un des grands défis de l'analyse protéomique qui consiste à caractériser les isoformes. Les résultats présentés ont permis d'évaluer le gain d'information pouvant être obtenu par ces deux méthodes orthogonales et complémentaires en termes de couverture de séquence et de caractérisation d'isoformes.

Enfin, un chapitre de cette partie sera consacré au développement et à l'évaluation du premier couplage CE-UV/MALDI-MS/MS avec système de dépôt de matrice totalement automatisé et

mis en place au laboratoire. Plusieurs développements instrumentaux ont été réalisés pour rendre ce couplage robuste, autonome et polyvalent. Pour ce faire, nous nous sommes basés sur la modification de deux systèmes commerciaux. Afin de contrôler la séparation électrophorétique, la cellule UV originelle de l'appareil de CE a été déportée pour permettre simultanément la détection UV et la collection de fraction. De plus, une modification de la cartouche CE a été optimisée afin de rendre le couplage universel en termes de mode de dépôt MALDI. Ainsi, la séparation, le suivi UV, la collecte de fraction et l'adjonction de matrice sont automatisés. Dans un premier temps, les performances du système ont été évaluées avec la séparation d'un mélange de 5 protéines entières. Les résultats obtenus ont montré la haute résolution et la grande répétabilité de séparation conservés en utilisant le couplage CE-UV/MALDI-MS. Dans un second temps, le système a été évalué sur un digestat trypsique de 9 protéines. Une comparaison avec le couplage référence HPLC-MALDI-MS a été réalisée sur l'approche de peptide mapping. Cela a permis de montrer une très grande complémentarité entre les deux approches. Enfin, dans le but de confirmer les performances du système pour la séparation de protéines entières de haut poids moléculaire, l'application à la séparation de mAbs intacts et digérés a été réalisée à l'aide du couplage CE-UV/MALDI-MS. Ce travail représente la première analyse de mAb intactes à l'aide d'un couplage de la CE avec une détection MS.

- **La deuxième partie de ce manuscrit (PARTIE II) sera consacrée à la caractérisation de la structure primaire d'anticorps monoclonaux et évaluation de biosimilarité à l'aide du couplage CESI-MS.**

Depuis 1986 et l'approbation du murumonab-CD3 par l'administration de régulation des aliments et des médicaments américaine (FDA), les anticorps monoclonaux (mAbs) représentent la classe d'agent thérapeutique qui enregistre la croissance la plus importante. Les mAbs du fait de leur spécificité pour l'antigène associé ouvrent de nouvelles perspectives en termes de traitement thérapeutique, de plus certaines de leurs propriétés sont favorables à l'application pour le traitement de pathologies (réductions des effets secondaires, pharmacocinétique et pharmacodynamique favorables). A l'heure actuelle, une quarantaine de mAbs ont été approuvés comme traitement thérapeutique dans le monde et 30 autres mAbs sont en Phase 3 d'essai clinique. Les champs d'application de ce type de protéines sont principalement l'oncologie, ils sont également utilisés dans le traitement de maladies inflammatoires, de maladies auto-immune ou plus récemment pour le traitement de la maladie d'Alzheimer. Les brevets protégeant la première génération de mAbs, dont plusieurs *blockbusters* (traitements générant plus d'un milliard de dollars de chiffre d'affaires par an),

tomberont dans les prochains mois dans le domaine public. La fin de cette protection offre l'opportunité à des entreprises tierces de produire des versions « bio-générique » de ces mAbs. Ces copies portent le nom de biosimilaires.

Les mAbs sont des glycoprotéines complexes potentiellement porteuses de nombreuses micro-hétérogénéités moléculaires dont les sources sont variées (procédés de production inadaptée, conservation prolongée). La complexité de ces protéines et leur utilisation en tant que traitement thérapeutique justifient la nécessité d'améliorer les méthodologies analytiques de caractérisation.

Ces travaux décrivent la mise en œuvre et le développement d'une méthode analytique de CZE-ESI-MS/MS, utilisant le système CESI-MS, permettant la caractérisation fine de la structure primaire de plusieurs anticorps monoclonaux (mAbs) dont notamment trastuzumab (Herceptin®, Roche) et cetuximab (Erbitux®, Merck KgaA). Les échantillons ont tout d'abord subis une digestion trypsique. Le protocole de digestion a été optimisé afin d'améliorer la compatibilité entre l'échantillon et la méthode d'analyse par CESI-MS/MS, ce qui a pour effet d'améliorer significativement le niveau de caractérisation. Le digestat de peptides a ensuite été séparée et caractérisée par isotachophorèse transitoire (t-ITP) et CESI-MS/MS. Les résultats obtenus ont montré la capacité de la méthodologie développée à obtenir systématiquement 100% de recouvrement de séquence en une seule injection pour chacun des mAbs étudiés. Les données de t-ITP CESI-MS/MS ont également été utilisées pour caractériser les glycosylations portées par ces différents mAbs. Les performances de la méthode développée, au regard de la caractérisation des glycosylations, sont directement reliées avec la capacité de la CE à séparer les différents glycopeptides en fonction du glycanne qu'ils portent. Enfin les PTMs *hot-spots* ont également pu être caractérisés pour chacun des mAbs. Dans chacun des cas, l'ensemble des PTMs *hot-spots* portés par la protéine ont pu être caractérisés. L'utilisation de la CE en tant que méthode séparative en amont de l'analyse MS a démontré sa pertinence en particulier dans le cadre de la caractérisation des PTMs. En effet les résultats obtenus ont montré la faculté de la méthodologie développée à séparer complètement un peptide modifié de son homologue resté intact. Ce type de sélectivité, délivrée par l'utilisation de la séparation électrophorétique, a pu être observée même dans le contexte de modification impactant relativement peu le peptide tel que la déamidation d'une asparagine (+0.98 Da) ou encore l'isomérisation de l'acide aspartique (aucune différence de masse).

Concernant l'évaluation de la biosimilarité des protéines thérapeutiques, les résultats obtenus sur deux couples d'échantillons (trastuzumab/biosimilaire, cetuximab/biosimilaire) pour la caractérisation de la séquence d'acides aminés ont confirmé pour chacun des mAbs analysés

la possibilité de caractériser la totalité de la séquence en une seule analyse. Les données CESI-MS/MS ont permis dans le cadre de l'étude trastuzumab/biosimilaire de mettre en évidence une différence d'un unique acide aminé entre les deux protéines. Les données ont ainsi permis de caractériser rapidement la présence d'une différence majeure au niveau de la séquence d'acides aminés du candidat biosimilaire, mais également de déterminer l'acide aminé en cette position dans le cas du candidat. Dans le cadre de l'étude de la biosimilarité cetuximab/biosimilaire, les données CESI-MS/MS ont permis de démontrer sans ambiguïté une complète similitude entre cetuximab et le candidat du point de vue de la séquence d'acides aminés. Concernant les différentes glycosylations portées par les mAbs étudiés. Les glycannes ont pu être caractérisés du point de vue de leur structure et les abondances relatives ont été estimées, afin de comparer les profils d'expression des glycosylations entre les mAbs originaux et leur candidat biosimilaire. Les comparaisons des profils de glycosylation entre les mAbs originaux et leur candidat biosimilaire démontrent la capacité de la méthode pour caractériser des différences mineures entre les profils de glycosylation. Pour cetuximab, les deux sites de glycosylations ont pu être caractérisés de manière totalement indépendante. Les données de CESI-MS/MS montrent des différences d'abondances entre cetuximab et son candidat biosimilaire. De plus, les données recueillies ont servi pour prouver qu'une partie des glycosylations exprimées par le candidat biosimilaire n'étaient pas présentes chez cetuximab, ce qui a abouti à l'impossibilité de considérer ces deux protéines comme étant biosimilaires. Les résultats de la caractérisation réalisée par t-ITP CESI-MS/MS ont permis de conclure sur la biosimilarité de chacun des candidats aussi bien dans le cas de trastuzumab ou cetuximab.

- **La troisième partie de ce manuscrit (PARTIE III) sera consacrée à la caractérisation d'anticorps monoclonaux par approche middle-up à l'aide du couplage indirect CE-MS.**

Tout comme dans la partie décrite précédemment, la caractérisation de plus en plus approfondie de mAbs à l'aide du couplage CE-MS, sera au cœur de cette partie. En effet, les méthodes analytiques dans leur ensemble doivent être capables de couvrir un large champ de la structure des mAbs : séquence d'acides aminés, glycosylation, modifications post-traductionnelles (PTMs), ponts disulfures. Concernant plus particulièrement les glycosylations qui sont connues pour être une classe de modification importante pouvant impacter les propriétés immunogéniques des mAbs, leur caractérisation est réalisée suivant différentes stratégies telles que la mesure de masse intacte, l'analyse middle-up des différents domaines des mAbs ou l'analyse protéomique. Cependant, aujourd'hui, aucune séparation de glycoformes de mAbs n'a encore été décrite à l'aide du couplage de référence LC-MS. La

partie II de ce manuscrit montre un réel intérêt du couplage CESI-MS pour la caractérisation de mAbs. Cependant, due à la présence de sel en forte concentration et à la non-volatilité des électrolytes support (BGE) nécessaire pour séparer des variants de charges de mAbs, le développement du couplage de la CZE avec la MS est très limité pour l'analyse de protéines entières. Faisant suite à l'étude précédemment introduite sur le développement du couplage CZE-UV/MALDI-MS (III - Développement et évaluation du couplage CE-UV/MALDI-MS pour l'analyse de protéines intactes et d'applications protéomiques) et dans le but de pallier le manque de résolution du MALDI-MS, nous avons développé une stratégie pour réaliser un couplage indirect CZE-UV/ESI-MS basée sur l'utilisation de notre robot CZE-UV/collection de fraction suivie d'une infusion des différentes fractions par CESI-MS utilisé en tant que plateforme nanoESI. L'un des objectifs de cette étude a été de rendre le BGE compatible avec la détection ESI-MS. Pour cela, des conditions asymétriques mettant en jeu différents BGE aux extrémités du capillaire ont été développées. Le cetuximab a pour particularité de contenir deux sites de glycosylation répartis sur la chaine lourde (HC) : un localisé sur le domaine Fc/2 (Asn^{299}) et le second localisé sur le domaine F(ab')2 (Asn^{88}). Cela représente un échantillon de choix pour notre étude qui a permis de caractériser ces deux domaines séparément à l'aide d'une stratégie middle-up par digestion IdeS sans réduction. Les performances de notre stratégie ont été évaluées sur la caractérisation et la séparation des variants de masses dus à la perte d'une lysine en C-terminal de la HC et sur les variants de charges de glycoformes du cetuximab. De plus, nous avons cherché à comprendre la nature de la séparation des variants de charges correspondant au domaine Fc/2. En effet, l'étude MS de la séparation électrophorétique montrant 3 pics de variant de Fc/2, a permis de caractériser deux variants de masse basés sur la perte d'une lysine en C-terminal de la HC du cetuximab. Cependant la présence d'un troisième pic correspondant à un mélange de ces deux variants de masse implique plusieurs questions sur l'origine de ce phénomène : diffusion lors du dépôt, rôle des glycosylations, rôle de la perte de lysine ou formation d'agrégats. Les résultats détaillés dans cette partie permettront donc d'apporter toutes les réponses à ces questions, notamment sur la formation de dimères de Fc/2.

- **La dernière partie de ce manuscrit (PARTIE IV) sera consacrée à la caractérisation structural d'anticorps conjugués par CESI-MS à différents niveaux : intact, middle-up et peptide.**

Les anticorps conjugués (ADCs) sont une catégorie émergeante de molécules biothérapeutiques dérivés des anticorps monoclonaux (mAbs). Ils combinent la forte activité anti-cancereuse des médicaments cytotoxiques ainsi que la sélectivité et les propriétés

pharmacologiques favorables des mAbs. La synthèse des ADCs est obtenue par une liaison chimique entre le médicament et le mAbs. Aujourd'hui, deux ADCs (brentuximab vedotin et trastuzumab emtansine) ont été approuvés et mis sur le marché par la FDA, et plus de 50 candidats sont actuellement en cours d'essais cliniques ce qui montre l'intérêt croissant de l'industrie biopharmaceutique pour ces composés. Le principe d'un ADCs est basé sur l'utilisation de la forte spécificité du mAbs pour cibler un type de cellules tumorales et ainsi permettre par relargage, l'application de médicaments extrêmement puissants. Cela limite l'exposition des cellules saines et donc permet d'optimiser les traitements notamment avec la diminution des effets secondaires.

L'hétérogénéité des ADCs, générée par la variation du nombre et la position de médicaments conjugués sur le mAbs durant sa production, implique une réponse immunitaire différente suivant la valeur du DAR. Tout comme les mAbs, les ADC sont des glycoprotéines complexes potentiellement porteuses de nombreuses micro-hétérogénéités moléculaires dont les sources sont variées (procédés de production inadaptée, conservation prolongée). La complexité de ces protéines et leur utilisation en tant que traitement thérapeutique justifient la nécessité d'améliorer les méthodologies analytiques de caractérisation.

Dans cette partie, nous avons donc appliqué une approche globale mettant en œuvre l'utilisation du couplage CESI-MS en tant que plateforme d'infusion nanoESI et méthode séparative pour obtenir la caractérisation structurale à différents niveaux d'un ADCs. L'ADCs sélectionné comme modèle standard pour cette étude est le brentuximab vedotin (BV) qui est l'un des deux ADCs approuvés par la FDA (2011) pour le traitement du lymphome d'Hodgkin.

Le premier niveau de caractérisation décrit dans cette partie consiste à évaluer le couplage CESI-MS en tant que plateforme nanoESI en MS native pour la détermination du DAR moyen du BV et pour obtenir la distribution de conjugués sur le mAbs. Ensuite, dans le but d'approfondir la caractérisation du BV en utilisant le même système, nous avons réalisé une approche middle-up par digestion IdeS. La réaction enzymatique de l'ADCs avec l'IdeS suivi de son passage en MS permet d'obtenir deux types de fragments distincts (Fc/2 et F(ab')2) ainsi que de nombreuses informations de structure. Enfin, une méthode CESI-MS basée sur une approche bottom-up protéomique a été développée dans le but d'obtenir la structure primaire de l'ADCs ainsi que les peptides liés aux médicaments. En effet, due aux propriétés d'hydrophobicités du médicament, nous avons dans un premier temps optimisé les conditions du protocole de digestion trypsique. Puis, à l'aide du couplage CESI-MS, nous avons réalisé la caractérisation simultanée de la séquence en acide aminé de l'ADCs ainsi que son profil de glycosylation et l'identification des peptides conjugués.

PARTIE I : Développement, mise en place et évaluation du couplage CE-MS

I - Etude du couplage CESI-MS et application en tant que plateforme d'infusion nanoESI pour l'étude de complexes non-covalents biologiques

1. Introduction

L'étude des structures de protéines maintenues sous leur forme native ainsi que des complexes protéiques permet la compréhension des mécanismes biochimiques. Les interactions non-covalentes jouent un rôle prépondérant dans un grand nombre de processus biochimiques, en effet la formation de complexes protéiques est généralement à l'origine de la spécificité de nombre de mécanismes de reconnaissance, liaison ou transport. La spectrométrie de masse (MS), notamment avec les développements des sources d'ionisation à électrospray (ESI) s'est affirmée comme une technique de choix dans le cadre d'études de complexes non-covalents. L'ESI en tant que source d'ionisation douce permet le transfert en phase gazeuse des complexes protéiques tout en maintenant les interactions non-covalentes.

Afin de maintenir les protéines et les complexes non-covalents sous leur forme native, il est indispensable d'utiliser des tampons aqueux ne perturbant pas la structure quaternaire des protéines et dont la volatilité est suffisante pour être compatible avec l'ionisation par ESI : l'acétate d'ammonium est de ce fait l'un des tampons les plus couramment utilisé, les concentrations pouvant variées de 20 mM à 1 M selon les cas. L'utilisation de telles concentrations n'est pas sans conséquence et l'ESI-MS native souffre généralement d'un défaut de l'intensité du signal et d'une sensibilité réduite en raison des effets de suppression d'ions dus aux sels. L'une des alternatives généralement utilisée pour augmenter l'intensité du signal est d'utiliser des échantillons dont la concentration en analytes est relativement élevée, aux risques de favoriser des interactions non spécifiques ou l'agrégation de protéines. De plus, le volume d'échantillon nécessaire pour réaliser une acquisition est relativement important lors de l'utilisation d'une source ESI conventionnelle (environ 100 µL). L'utilisation de quantités importantes d'échantillon peut s'avérer être un frein notamment dans le cadre d'étude d'échantillons biologiques dont la disponibilité est limitée. Il apparait donc nécessaire de pouvoir améliorer le rapport signal/bruit tout en diminuant la quantité d'échantillon consommée par acquisition afin de pouvoir profiter au maximum des performances de la MS pour l'étude de ces complexes.

Dans le but de pallier à ces limitations, Wilm et Mann ont introduit en 1994 une source dite nano électrospray (nanoESI). En nanoESI, les échantillons sont injectés dans la source par l'intermédiaire d'un capillaire dont le diamètre interne est faible (de 1 à 5 µm), il est ainsi

possible d'infuser l'échantillon dans le spectromètre de masse à de faibles débits diminuant ainsi la quantité d'échantillon nécessaire pour réaliser une injection. L'utilisation d'une source nanoESI favorise la formation de gouttelettes de diamètre plus faibles par rapport à l'ESI classique. La formation de gouttelettes plus petites favorise le processus d'ionisation car elles facilitent les divisions de Rayleigh durant le processus de désolvatation. La conséquence directe au niveau des spectres MS est une sensibilité du signal améliorée et une réduction du phénomène de suppression d'ions.

L'objectif de cette étude est d'utiliser le système CESI-MS en tant que plateforme semi-automatisée d'infusion nanoESI et de démontrer la compatibilité de cette approche pour l'étude de protéines en conditions natives et de complexes non-covalents par MS native. Il ne s'agit pas dans ce cas d'utiliser l'électrophorèse capillaire (CE) en tant que méthode séparative mais d'exploiter les caractéristiques intrinsèques du système CESI-MS pour produire de manière robuste et parfaitement contrôlée un nanoESI : diamètre interne du capillaire faible (30 µm), possibilité de réaliser des injections d'échantillon contrôler à des débits de l'ordre du nL/min grâce au système d'injection équipant la CE.

Les travaux se sont tout d'abord concentrés sur l'étude des caractéristiques de l'ESI formé par l'intermédiaire du système CESI-MS : débits d'infusions accessibles, influence du débit d'infusion sur la sensibilité en particulier, et de comparer les performances d'ionisation avec celle d'une source ESI conventionnelle. Les résultats obtenus sur des protéines modèles ont permis de mettre en évidence la capacité du système CESI-MS à maintenir un ESI stable et parfaitement contrôlé pour des débits d'infusions d'échantillon de 900 nL/min abaissés jusqu'à 4 nL/min. Les études de sensibilité, réalisées sur une protéine modèle en conditions natives, ont montré qu'une réduction du débit d'infusion permet d'augmenter de manière importante la sensibilité du signal MS : par l'intermédiaire du système CESI-MS, il a été possible d'améliorer la sensibilité 46 fois en réduisant le débit d'infusion de 457 à 14 nL/min. De plus le rapport signal/bruit s'est vu amélioré d'un facteur proche de 5 par la réduction du débit d'infusion. Ces résultats illustrent le gain en sensibilité pouvant être apporté à la MS native par ce type d'approche instrumental. L'étude du débit d'infusion a également mis en évidence que le débit appliqué par le système d'injection délivre une infusion parfaitement contrôlée et reproductible.

L'infusion réalisée par l'intermédiaire du système CESI-MS a ensuite été comparée à l'infusion en ESI classique. Dans le cadre de cette comparaison, l'étude d'une protéine a été réalisée en conditions natives. L'anneau beta est une protéine présente *in vivo* sous la forme d'un homodimère de 85 kDa. Elle a pour fonction d'améliorer la processivité entre l'ADN et l'ADN polymérase. En plus de l'étude de la protéine sous forme native, la formation du complexe entre l'anneau beta et un peptide synthétique (P14) a été aussi étudiée. Le peptide P14 a pour

particularité d'interagir en formant un complexe non-covalent avec l'anneau beta par l'intermédiaire du même site consensus que celui utilisé par l'ADN *in vivo*. Les résultats obtenus ont ainsi permis de déterminer les différentes stœchiométries du complexe formé entre l'anneau beta et le peptide P14. La comparaison des spectres MS en infusion ESI conventionnelle et infusion CESI-MS du même échantillon a permis de mettre en évidence une intensité du signal environ 4 fois plus importante dans le cas de l'infusion CESI-MS. De plus le débit d'injection utilisé pour l'infusion CESI-MS (50 nL/min) permet de diminuer de manière drastique la quantité d'échantillon nécessaire pour une acquisition : celle-ci a pu être réduite 80 fois en comparaison avec l'infusion ESI classique. Une réduction si importante de la quantité d'échantillon offre la possibilité de réaliser des acquisitions prolongées. Ces expérimentations ont ainsi permis de démontrer la compatibilité de ce type d'approche par l'intermédiaire du système CESI-MS avec l'étude de protéines en conditions natives mais également la faculté du système CESI-MS à maintenir les interactions non-covalentes que l'on observe dans les complexes protéine-protéine ou protéine-ligand.

Enfin le système CESI-MS a été mis en œuvre afin de réaliser l'étude de complexes non-covalents de haut poids moléculaire : les hémocyanines sont des protéines impliquées dans le transport d'oxygène chez les arthropodes et les mollusques, elles forment des complexes multi-hexamèriques dont chaque sous-unité possède une masse aux alentours de 75 kDa. Les résultats obtenus ont montré la capacité du système CESI-MS à maintenir des interactions non-covalentes qui structurent ces complexes. Il a ainsi été possible par l'étude d'échantillons d'hémolymphes de *carcinus maenas* d'observer les formes 6-mers, 12-mers et 18-mers d'hémocyanines dont les masses sont respectivement de 445 kDa, 900 kDa et 1.34 MDa. Les infusions ont pu être réalisées avec un débit de 10 nL/min soit un volume total de 75 nL, ce qui représente un volume d'échantillon réduit 400 fois par rapport à une acquisition en ESI classique équivalente.

Les résultats présentés dans cette partie ont fait l'objet d'une publication dans le journal scientifique *Analytical and Bioanalytical Chemistry (*Springer 2014).

2. Novel sheathless CE-MS interface as an original and powerful infusion platform for nanoESI study: from intact proteins to high molecular mass noncovalent complexes

Rabah Gahoual, Jean-Marc Busnel, Philippe Wolff, Yannis Nicolas François, Emmanuelle Leize-Wagner ***Analytical and Bioanalytical Chemistry*** 2014, 406(4), 1029-1038.

Abstract

Development of nano-electrospray (nanoESI) sources allowed to increase significantly the sensitivity which is often lacking when studying biological noncovalent assemblies. However, the flow rate used to infuse the sample into the mass spectrometer cannot be precisely controlled with nanoESI and the robustness of the system could represent an issue. In this study, we have used a sheathless capillary electrophoresis–mass spectrometry (CESI) prototype as a nanoESI infusion device. The hydrodynamic mobilization of the capillary content was characterized and the ability of the system to generate a stable electrospray under controlled flow rate conditions ranging from 4 up to 900 nL/min was demonstrated. The effect of the infusing flow rate on the detection of an intact model protein analyzed under native conditions was investigated. Results demonstrated a significant increase in sensitivity of 46-fold and a signal-to-noise ratio improvement of nearly 5-fold when using an infusing flow rate from 456.9 down to 13.7 nL/min. The CESI prototype was further used to detect successfully the β ring homodimer in its native conformation. Obtained results were compared with those achieved with conventional ESI. Intensity signals were increased by a factor of 5, while sample consumption decreased 80 times. β ring complexed with the P14 peptide was also studied. Finally, the CESI interface was used to observe the quaternary structure of native hemocyanins from Carcinus maenas crabs; this high molecular complex coexisting under various degrees of complexation and resulting in masses ranging from 445 kDa to 1.34 MDa.

Introduction

Structural determination of intact proteins and protein complexes has a major contribution to understanding fundamental biochemical pathways. Noncovalent interactions are decisive for numerous kinds of in vivo biochemical processes as assemblies of several proteins generally allow specific recognition, binding, or transport. Mass spectrometry (MS) since the development of electrospray ionization (ESI) has grown into a major technique to study noncovalent interactions [1–3]. Indeed

ESI is a soft ionization technique which allows transfer of noncovalent protein complexes to the gas phase without losing its quaternary structure [4]. MS is able to provide access to structural information [5], stoichiometry [6], and binding constant of noncovalent complexes [7]. To detect native complexes by MS, it is mandatory to use aqueous buffers which do not disrupt the quaternary structure of the protein or the complex. Additionally, used buffers need to be relatively volatile to enable ESI ionization/desorption processes. Ammonium acetate is one of the most commonly used buffers as it fulfills those requirements [8]. Salt concentrations necessary to maintain the complex in native conditions may vary from 20 to 100 mM or even higher. Such levels of salt concentration are quite significant and as a direct consequence, ESI-MS in native conditions usually results in poor signal intensity and sensitivity mainly due to ion suppression effect [9, 10]. A typical alternative to enhance signal intensity is to use a higher sample concentration (superior to 10 µM); however, such conditions risk to favor non-specific interactions and protein aggregation. The total volume of sample required to perform an experiment is also an important issue. In the context of the study of noncovalent complexes by classical ESI, the volume of sample to realize one acquisition is usually in the range of 100 µL. Such important volume may be unfavorable for studying biological samples which are only available in scarce amounts. Even if classical ESI has given the possibility to observe noncovalent edifices by MS, it was necessary to improve signal-to-noise ratio while concomitantly reducing sample consumption to realize the full potential of MS-based studies in this field.

To address both issues, Wilm and Mann introduced in 1994 a nano-electrospray ionization source (nanoESI) causing a major advance in MS [11].With nanoESI, the sample is introduced in the MS through a tapered capillary with reduced internal diameter (from 1 to 5 µm). Such device being compatible with low flow rates (typically a few dozens of nanoliters per minute), it allows to significantly lower the volumes of infused sample while maintaining a similar acquisition time. In nanoESI, smaller droplets are produced increasing sensitivity and reducing ion suppression by improving analyte ionization efficiency as they facilitate Rayleigh divisions throughout the desolvation process. Using nanoESI, they demonstrated the MS detection of a

synthetic peptide at a concentration of 1 µM in 100 mM NaCl [12]. Thanks to the significant reduction in flow rates as compared with classical ESI, sample consumption was reduced to at least 2 orders of magnitude in nanoESI. Typically, a nanoESI experiment is performed using a borosilicate glass capillary pulled to reduce the internal diameter of its extremity to ensure good compatibility with low flow rates. Capillary tip is then gold plated allowing grounding. Sample is introduced in the capillary and nanoESI is initiated by applying a potential between the capillary and the MS inlet [12]. Note that no pumping system is used to introduce the sample into the source, therefore infusion cannot be precisely controlled and flow rate robustness may be an issue. Although technical developments have been realized allowing automated nanoESI to be performed using chip-based automated infusion systems [13], the robustness of such system may be problematic especially in the case of samples containing relatively high ionic strength buffers typically used to maintain noncovalent complexes. Indeed, to the best of our knowledge, such automated systems have not been used yet to study high molecular mass noncovalent assemblies (superior to 250 kDa) [14]. That observation limits unfortunately the use of automated systems for high molecular mass and complex assemblies which have been subject to intensive research work in the recent years [5].

Recently, a novel interface (referred as sheathless capillary electrophoresis (CESI)) allowing the hyphenation of capillary electrophoresis (CE) with ESI-MS in a single dynamic process within the same device has been developed by Beckman Coulter. It is based on a design previously described by Moini [15]. This sheathless interface uses a fused-silica capillary whose outlet has been etched using hydrofluoric acid (HF) making it porous to the electrical transport of small ions. Detailed description of this interface has been given by Haselberg et al. [16]. It has already been used to perform successful CE-MS experiment in proteomics [17–23], metabolomics [24], and intact proteins [16, 25, 26] demonstrating drastically increased sensitivity compared with sheath-liquid CE-MS due to low operating flow rates of 10 to about 100 nL/min.

After having studied the suitability of the CESI prototype to be used as a nanoESI infusion device, the aim was to assess its potential as compared with either conventional ESI or classical approaches to nanoESI under very demanding conditions such as native MS of large noncovalent complexes. To this end, parameters such as ease of use, sensitivity gain, robustness, and sample consumption have been considered. Electromigration separation was not involved in this work properly but the intrinsic characteristics of this system (miniaturized aspect and low infusion flow rate) and the possibility to realize accurate hydrodynamic infusion were used to infuse the sample in the source and generate fully controlled nanoESI experiments. Then, a model protein, β ring, was studied in native conditions with both CESI and conventional ESI to determine the ability of the CESI interface to maintain noncovalent

complexes such as protein–protein and protein–ligand interactions. Finally, as a last step, a biological high molecular mass complex was successfully observed using the CESI prototype through the successful detection of the considered complex assembly, maintained by weak interactions having a molecular mass superior to 1.3 MDa; CESI was demonstrated to be well suited for such studies. The compatibility of the system with very low flow rates and its robustness enabled the achievement of producing high-quality mass spectra and accurate mass determination while consuming very limited amounts of samples.

Materials and methods

Reagents

Myoglobin and all chemicals used were of analytical grade or high purity grade and purchased from Sigma-Aldrich (Saint Louis, MO, USA). Water used to prepare buffers and sample solutions was obtained using an ELGA purelab UHQ PS water purification system (Bucks, UK). Desalting steps were performed using Amicon Ultra 10 K centrifugal filters purchase from Millipore (Molsheim, France). β ring samples were kindly provided by Dominique Burnouff group (IBMC, Strasbourg, France), Carcinus maenas hemolymph samples were supplied by Frank Zal group (Roscoff, France).

Production of the β ring and P14 peptide

The Pseudomonas dnaN gene was cloned into pET15b plasmid (Invitrogen) using standard protocols. The resulting Ntagged protein was expressed in BL21 Pseudomonas cells after isopropyl β-D-1-thiogalactopyranoside induction (0.1 mM) at 28 °C. The β protein fraction was first enriched on a Ni-NTA column, eluted with an histidine step (300 mM) and further purified on a MonoQ column in a buffer containing

20 mM Tris HCl (pH 7.5), 0.5 mM EDTA, and 10 % glycerol, using a gradient from 0 to 0.5 M NaCl. The pure protein was dialyzed and concentrated (around 300 µM) against a buffer containing 20 mM Tris HCl (pH 7.5), 0.5 mM EDTA, and 10 % glycerol. P14 peptide (Ac-Q Cha DL diClF, M=785 Da) was synthesized using a procedure described previously by Wolff et al. [27].

Hemocyanin sample preparation

Hemolymph sample collection and preparation were adapted from Sanglier et al. [6]. Prior to use, frozen C. maenas hemolymph samples were thawed at 4 °C and further centrifuged at 12,000 rpm during 15 min. The obtained supernatant was collected and considered for a further treatment. Collected supernatants were desalted using Amicon centrifugal filters (cut

off=10,000 Da) in 20 mM ammonium acetate (pH=6.8). Nine desalting steps were performed at 4 °C and 14,000×g for 20 min each step. Ammonium acetate allows the native structure of proteins to be preserved and be compatible to ESI-MS. CE as infusion platform. The infusion experiments were carried out with a PA 800 plus capillary electrophoresis (CE) system from Beckman Coulter (Brea, CA) equipped with a temperature controlled autosampler and a power supply able to deliver up to 30 kV. Prototype fused-silica capillaries (total length, 100 cm; 30 µm i.d.) whose outlet end (about 3 cm) was etched with HF were used for all related CESI experiments and initially provided by Beckman Coulter (Brea, CA, USA). It is certainly noteworthy to specify here that the inner lumen of these capillaries is not tapered and presents an i.d. of 30 µm throughout its entire length. New capillaries were initially conditioned by flushing them for 10 min with MeOH, 10 min with 0.1 M sodium hydroxide, 10 min with 0.1 M hydrochloric acid, water for 20 min, and finally for 10 min with background electrolyte (BGE) each flushing step being conducted at 75 psi (5.17 bar). BGE used ammonium acetate at 20 or 50mM depending on the sample at physiological pH (pH ranging from 6.7 to 7). Hydrodynamic injection provided by the CE pressure delivery system was used to infuse the sample through the CESI prototype.

MS and data analysis

For sheathless CE-ESI-MS experiments, the CE system was coupled to a microTOF-Q II (Bruker Daltonics, Bremen, Germany) mass spectrometer for general characterization and protein–ligand applications and to a microTOF (Bruker Daltonics, Bremen, Germany) mass spectrometer for analysis of the hemocyanin samples. MS transfer parameters were optimized using the actual sample directly infused via the CE system using a pressure of 5 psi (340 mbar). MS parameters were optimized so that heavy ions could be properly transferred to the TOF analyzer while avoiding fragmentation. In the case of the microTOF-Q MS, ion funnels were set at values of 300 and 400 Vpp. When using the microTOF MS, capillary exit was set at a value of 140 V. The electrospray voltage (capillary voltage) typically ranges from −1.2 to −1.8 kV depending on the composition of the buffer. No nebullizer gas is required because of the design of the CESI interface. Dry gas was set at 1.5 L/min and source temperature at 180 °C.

Evaluation of flow rate and infusion stability

While the capillary was initially filled with 10 % acetic acid, a sample of myoglobin at a concentration of 1 µM (16.95 ng/µL) in 10 % acetic acid was infused through the CE system using different pressure (from 20 to 0.4 psi). To estimate infusion flow rates, the time necessary for the sample to fill the entire capillary was measured following the MS signal. The experiment was repeated in triplicate for each pressure.

Evaluation of CESI sensitivity

Myoglobin sample at a concentration of 250 nM in 20 mM ammonium acetate (native conditions) was infused through the CE system using different pressure (from 100 to 0.4 psi). For each pressure, sample mass spectrum was recorded for 4 min. An average mass spectrum was then generated using the last 3 min (first minute was excluded to ensure flow rate signal stabilization). Signal intensity and sensitivity (amount of sample introduced to the MS) were plotted against infusion flow rate. The full experiment was replicated four times.

Results and discussion

Evaluation of CESI interface as a nanospray emitter. As a preliminary experiment, the CESI interface was studied regarding the range of infusion flow rate accessible while maintaining a stable electrospray. CE experiments are usually performed in a reduced volume of a few nanoliters because the total capillary volume did not exceed a few microliters. Therefore, CE systems are designed to accurately execute injection of small volumes of sample (<1 µL injected), the pressure system can also be used to infuse a sample solution through the capillary at very low flow rates in a stable and robust way. Hydrodynamic injection is made by applying a pressure over the contents of a vial located at the inlet of the capillary.While flow rates can theoretically be calculated using the Hagen–Poiseuille Eq. (1), it was important to experimentally measure the actual flow rates under the considered nanoESI conditions, such measurements allowing to take into account the viscosity of the sample solution and further the potential suction effect induced by the ESI process itself. Indeed the capillary tip is positioned at a distance of about 2 mm from the MS entrance; this proximity requires to determine if at low flow rates the potential applied in the source does not create a suction effect of the sample similar to those observed with borosilicate glass capillaries in nanoESI [12]. That effect could potentially disrupt the pressure applied by the CE making flow rate control impossible.

Myoglobin in 10 % acetic acid was used as a model protein to evaluate flow rates. Sample was injected through the capillary at a given pressure. The capillary, initially only filled with buffer, is progressively filled with the sample until the myoglobin is detected by the MS. The time necessary to record MS signal corresponding to myoglobin (detection time) is obtained and considered for flow rate calculation. Viscosity of the buffer was then calculated using the Hagen–Poiseuille Eq. (1). Flow rates were calculated from the detection time and the capillary total volume.

$$\eta = \frac{d^2{}_c \Delta P t}{32 L^2} \quad (1)$$

Where d_c is the capillary internal diameter, ΔP the applied pressure, t the detection time, L the total length of the capillary.

The measured flow rate were then plotted as a function of the pressure used for mobilization (Fig. 1). Results obtained show the capacity of the system to maintain a stable electrospray with a pressure as low as 0.4 psi corresponding to a flow rate of 3.8 nL/min. Moreover to the possibility to realize sample infusion at very low flow rates, the system has also demonstrated a significant robustness as the relative standard deviation calculated amongst the repetitions is in every case inferior to 1.1 %. Additionally, the linear regression showed a correlation coefficient superior to 0.999. The proportional relationship between the applied pressure and the flow rate clearly illustrates the fact that no significant suction effect is observed under the tested flow rates, which means that the injecting flow rate is only induced by the CE hydrodynamic injection system. This trend, in contrast to other nanoESI source, could be explained by the internal diameter of the capillary to prevent sample aspiration due to the electrical field. Indeed, the use of a capillary having an internal diameter of 30 µm increases the surface tension of the liquid preventing it from being pulled out of the capillary by the electrical field [11]. In addition, the CE system is able to deliver fully controlled hydrodynamic injections even at infusion flow rates of as low as 4 nL/min.

As a second step to characterize the electrospray generated through the CESI interface, the influence of flow rate on sensitivity under native conditions was studied. Effect of flow rate on sensitivity in ESI has already been discussed [18–28]. Under conventional ESI flow rates, native MS usually lacks sensitivity as it is intensively subject to ion suppression and as such typically benefits from the nanoESI regime. Using the CESI prototype as an infusion device, sensitivity was evaluated using myoglobin at a rather low concentration (250 nM) in native conditions (20 mM ammonium acetate, pH 6.8 as buffer). The sample was infused into the MS through the system at different flow rates ranging from 3.8 to 913 nL/min. Sensitivity defined as the ratio of MS spectrum intensity to the quantity of introduced sample in the source has then been plotted against flow rate (Fig. 2).

Results collected show a significant increase of the sensitivity as the infusion flow rate is reduced. As a consequence, sensitivity was increased by a factor of 46 while reducing the infusion flow rate from 456.9 to 13.7 nL/min. Even greater values have been obtained for lower flow rates. Signal-to-noise ratio is also significantly enhanced with respective values of 189.5 at 456.9 nL/min and 920.4 at 13.7 nL/min (data not shown) demonstrating reduction of the ion suppression effect. Such improvement in sensitivity, even for unfavored conditions as native

MS is a huge asset in the perspective of noncovalent complexes study. In agreement with data presented elsewhere, an increase in sensitivity was observed when decreasing the infusion flow rate [18]. In the referenced work, the described interface provided a stable spray at flow rates of as low as 10 nL/min below, thereby allowing to work not only in the concentration-sensitive region of ESI-MS but also in the mass-sensitive one.

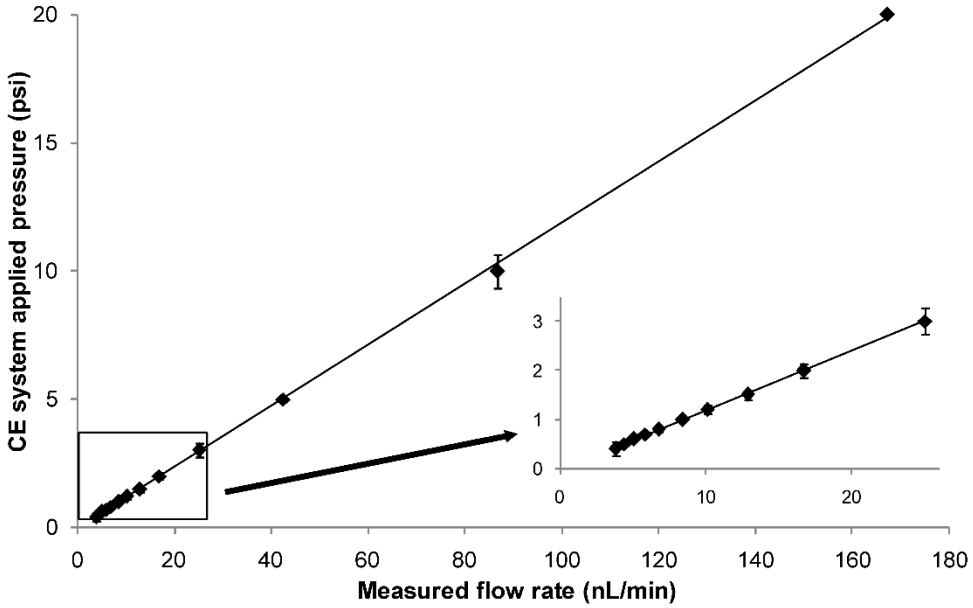

Figure 1. Evolution of the infusing flow rate as a function of the applied pressure. Experimental conditions: bare fused-silica capillary with porous tip: total length, 92 cm (30 µm i.d., 150 µm o.d.). Sample: myoglobin, 1 µM (in 10 % AA). MS: capillary voltage, −1,400 V (considered m/z 848.94)

As a direct consequence, lower ion suppression is expected to provide higher ionization efficiencies. In the context of the study presented, this consequence directly impacts the compatibility of the system to work under native conditions. This condition is well exhibited in Fig. 2 where an increase in sensitivity can be observed when flowrate decreases. Within the mass-sensitive region where the signal is proportional to the amount of molecules injected per unit of time, the sensitivity should theoretically remain constant. Nevertheless, it can be observed on Fig. 2 that the sensitivity keeps increasing when decreasing the flow rate. The related experiment being achieved under native conditions, we believe that this phenomenon is related to the ion suppression whose magnitude decreases with the flow rate.

Other nanoESI source uses capillaries having internal diameter ranging from 1 to 10 µm. With the CESI system the use of a 30 µm i.d. capillary allows to increase drastically the sensitivity as well. This could be explained by the fact that the flow rate is fully controlled. Mann and Wilm modelized the droplets emission radius from the Bernoulli equation; they demonstrated that this radius was proportional to a two thirds power of the flow rate [11], meaning that the droplet emission radius may be more influenced by the flow rate than the internal diameter; the ability

of the system to operate at a controlled nano-flow rate should not increase the emission radius allowing to maximize ionization by still focusing the droplet emission zone toward the MS inlet. Flow rate measures and sensitivities were repeated several times; they demonstrated the repeatability of the method which is shown in Fig. 2 by the relatively low deviation recorded from one experiment to another.

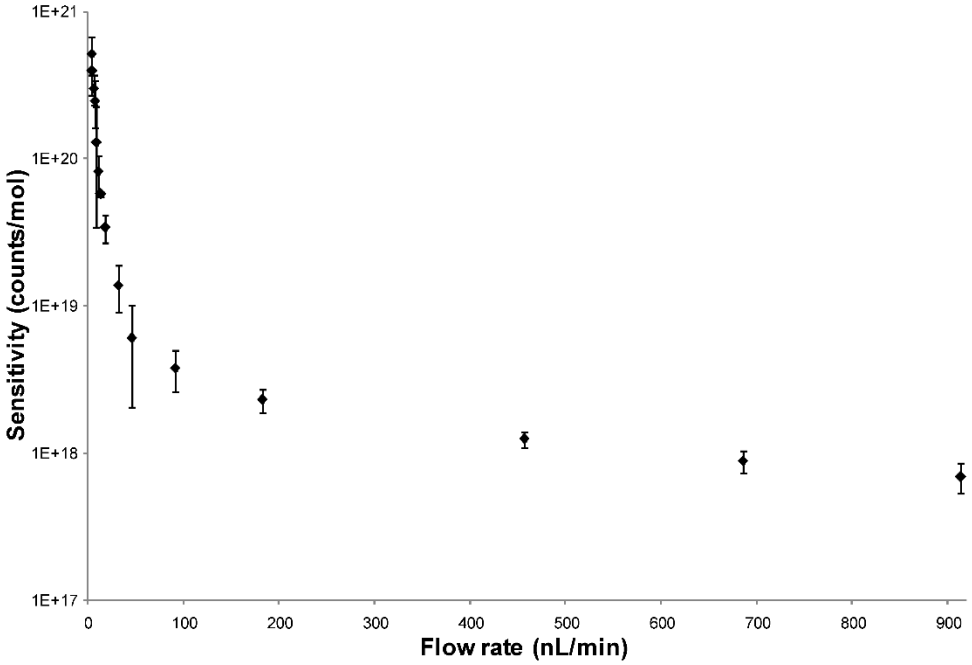

Figure 2. Evolution of signal sensitivity (logarithmic scale) as a function of the infusion flow rate in native conditions. Experimental conditions: bare fused-silica capillary with porous tip: total length, 92 cm (30 µm i.d., 150 µm o.d.). Sample: myoglobin, 250 nM (in 20 mM ammonium acetate). MS: capillary voltage, −1,400 V (considered m/z 2,196.8)

These results clearly demonstrate the capacity of the considered CESI-MS system to generate a so-called nanoelectrospray with the capabilities to have a full control on flow rate over an extended range going from around 4 to 500 nL/min and even more. In native conditions, sensitivities could be increased by a factor over 46 while improving as well the signal-to-noise ratio. These results clearly indicate the possibility to enhance sensitivity for detecting noncovalent complexes using this system.

Protein–ligand complexes characterization by CESI-MS and comparison to conventional ESI-MS. Study of biological protein–protein and protein–ligand noncovalent interactions are important and subject to intensive research as they are involved in many in vivo biochemical processes. Results presented previously illustrate the ability of the CESI prototype to be operated at very low flow rates thereby drastically increasing sensitivity. The suitability of the system for noncovalent complexes study was evaluated using a model protein. Pseudomonas processivity factor, also referred as β ring is a homodimer of about 85.45 kDa that encircles

the bacterial DNA and anchors the replicative DNA polymerase, which consequently acquires its high processivity. The β ring is also a landing platform for many other enzymes which contribute to DNA metabolism. A sample of β ring in 50 mM ammonium acetate (pH 7.0) was infused into the MS using either a conventional ESI source operating at a flow rate of 4 µL/min or the CESI prototype operating at 50 nL/min. MS transfer parameters were identical. Mass spectra obtained are presented in Fig. 3. The analytes detected have deconvoluted masses of 85,457.2±1.7 and 85,457.6±1.8 Da respectively using the conventional ESI source and the CESI interface (expected mass, 85,456.0 Da). These results allowed the accurate detection of the protein in its native homodimeric form with both sources.

As compared with conventional ESI, in terms of signal intensities, mass spectra obtained via CESI exhibited, for the main peak, an intensity of roughly five times higher. With both sources, the same four-charge states comprised between (20+) and (23+) could be resolved allowing intact protein mass determination with a satisfying residue inferior to 2 Da.

As compared with ESI, the quantity of β ring sample used to complete an experiment was reduced by a factor of 80 while using the CESI prototype. The total volume injected to the source was only 500 nL; this low value is clearly an advantage for noncovalent complexes mass analysis only available in low amounts as it allows to significantly increase the acquisition time to enhance resolution.

The study of noncovalent complexes induces the mass analysis of protein-ligand mixture at different concentrations to determine for example the stoichiometry or the binding constant of a given complex [13]. β ring is an interesting model as all the partners of the ring interact with a same binding pocket located on each monomer [29]. In this work, the ligand chosen was the P14 peptide which has a high affinity for the binding pocket of the β ring. The CESI prototype was then tested using a mixture of β ring and P14 peptide in ammonium acetate buffer (see “Material and methods”).

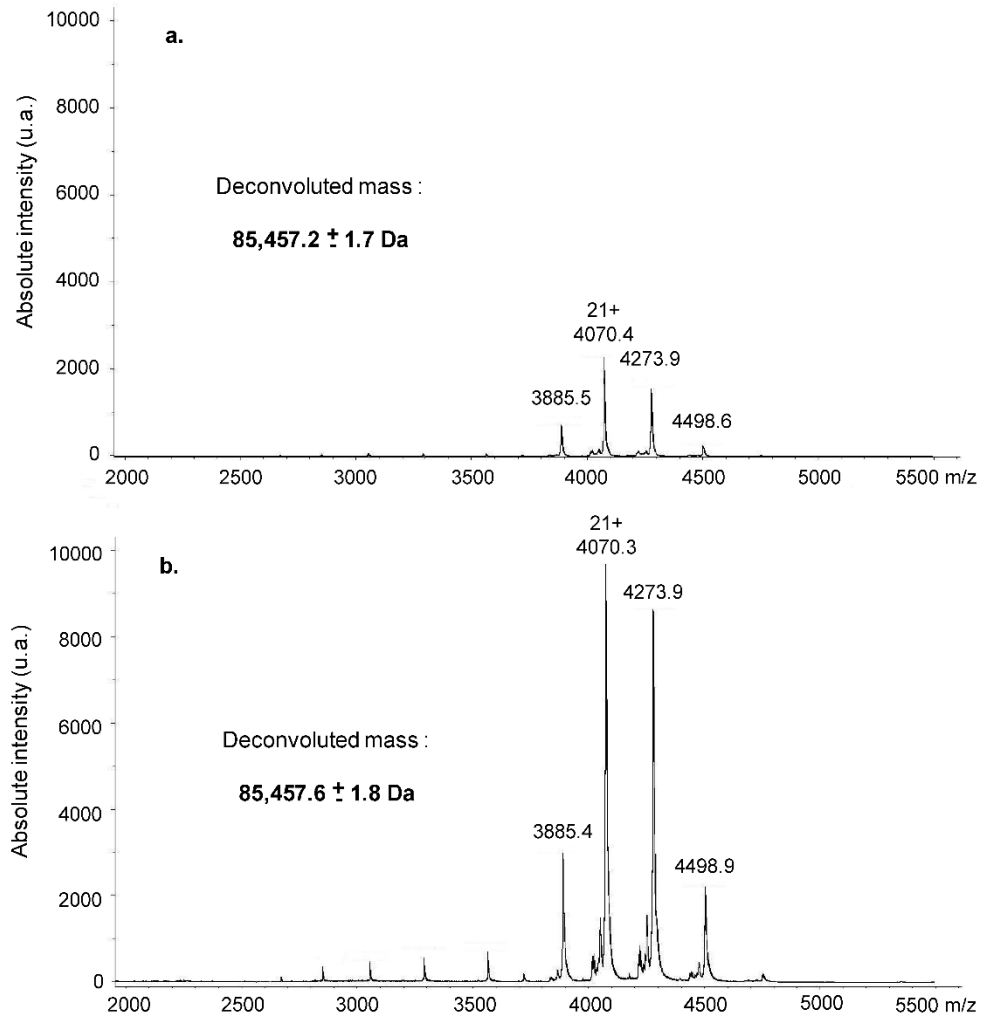

Figure 3. Comparison of β ring mass spectra in native conditions by a ESI-MS and b CESI-MS. Experimental conditions: β ring sample (1 µM in 50 mM ammonium acetate); a) conventional ESI source: flow rate, 4 µL/min; capillary voltage, −4,500 V; dry gas, 4 L/min; source temperature, 180 °C; b) bare fused-silica capillary with porous tip (30 µm i.d.): flow rate, 50 nL/min; capillary voltage, −1,500 V; dry gas, 1.5 L/min; source temperature, 180 °C

After a 5-min incubation period, the mixture was infused into the MS at a flow rate of 20 nL/min. The mass spectrum obtained presented in Fig. 4 exhibits three different envelope. The first one gives a deconvoluted mass of 85,456.5±1.1 Da corresponding to the β ring homodimer (expected mass, 85,456.0 Da). The second profile has a deconvoluted mass of 86,242.5±1.5 Da, meaning a mass difference of 785.8 Da as compared with the β ring homodimer profile (expected mass, 85,241.0 Da). This difference corresponds to the mass of the P14 peptide. Therefore, the second profile is associated to the noncovalent complex between the β ring homodimer and one P14 peptide. The last profile presented a deconvoluted mass of 87,028.3±0.8 Da; it corresponds to the complex between the β ring homodimer and two P14 peptides (expected mass, 87,026.0 Da).

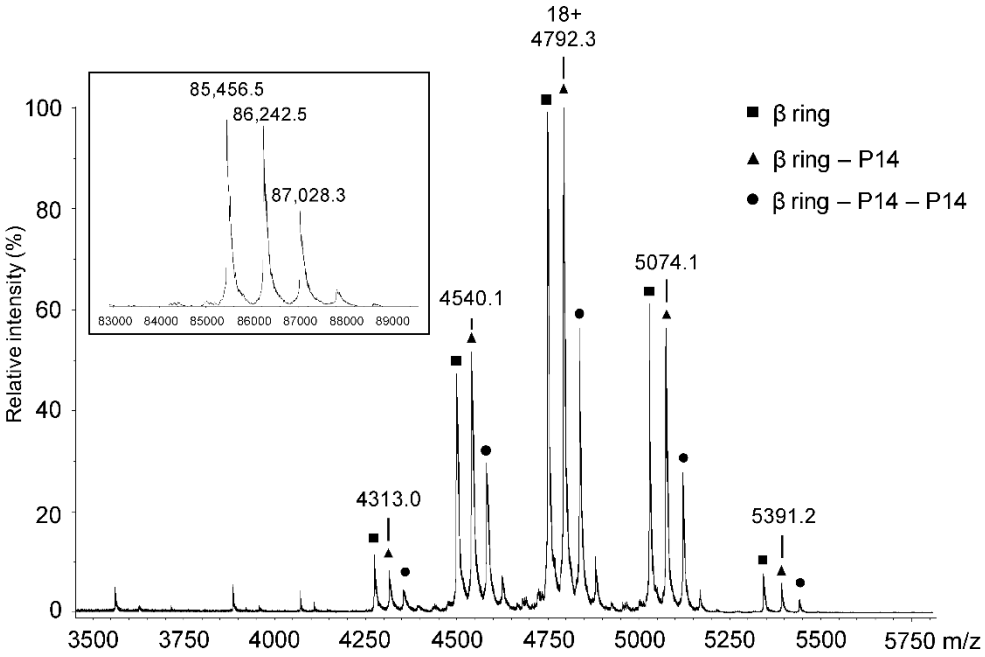

Figure 4. β ring and P14 peptide mixture raw and deconvoluted mass spectra in native conditions. Experimental conditions: β ring, 0.3 µM with 0.11 µM P14 peptide (in 50 mM ammonium acetate, 1 % DMSO). Bare fused-silica capillary with porous tip (30 µm i.d.): flow rate, 20 nL/min; capillary voltage, −1,600 V; dry gas, 1.5 L/min; source temperature, 180 °C

Along with the protein homodimer naturally formed in native conditions, two different protein–ligand complexes could be successfully resolved. The CESI prototype allowed to operate with reduced ion suppression phenomena, resulting in a significant signal-to-noise ratio value of 138.7, proper signal intensity and detection of several charge states. In addition, minimizing salt effect enables accurate detection of three different complexes even if those complexes have masses superior to 85 kDa and mostly m/z differences of only a few tens units. The infusion flow rate, reduced to 20 nL/min was again fully controlled by the CE system, thereby stabilizing the MS signal and increasing the robustness of the system. In this case, for a 10min acquisition, only 200 nLwere needed to acquire the presented mass spectrum (Fig. 4) and much longer acquisition could have been performed with less than 1 µL. During all experiments, no clogging of the capillary could be observed.

After each acquisition, the capillary was flushed with various buffers for several minutes and could be reused to infuse another sample without experimenting any cross-contamination. As compared with borosilicate capillaries used in conventional nanoESI which have internal diameter ranging from 1 to 5 µm, the internal diameter of the capillaries used in CESI is 30 µm; this difference significantly decreases the possible occurrence of clogging, which is a very valuable advantage in this field.

By the mean of this experiment, the system has been proven to be robust and easily operated. Additionally, a single capillary could be used to perform several infusion under denaturing or native conditions.

High molecular mass noncovalent complexes. Along with the development of biological noncovalent complexes study by MS, researches have been continuously driven toward bigger and more complex assemblies. Such assemblies are usually available in low quantities and require a buffer heavily concentrated in salt to increase ionic strength and prevent dissociation, the ESI ionization process further allowing the conservation of assemblies even if maintained by quite weak interaction and this regardless their molecular mass. Considering that time-of-flight mass spectrometry has virtually an unlimited mass range analyzing ability, the detection of a wide range of noncovalent assemblies is possible even those having molecular masses of several megadaltons [30]. To confirm the capacity of the system to analyze very high molecular mass noncovalent complexes, a sample of hemocyanin was studied. Hemocyanin is a protein present in mollusks and arthropods hemolymph, and is involved in oxygen transport. One particularity of this protein is that it may contain one or several hexamers whose subunits have a mass approximately 75 kDa [6]. Elements composing hemocyanin are maintained by weak noncovalent interactions. For this reason those complexes are usually difficult to observe in ESI-MS.

Hemocyanin directly extracted from C. maenas hemolymph was analyzed using the considered CESI-MS system. In this case, the sample was infused into the MS using a flow rate of 10 nL/min in order to maximize the enhancement of sensitivity. A typical mass spectrum obtained is presented in Fig. 5. It exhibits three different profiles. The first one detected in the range of 10,600–11,800m/z gives a deconvoluted mass of 445,242±33 Da and corresponds to the hexameric form of hemocyanin. The second profile detected between 13,600 and 15,550 m/z provides a deconvoluted mass of 899,496±505 Da which corresponds to the dodecamer form of hemocyanin. The last one detected in the range 16,100–17,900 has a deconvoluted mass of 1,337,489±456 Da and can therefore be attributed to the octodecamer form of the protein. Mass measures obtained with the CESI-MS system are in complete agreement with the values described in the literature [6]. In addition, the standard deviation on measured masses obtained in this work demonstrates the sensitivity of the system, indeed the deviation is always inferior to 0.05 % which may only be due to the formation of complexes having different type of hemocyanins as it was described before [6].

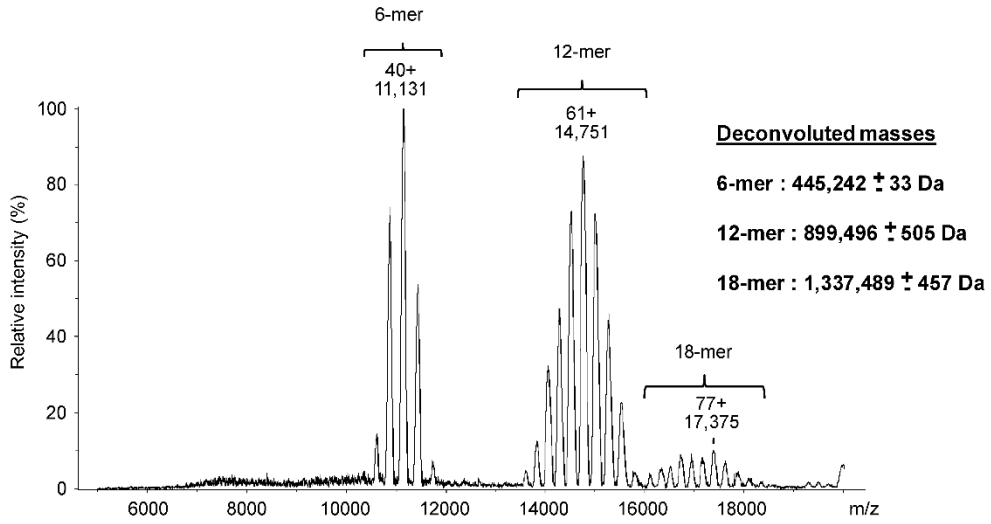

Figure 5. Hemocyanin mass spectra in native conditions. Experimental conditions: hemocyanin, 50 µM approx. (in 20mM ammonium acetate; pH=6.7). Bare fused-silica capillary with porous tip (30 µm i.d.): flow rate, 10 nL/min; capillary voltage, −1,400 V; dry gas, 1.2 L/min; source temperature, 140 °C

The controlled flow rate delivered by the CE system permitted to generate a stable spray during the acquisition. Thereby allowing the system to exhibit a substantial robustness by enabling intensive infusion and extended acquisition duration. Results presented on this mass spectrum demonstrate the detection of the three forms of hemocyanin present in the sample. Those noncovalent assemblies have high molecular masses ranging from 445 kDa to 1.34 MDa. Each complex could be structurally maintained with native conditions used and the ionization process delivered by the CESI prototype was suitable to profile those assemblies and especially the octodecamer hemocyanin are maintained by weak interaction and are typically difficult to ionize. To our knowledge, the only way to observe above 1 MDa noncovalent complexes in native conditions is to use borosilicate capillaries.

However, the method developed in this work is the first one to be able to perform fully controlled and automated infusion in nanoESI and observe high molecular mass complexes above 1 MDa. One of the main improvements provided by this method, as compared with other nanoESI platforms, is its ease of use. The sample can be automatically infused into the MS without any other prior treatment of the capillary than a short washing. During the experiments reported here, no clogging of the capillary could be observed reinforcing the ease of use of this system while it is known to be an issue with other nanoESI platform especially when using high ionic strength buffers. Despite demonstrating the possibility to use the CESI interface to produce a nanoESI with a flow rate as low as 4 nL/min, characterization of the system demonstrated the possibility to obtain a fully controlled nanoESI flow rate. It is demonstrated that the developed method provides the ability to maintain the structure of noncovalent

complexes throughout the nanoESI process to subsequently observe them in the MS. However, the utilization of the CESI prototype enabled the detection of the high molecular mass-considered complexes with significant intensities. Those results prove the ability of the system, under the tested conditions, to maintain the existing interactions enabling the detection of noncovalent complexes, realizing the full potential of nanoESI a soft ionization process. Deconvoluted masses could be determined with an acceptable precision due to the detection, with an adequate resolution, of five to ten different charge states depending on the considered complex. However, buffer salt concentration was significant (20 mM); the low flow rate at which the experiments were performed and the overall geometry of the interface provided a rather low background noise fully compatible with complexes detection. Intensity improvement, as compared with conventional ESI, are comparable to those obtained previously, the observed behavior being in agreement with the higher ionization efficiency reported using nanoESI [31]. The total volume of sample infused to perform an acquisition was about 74 nL which represent a 400 times reduction in terms of sample consumption as compared with standard ESI. This point is critical as the volume of hemolymph that can be extracted from C. maenas is quite limited.

Like in the previous experiment, no clogging of the capillary could be observed during the experiment. The capillary was simply flushed after each experiment with buffer and could be directly used after again. A single capillary could be used for several days to infuse different samples without experimenting any significant variation of the mass spectra obtained.

Conclusions

In this study, we have demonstrated the suitability of a prototype CESI-MS system to be used as a nanoESI infusion platform. This work was specifically designed to assess the capabilities of the CESI prototype for the study through infusion of intact proteins and high molecular masses noncovalent assemblies. First, the sample injection device of the system was characterized and its capacity to hydrodynamically generate a stable electrospray with fully controlled flow rates ranging from less than 4 nL/min to several hundreds of nanoliters per minute. The effect of the flow rate on the detection of an intact model protein under native conditions demonstrated an increased sensitivity when reducing the flow rate as a result of the reduction of ion suppression and increased ionization efficiency. Decreasing the flow rate from 456.9 to 13.7 nL/min increased the sensitivity by a factor of 46 while the signal-to-noise ratio improved from 189.5 to 920.4, this result illustrates the ability of the system to be operated at very low flow rates to increase drastically sensitivity. Using the CESI interface as a nano-electrospray emitter, the study of β ring under native conditions enabled the detection of the

protein homodimer exhibiting a mass of 85.5 kDa. The study of the β ring and P14 peptide complexes was then performed and showed the detection of two different protein–ligand complexes demonstrating the suitability of the CESI interface for noncovalent complexes detection. In this particular case, the stoichiometry of the complexes between the β ring and the P14 peptide was determined. During the experiments, the system exhibited high sensitivity, robustness, and easy handling as no clogging could be observed. As a direct consequence of the very low flow rates used, the quantity of sample necessary to perform an acquisition could be drastically reduced with in the most favorable case a 400 fold reduction in comparison to conventional ESI. As the CESI prototype is designed to separate experiments, results demonstrated the flexibility of the system as it can be used to infuse sample into the MS. It opens new fields of applications such as stability constant determination especially with the affinity. Capillary electrophoresis hyphenised to MS. Finally, the CESI prototype was used for the study of hemocyanin under native conditions. These experiments demonstrated the possibility to successfully detect high molecular mass complexes up to 1.34 MDa. This approach reveals the complete suitability and robustness of the CESI prototype as a nanospray emitter for the study of noncovalent complexes by MS. Additionally, the results obtained suggest that the utilization of this system could improve the quality of membrane protein studies as those typically require important ionic strength buffer while being available in scarce amounts.

Acknowledgments

Rabah Gahoual would like to thank the MRT for funding his Ph.D. work. LSMIS would like to thank Beckman Coulter Inc. for lending a prototype capillary electrophoresis system and for their global support, especially Michel Anselme, Hans Dewald and Edna Betgovargez, and Bruker Daltonics Inc. for the loan of a MicroTOFQ MS. The authors would like to acknowledge Dr. Dominique Burnouf (IBMC, Strasbourg) for providing β ring samples and Dr. Frank Zal from Biological Station of Roscoff (France) for hemocyanin samples.

References

[1] **Loo JA** (1997) Mass Spectrom Rev 16:1 – 23

[2] **Heck AJ, van den Heuvel RH** (2004) Mass Spectrom Rev 23:368 – 389

[3] **Sharon M, Robinson CV** (2007) Annu Rev Biochem 76:167 – 193

[4] **Benesch JLP, Ruotolo BT, Simmons DA, Barrera NP, Morgner N, Wang L, Saibil HR, Robinson CV** (2010) J Struct Bio 172:161–168

[5] **Benesch JLP, Aquilina JA, Ruotolo BT, Sobott F, Robinson CV** (2006) Chem Biol 13:597–605

[6] **Sanglier S, Leize E, Van Dorsselaer A** (2003) J Am Soc Mass Spectrom 14:419 – 429

[7] **Greig MJ, Gaus H, Cummins LL, Sasmor H, Griffey RH** (1995) J Am Chem Soc 117:10765 – 10766

[8] **Hilton GR, Benesch JLP** (2012) J R Soc Interface 9:801 – 816

[9] **Constantopoulos TL, Jackson GS, Enke CG** (1999) J Am Soc Mass Spectrom 10:625 – 634

[10] **Sterling HJ, Batchelor JD, Wemmer DE, Williams ER** (2010) J Am Soc Mass Spectrom 21:1045 – 1049

[11] **Wilm MS, Mann M** (1994) Int J Mass Spectrom Ion Processes 136:167 – 180

[12] **Wilm MS, Mann M** (1996) Anal Chem 68:1–8

[13] **Touboul D, Maillard L, Grässlin A, Moumne R, Seitz M, Robinson J, Zenobi R** (2009) J Am Soc Mass Spectrom 20:303 – 311

[14] **Painter AJ, Jaya N, Basha E, Vierling E, Robinson CV, Benesch JLP** (2008) Chem Biol 15:246 – 253

[15] **Moini M** (2007) Anal Chem 79:4241 – 4246

[16] **Haselberg R, Ratnayake CK, de Jong GJ, Somsen GW** (2010) J Chromatogr A 1217:7605 – 7611

[17] **Faserl K, Sarg B, Kremser L, Lindner H** (2011) Anal Chem 83:7297 – 7305

[18] **Busnel JM, Schoenmaker B, Ramautar R, Carrasco-Pancorbo A, Ratnayake C, Feitelson JS, Chapman JD, Deelder AM, Mayoboroda OA** (2010) Anal Chem 82:9476–9483

[19] **Heemskerk AAM, Busnel JM, Schoenmaker B, Derks RJE, Klychnikov O, Hensbergen PJ, Deelder AM, Mayboroda OA** (2012) Anal Chem 84:4552–4559

[20] **Wang Y, Fonslow BR, Wong CCL, Nakorchevsky A, Yates JR III** (2012) Anal Chem 84:8505 – 8513

[21] **Zhu G, Sun L, Yan X, Dovichi NJ** (2013) Anal Chem 85:2569 – 2573

[22] **Heemskerk AAM, Wuhrer M, Busnel JM, Koeleman CAM, Selman MHJ, Vidarsson G, Kapur R, Schoenmaker B, Derks RJE, Deelder AM, Mayboroda OA** (2013) Electrophoresis 34:383 – 387

[23] **Gahoual R, Burr A, Busnel JM, Kuhn L, Hammann P, Beck A, François YN** (2013) mAbs 5:479 – 490.

[24] **Ramautar R, Busnel JM, Deelder AM, Mayoboroda OA** (2012) Anal Chem 84:885 – 892

[25] **Haselberg R, Harmsen S, Dolman MEM, de Jong GJ, Kok RJ, Somsen GW** (2011) Anal Chim Acta 698:77 – 83

[26] **Haselberg R, de Jong GJ, Somsen GW** (2013) Anal Chem 85:2289 – 2296

[27] **Wolff P, Olieric V, Briand JP, Chaloin O, Dejaegere A, Dumas P, Ennifar E, Guichard G,Wagner J, Burnouf DY** (2011) J Med Chem 54:4627 – 4637

[28] **Schmidt A, Karas M, Dülcks T** (2003) J Am Soc Mass Spectrom 14:492 – 500

[29] **Burnouf DY, Olieric V, Wagner J, Fujii S, Reinbolt J, Fuchs RPP, Dumas P** (2004) J Mol Biol 335:1187 – 1197

[30] **Baclayon M, Shoemaker GK, Uetrecht C, Crawford SE, Estes MK, Venkataram Prasad BV, Heck AJR, Wuite GJL, Roos WH** (2011) Nano Lett 11:4865 – 4869

[31] **El-Faramawy A, Siu KWM, Thomson BA** (2005) J Am Soc Mass Spectrom 16:1702 – 1707

3. Additional discussions

The next paragraph is intended to present and discuss additional results which could be obtained during the characterization of the nanoESI generated by the intermediate of the CESI-MS system. Those results were not included in the publication presented previously.

3.1. Evaluation and optimization of the capillary position

a) Experimental settings description

To evaluate the influence of the position of the capillary, a sample of model protein (myoglobin 1µM) in denaturing conditions (H_2O/CAN 50/50) was injected to the MS using the CESI-MS as the infusion platform. The flow rate used for those injection was 50 nL/min. The injection was performed using different length separating the capillary and the MS inlet (from 1 to 3.5 mm). In each case the MS signal was recorded while ranging the ESI voltage from -1100 V to -2000 V. The signal corresponding to the protein was recorded (investigated m/z: 808.16 – 893.15).

b) Results and discussion regarding capillary position

The objective here is to optimize to the position of the capillary inside the source as it can be adjusted because the MS adapter in fixed on a mobile xyz axis. The results obtained are compiled in the figure 6.

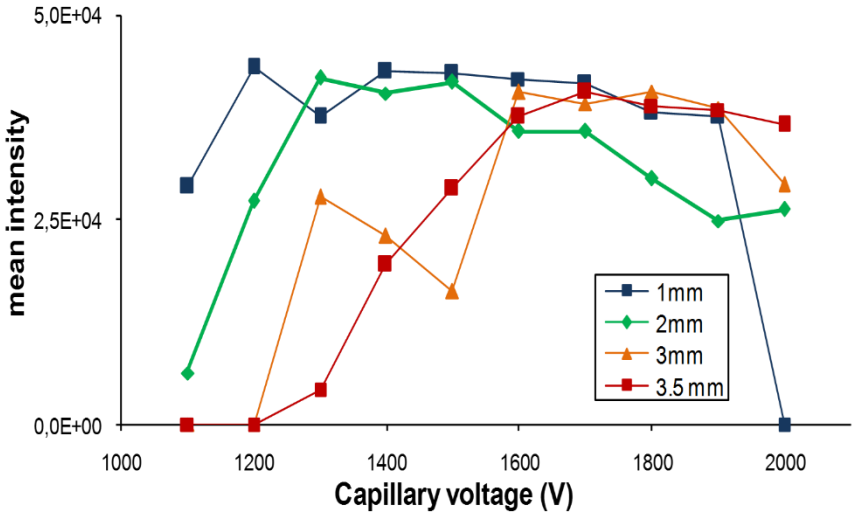

Figure 6. Graphical representation of the intensity corresponding to the signal of the model protein depending on ESI voltage applied, for several position of the separation capillary

Results recorded during the experiments showed first that the position of the capillary inside the source does not influence significantly the maximum MS intensity which can be achieved for a given sample. However the position is impacting the ESI potential which needs to be applied in order to generate the ESI. Indeed, if the distance between the CE capillary and the

MS inlet is increased, the ESI potential have to be increased also in order to obtain the best signal.

In the context of using the CESI-MS as infusion platform, in regard to those results, it was decided to use systematically a distance of 2 mm because it allows to maintain a spray for a wide range of ESI potential. That behavior is indicating that distance is compatible with various buffer having different volatilities.

3.2. Application of the CESI-MS as an infusion platform

Results described previously allowed to characterize the nanoESI generated by the CESI-MS system. Also the characteristics of the system (controlled and accurate injections, tolerance to clogging, multiple uses of capillary) proved its compatibility to non-covalent complexes studies in native MS.

From those results, the CESI-MS has been progressively adopted as a routine instrument to perform sample infusion in nanoESI and highly demanding experiments in native MS could be performed. Especially, experiments on ribosomes could be performed. Those experiments are particularly demanding because in order to be maintained, those analytes requires high concentration of salts (up to 1M) which would normally induced a consequent ion suppression effect rendering impossible the detection of the complex. As it is emphasize in figure 7, the use of the CESI-MS allowed to detect ribosomes.

That results is quite encouraging in the integration of the system for the study of highly complex non-covalent edifices such as ribosomes.

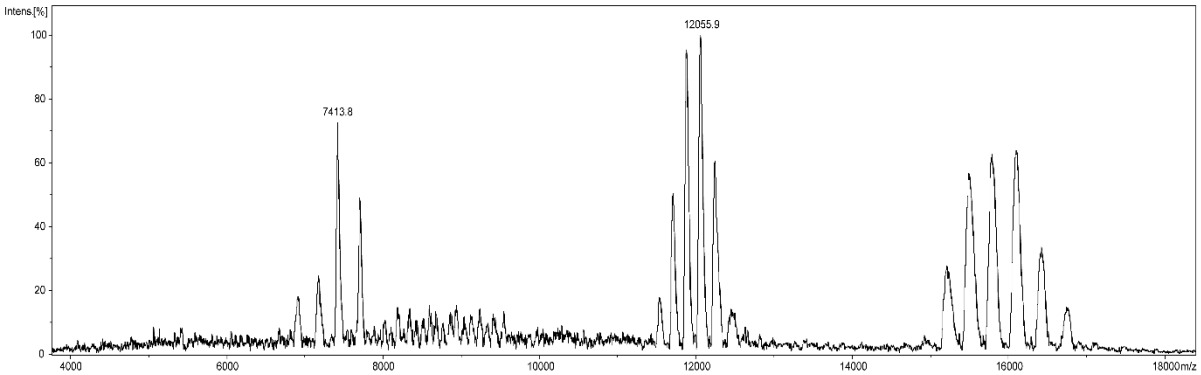

Figure 7. MS spectrum of 30S ribosome (buffer ammonium acetate 500 mM) obtained using the CESI-MS injection system (flow rate 50 nL/min)

Other experiments obtained in collaboration with the group of Prof. Compain about the evaluation a series of cyclopeptoid-based iminosugar clusters to finely probe the ligand

content-dependent increase in α-mannosidase inhibition, have been published in 2016 in *Chemistry – A European Journal* (Wiley 2016).

4. Conclusions

Les travaux présentés dans cette partie se sont, dans un premier temps, focalisés sur la caractérisation de l'ESI généré par l'intermédiaire du couplage CESI-MS. Pour ce faire, le système CESI-MS a été mis en œuvre en tant que plateforme d'infusion d'échantillons. L'objectif de cette étude est de profiter du système d'injection d'échantillon miniaturisé équipant la CE et des caractéristiques intrinsèques du capillaire afin de favoriser le processus d'ionisation en ESI et d'améliorer la qualité et la sensibilité du signal MS. Les résultats obtenus ont permis de démontrer que le système CESI-MS est capable de produire un ESI stable pour des débits d'infusion d'échantillon inférieurs à 4 nL/min, le système génère par conséquent un nanoESI. De plus les études effectuées ont permis de démontrer que le système d'injection hydrodynamique équipant la CE peut réaliser des infusions d'échantillons parfaitement contrôlées et reproductibles même pour des débits inférieurs à 10 nL/min. Au cours des expériences, aucun effet d'aspiration de l'échantillon n'a pu être observé. L'absence de phénomènes d'aspiration, potentiellement dû au champ électrique présent dans l'ESI et à la proximité entre l'extrémité du capillaire et l'entrée de spectromètre de masse, participe au bon contrôle du débit d'infusion de l'échantillon.

L'influence du débit d'infusion de l'échantillon par le système CESI-MS a de la même manière été évaluée dans des conditions de MS native, domaine qui souffre généralement d'un défaut de sensibilité. Les résultats obtenus ont montré que la sensibilité du signal peut être améliorée d'un facteur 46 par la diminution du débit d'infusion de 457 à 14 nL/min tout en améliorant le rapport signal/bruit d'un facteur proche de 5, ce résultat illustre l'amélioration de la qualité spectrale induite par l'ionisation en nanoESI. En plus du débit d'infusion faible, les caractéristiques du couplage CESI-MS, dont notamment le diamètre interne du capillaire, permettent de former des gouttelettes dont le diamètre réduit favorise l'efficacité d'ionisation. La possibilité d'automatiser l'infusion d'échantillon et l'amélioration de la sensibilité du signal de manière significative sont autant d'aspects qui suggèrent l'utilisation du système CESI-MS en tant que plateforme d'infusion nanoESI pour l'étude de protéines et de complexes non-covalents en conditions natives.

Le système CESI-MS a été utilisé pour réaliser l'étude d'un complexe protéique sous la forme d'un homodimère de 85 kDa, le même échantillon a également été analysé par ESI conventionnel afin de comparer les deux types de sources. Les résultats obtenus ont montré

une intensité du signal 4 fois plus importante dans le cas de l'infusion par CESI-MS par rapport à l'infusion ESI classique. De plus, l'utilisation du débit réduit a permis de diminuer 80 fois la quantité d'échantillon nécessaire pour réaliser une acquisition. Les interactions entre ce même complexe protéique et un peptide synthétique (P14) ont pareillement été étudiées. Les résultats obtenus ont permis de déterminer les trois différentes stœchiométries adoptées par ce complexe, la limitation des effets de suppression d'ions participe à l'obtention de spectres MS de qualité ce qui dans le cadre de cette étude a permis d'obtenir les masses exactes des différents complexes avec une précision largement satisfaisante. Ces résultats illustrent la capacité du système CESI-MS à réaliser l'ionisation de ce type de complexes tout en maintenant les interactions non-covalentes et par conséquent la compatibilité de cette approche instrumentale avec l'étude de complexes non-covalents.

Le système CESI-MS a également été utilisé pour étudier des complexes non-covalents de haut poids moléculaires. L'analyse d'échantillons d'hémolymphes de *Carcinus Maenas* par CESI-MS a été réalisée. Il a été ainsi possible d'observer les formes 6-mers, 12-mers et 18-mers d'hémocyanines dont les masses sont respectivement de 445 kDa, 900 kDa et 1.34 MDa. De tels résultats illustrent parfaitement la compatibilité de l'infusion par CESI-MS pour ce type d'étude. L'ionisation par l'interface CESI-MS délivre un rendement d'ionisation important tout en maintenant ce type de complexes structurés par des liaisons non-covalentes. Les infusions, réalisées avec un débit de 10 nL/min, représentent une quantité de 75 nL par acquisition. La quantité d'échantillon disponible est généralement un facteur limitant pour ce type d'analyse et l'infime quantité utilisée apparait donc comme un atout majeur. De plus les infusions ont pu être réalisées sans observer d'obstruction du capillaire ou de contamination croisée. Un simple rinçage du capillaire en utilisant le même tampon permet d'éliminer l'échantillon et plusieurs dizaines d'infusions ont pu être réalisées avec le même capillaire.

Ces travaux ont permis de démontrer la robustesse du système CESI-MS pour réaliser des infusions et le gain en sensibilité significatif pouvant être délivré par l'interface de couplage CESI-MS. Ces travaux ont initié plusieurs projets au laboratoire qui utilisent le système CESI-MS en tant que plateforme d'infusion nanoESI dans le cadre d'étude de complexes non-covalents en conditions natives. En plus de la possibilité d'utiliser le système CESI-MS en tant que plateforme d'infusion semi-automatisée d'échantillons, ces travaux mettent en évidence la capacité d'ionisation du couplage CESI-MS. Les expériences effectuées suggèrent que ce système de couplage fourni une excellente ionisation en ESI et font apparaitre le potentiel en terme de sensibilité que ce type de couplage peut fournir pour des analyses CE-ESI-MS, notamment du fait de l'absence de liquide additionnel. Il convient d'étudier la sensibilité que le couplage CESI-MS pour la caractérisation de peptides et de protéines dans le cadre d'analyses conventionnelles utilisant cette fois la CE en tant que méthode séparative.

II - Etude comparative des performances des couplages CESI-MS et NanoLC-MS sur l'analyse protéomique d'un extrait mitochodrial de *Saccharomyces cerevisiae*.

1. Introduction

Ces dernières années, l'analyse protéomique par spectrométrie de masse (MS) a été reconnue comme l'une des plus importantes méthodes pour l'identification de protéines notamment dans les recherches de leur fonctionnalité, de leurs possibles interactions dans la cellule ainsi que dans la découverte de nouveaux biomarqueurs pour la recherche médicale. Dans le but d'optimiser la détection et l'identification d'analytes et afin d'obtenir un plus grand nombre d'informations structurales, la MS couplée à une méthode séparative représente l'outil de choix pour l'analyse de mélanges complexes. Aujourd'hui, la chromatographie liquide (LC) apparait comme la méthode de référence pour le couplage avec les spectromètres de masse. Les avancées technologiques en LC, tel que la réduction des débits de la phase mobile ou l'utilisation de ultra hautes pressions, ont permis l'émergence de méthodes séparatives performantes totalement compatible avec la croissante complexité des échantillons. Ainsi, la NanoLC couplée avec un spectromètre de masse en tandem (NanoLC-MS/MS) est aujourd'hui la méthode la plus utilisées pour des applications d'analyses protéomiques. Les analytes sont élué sur une colonne phase inverse par l'augmentation de la concentration en solvants organiques dans la phase mobile. La séparation des peptides est ainsi basée suivant leur hydrophobicité. L'électrophorèse capillaire de zone (CZE) est une seconde technique séparative pouvant être couplée avec un spectromètre de masse à source d'ionisation électrospray (ESI). Le principe de la séparation des composés sous champs électrique est basé sur la mobilité électrophorétique des analytes qui dépend de la charge des molécules ainsi que leur rayon hydrodynamique. Ces dernières années, l'intérêt pour la CZE est en constante croissance notamment grâce aux améliorations technologiques permettant d'optimiser le couplage CZE-MS/MS par la mise en place de séparation à très faible débit (> 4 nL/min). Comme il a été démontré dans le chapitre précédent, comparé à la NanoLC-MS/MS, l'apport de ces très faibles débits permet une amélioration significative de la sensibilité et de la forte réduction des phénomènes de suppression d'ions.

En tenant compte de toutes ces considérations, il apparait évident que la mise en œuvre d'un protocole combinant la CZE-MS/MS et la NanoLC-MS/MS pourrait augmenter le niveau

d'informations obtenues due à l'orthogonalité de ces deux méthodes séparatives basée sur des principes physico-chimiques différents. En effet, due à la complexité croissante des échantillons et aux propriétés intrinsèques des peptides (taille, point isoélectrique, hydrophobicité), l'utilisation d'une seule méthode ne permet pas d'obtenir l'identification de la totalité des analytes. Un protocole expérimental mettant en jeu deux campagnes d'analyses par CZE-MS/MS et NanoLC-MS/MS sur le même échantillon peut être facilement imaginé. L'étude croisée des résultats obtenus permettra d'augmenter le niveau de confiance dans l'identification des protéines de cet échantillon.

Dans cette étude, nous proposons de réaliser l'évaluation d'un protocole expérimental combinant la CZE-MS/MS ainsi que la NanoLC-MS/MS afin d'explorer le protéome mitochondrial de *Saccharomyces cerevisiae*. Les expériences de CZE-MS/MS seront réalisées à l'aide de l'interface CESI-MS précédemment décrite. La levure de *Saccharomyces cerevisiae* est souvent utilisée comme système modèle dans le développement et la validation de nouvelles approches instrumentales. Ce choix s'explique principalement par le fait que de nombreux processus cellulaires essentiels sont conservés entre la levure et d'autres organismes d'intérêt comme l'humain. De nombreux avantages à l'utilisation de ce système découlent du fait que la levure est l'un des plus simples organismes eucaryotes. Ce grand intérêt en biologie moléculaire est illustré par le fait que ce système est le premier organisme eucaryote dont le génome a été totalement séquencé. La mitochondrie est une organelle intracellulaire complexe ayant un rôle crucial pour le fonctionnement de la cellule notamment dans le métabolisme normal de la cellule, l'énergie cellulaire, le maintien de l'homéostasie ionique et la mort cellulaire programmée. Afin de comprendre le rôle de cette organelle complexe, spécifiquement dans la recherche médicale, il est important d'obtenir le maximum d'information permettant de caractériser sa composition protéique. Aujourd'hui encore et malgré les avancées technologiques importante des techniques analytiques, l'identification complète des protéines d'un système complexe reste un défi majeur en biologie cellulaire. Quelques études ont été menées dans le passées sur l'analyse protéomique de mitochondrie. Ces études ont permis de répertorier entre 1000 et 1500 protéines différentes pour cette organelle, soit dans la levure soit dans des échantillons humains. Afin d'obtenir les meilleurs résultats et d'éviter de possible contamination, une préparation de mitochondrie purifiée doit être réalisée obligatoirement.

Le but de l'étude décrite dans ce chapitre, est de démontrer l'amélioration en termes d'identification protéique par une approche protéomique d'un extrait mitochondrial de *Saccharomyces cerevisiae* en combinant l'analyse par CZE-MS/MS et NanoLC-MS/MS. Cette étude a été réalisée sur une faible quantité d'échantillon (100 ng) représentant environ 1000 protéines. Les résultats d'identification obtenus en CZE-MS/MS et en NanoLC-MS/MS ont été

comparés. Les spécificités de chaque méthode séparative ont été analysées suivant différentes propriétés physico-chimiques des peptides identifiés : le poids moléculaire, le ratio charge sur masse, le point isoélectrique, la couverture de séquence ainsi que la qualité spectrale MS/MS. Une analyse de la localisation cellulaire de toutes les protéines identifiées a également été effectuée. De plus, l'accent a été mis sur l'un des grands défis de l'analyse protéomique qui consiste à caractériser les isoformes. Les résultats présentés dans cette étude ont permis d'évaluer le gain d'information pouvant être obtenu par ces deux méthodes orthogonales et complémentaires en termes de couverture de séquence et de caractérisation d'isoformes.

Les résultats présentés dans cette partie ont fait l'objet d'une publication dans le journal scientifique *Journal of Chromatographic Science* (Oxford Journals 2016).

2. Improvement of mitochondria extract from *Saccharomyces cerevisiae* characterization in shotgun proteomics using sheathless capillary electrophoresis coupled to tandem mass spectrometry

Marianne Ibrahim, Rabah Gahoual, Ludovic Enkler, Hubert Becker, Johana Chicher, Philippe Hammann, Yannis Nicolas François, Emmanuelle Leize-Wagner ***Journal of Chromatographic Science*** 2016, 54(4), 653-663.

Abstract

In this work, we describe the characterization of a quantity-limited sample (100ng) of yeast mitochondria by shotgun bottom-up proteomics. Sample characterization was carried out by sheathless capillary electrophoresis, equipped with a high sensitivity porous tip and coupled to tandem mass spectrometry (CESI-MS/MS) and concomitantly with a state-of-art nano flow liquid chromatography coupled to a similar mass spectrometry (MS) system (nanoLC-MS/MS). With single injections, both nanoLC-MS/MS and CESI-MS/MS 60 min-long separation experiments allowed to identify 271 proteins (976 unique peptides) and 300 proteins (1765 unique peptides) respectively, demonstrating a significant specificity and complementarity in identification depending on the physicochemical separation employed. Such complementary, maximizing the number of analytes detected, presents a powerful tool to deepen a biological sample's proteomic characterization. A comprehensive study of the specificity provided by each separating technique was also performed using the different properties of the identified peptides: molecular weight (MW), mass-to-charge ratio (m/z), isoelectric point (pI), sequence coverage or MS/MS spectral quality enabled to determine the contribution of each separation. For example, CESI-MS/MS enables to identify larger peptides and eases the detection of those having extreme pI without impairing spectral quality. The addition of peptides, and therefore proteins identified by both techniques allowed to increase significantly the sequence coverages and then the confidence of characterization. In this study, we also demonstrated that the 2 yeast enolase isoenzymes were both characterized in the CESI-MS/MS dataset. The observation of discriminant proteotypic peptides is facilitated when a high number of precursors with high-quality MS/MS spectra are generated.

Introduction

In recent years, mass spectrometry-based proteomics was recognized as one of the best tools for identifying proteins with new functions, mapping their interactions in a cellular context and discovering new biomarkers for medical research [1]. In order to maximize the number of analytes successfully ionized, detected and identified in mass spectrometry (MS), a physicochemical separation prior to MS analysis is often applied. Currently, liquid chromatography (LC) is the most widely used separation techniques that can be coupled on-line with a mass spectrometer [2]. In a classical nanoLC coupled to tandem mass spectrometry (nanoLC-MS/MS) experiment, which is the most usual method for proteomic application, analytes are eluted from a reverse-phase column by increasing the organic content of the mobile phase to separate peptides by hydrophobicity. Capillary Zone Electrophoresis (CZE) is another separating technique that has been successfully coupled to electrospray ionization [3]: with the use of an electrical field, ions are separated based on their electrophoretic mobility which is dependent upon the charge of the molecule and the analyte's hydrodynamic radius [4]. Capillary electrophoresis (CE) is gaining more and more interest mostly because of recent technical improvements regarding capillary zone electrophoresis–mass spectrometry (CZE-MS) coupling that allow the use of lower flow rates resulting in improved sensitivity [5,6] and also because ion suppression phenomenon is largely reduced compared to a nanoLC-MS analysis [7-9]. Recently, CE–MS has emerged as a highly efficient separation technique 10 and is considered as a powerful technique for the analysis of peptides and proteins [4,11-22].

Taking into considerations these observations, an ideal workflow would be the combination of orthogonal separative techniques based on different physiochemical principles. Indeed, as a significant fraction of proteins can escape detection in individual separation approaches [23], crossing over different techniques and their respective benefits would minimize that problem. Consecutive CZE-MS/MS and nanoLC-MS/MS analysis campaigns can be easily imagined on the same samples as further discussed in this study with the example of a quantity-limited sample from a yeast mitochondrial extract.

In this study, we used CZE-MS/MS as well as a classical nanoLC configuration to explore the S. cerevisiae mitochondrial proteome. CZE-MS/MS experiments were performed using a recently introduced CE-MS interface referred as CESI-MS. New instrumental approaches are often developed and validated on model systems such as the baker's yeast *Saccharomyces cerevisiae*, mainly because many essential processes are conserved between yeast and other organisms of interest like humans [24]. Numerous benefits derive from the fact that yeast is one of the simplest eukaryotes [25]. Because of its importance in molecular biology, it was the first eukaryotic organism for which the genome sequence was completed [26]. Mitochondrion

is a complex intracellular organelle, crucial for numerous cellular functions like normal cell metabolism [27], cellular energetic [28], maintenance of ion homeostasis [29] and programmed cell death [30,31]. To achieve the best results, a highly purified mitochondrial preparation is mandatory [32,33]. To understand the role of this complex organelle especially in disease, it is important to extensively characterize its protein composition: identifying the whole set of resident proteins within a complex organelle remains a major challenge in cell biology even if innovative technologies are emerging year after year. Previous proteomic analyses on purified mitochondria resulted in a repertoire of 1000 to 1500 different proteins for that organelle, in either yeast or human samples [34,35].

In the current work, the aim is to demonstrate the improvement in protein identification by shotgun bottom-up proteomics of mitochondria extract from *Saccharomyces cerevisiae* by using two different physicochemical separations prior to the MS analysis. The study is carried out on a limited quantity (100ng) of sample, roughly representing 1000 proteins. Identifications obtained by CESI-MS/MS were compared to those achieved in nanoLC-MS/MS. Specificity of each separating technique was determined observing different physicochemical properties of the identified peptides: molecular weight (MW), mass-to-charge ratio (m/z), isoelectric point (pI), sequence coverage as well as MS/MS spectral quality. Moreover, a cellular localization analysis of all identified proteins was also performed. Biological significance was finally investigated by comparing the protein identifications from this study to previously published studies. In addition, emphasis was placed onto a challenging area of proteomics research which is isoforms characterization. Taking together, the results presented in this study allowed us to assess the gain of information which can be achieved by two complementary and orthogonal techniques, in terms of peptide metrics, sequence coverage, spectrum quality and isoforms characterization.

Experimental

Materials and reagents.

Chemicals used were of analytical grade or high purity grade and purchased from Sigma-Aldrich (Saint Louis, MO, USA).Water used to prepare buffers, mobile phases and sample solutions was obtained using an ELGA purelab UHQ PS water purification system (Bucks, UK). Dithiothreitol (DTT) and iodoacetamide (IAA) were purchased from Sigma (St. Louis, MO, USA). Trypsin sequencing grade was obtained from Promega (Madison, WI USA). All other reagents and plastic ware were obtained from commercial sources.

Isolation of Mitochondria from S. cerevisiae

Yeast cells were grown in Yeast Peptone Dextrose (YPD) at 30°C up to OD600nm=1, 2.Mitochondria were isolated as detailed in the Supplemental Information section. Briefly, spheroplasts have been first generated by the enzymatic digestion of the cells wall by Zymolase 20T (Euromedex) and further broken by homogenization in a glass-Teflon potter. After the removal of cell debris and nuclei by centrifugation (1,500×g, 5 min at 4°C), mitochondria were recovered from the supernatant by additional centrifugations at 12,000×g for 15 min at 4°C, to give a final concentration of 5 mg/ ml. Mitochondrial fraction was treated by 10 strokes in a glass-Teflon potter and loaded onto a four-step sucrose gradient, from which purified mitochondria were recovered from the 60 %/ 32 % interface.

In solution protein digestion

10µg of protein were diluted in 75µL ammonium bicarbonate buffer (50mM) and incubated with 5µL DTT 0.1M for protein denaturation (10min, 95°C). After cooling down, 10µL IAA 0.1M were added and the samples were incubated for 20min at room temperature in the dark to allow alkylation of reduced cysteine residues. For protein digestion, 500ng (1:20) trypsin (Promega, V5111) was added and sample was incubated overnight at room temperature. For nano-LC/MSMS analysis, 3µg of the clear supernatant were transferred into glass inserts and evaporated to dryness using a miVac DNA concentrator (Genevac, NY, USA). The sample waseither reconstituted in 15µL formic acid (0.1% v/v) for nanoLC-MS/MS analyses or in ammonium acetate 50 mM (pH 4.0) for CESI-MS/MS analyses.

NanoLiquid Chromatography

Samples were transferred into glass inserts, compatible with the LC autosampler system (ultra nanoLC-2D plus, Eksigent, UK). Five microliters of each sample were loaded on a nanoFlexcHiP module consisting of a Trap-and-elute jumper (fluidics configuration design), a ChromXP-C18 trap cHiP column (0.5cm x 200µm i.d., 3µm, 120Å, Eksigent) and a ChromXP-C18 analytical cHiP column (15cm x 75µm i.d., 3µm, 120Å, Eksigent). The separation method consisted in a 60 min run at a flow rate of 300 nL/min using a gradient of two solvents: A (99.9% water: 0.1% formic acid) and B (99.9% acetonitrile: 0.1% formic acid). After a 10 min step to trap the peptides on the pre-concentration column at 20µL/min, they were eluted from the analytical column as follows: 0-35 min, 5-40% B; 35-36 min, 40-80% B; 36-40 min, 80% B; 40-41 min, 80-5% B; 41-60 min, 5% B.

Capillary Electrophoresis

The CE experiments were performed with a CESI8000 capillary electrophoresis system from Sciex separations (Brea, CA) equipped with a temperature controlled autosampler and a power supply able to deliver up to 30 kV. Bare fused-silica capillaries (total length 91 cm; 30µm i.d.) with characteristic porous tip on its final 3 cm supplied by Sciex separation allowed electric contact between both electrodes of the CE system. New capillaries were flushed at 75 psi (5.17 bar) for 10min with methanol, then 10min with 0.1M sodium hydroxide, followed 10min with 0.1M hydrochloric acid and water for 20min. Finally, the capillary was flushed 10min at 75 psi with BGE which was acetic acid 10%. Hydrodynamic injection (410 mbar for 1min) corresponding to a total volume of 55 nL of sample injected was used. Separations were performed using an electric field of +219.8 V/cm^{-1}.

Mass spectrometry

The LC and CESI systems were coupled to a 5600Triple TOFTM system (AB Sciex, CA, USA) with nanospray source operating in positive ESI mode. Interchangeability of the LC and CESI systems is easy and takes only 20 min. The mass spectrometer was operated in the IDA mode (Information-Dependent Acquisition) and was externally calibrated each 6 hours with 25fmol of an E.coli Beta-galactosidase digest. This standard also allowed to regularly check the performances and reproducibility of the mass spectrometer. For the nanoLC-MS/MS, the following ESI source parameters were used: ion spray voltage, 2.3 kV; ion source heater, 150 °C; curtain gas, 22 psi and ion source gas 1,5 psi. While for the CESI-MS/MS and due to the nanospray properties, the ESI source parameters were: ion spray voltage, 1.75 kV (1.0 to 4.0 mm between the MS inlet and capillary tip); ion source heater, 75 °C; curtain gas, 5 psi (flow rate < 100 nL/min) and ion source gas 1, 0 psi. The MS survey scan was acquired over a 400-1250 m/z mass range with a 250ms accumulation time; whereas the product ion experiments mass range was extended to 100-2000 m/z with a 60ms accumulation time. The duty cycle time was therefore maintained below 2s with up to 30 precursor ions selected on the MS survey scan to undergo CID fragmentation (Collision Induced Dissociation), while respecting the following criteria: intensity greater than 150 counts per second, charge state between 2+ and 5+, not present in the dynamic exclusion list, exclusion after a period of 15s and 2 consecutive MS/MS spectra, and use of a rolling collision energy. Data were acquired with Analyst 1.6 software (AB Sciex) and the resulting .wiff files were primarily evaluated with PeakView 1.2 software (AB Sciex).

Data Analysis

For protein identification, the MS/MS raw data files were first converted to mascot generic format (mgf) using the executable MS-DataConverter.exe (AB Sciex). MS data processing was then performed with Proteinscape software (version 3.1, Bruker Daltonics). We used the Mascot algorithm (v2.2, Matrix Science, UK) for database searching, configured with the following parameters: (a) database: Swiss-Prot with Baker's Yeast taxonomy (version 2013-01-09, 18860 sequences), (b) enzyme: Trypsin/P, (c) maximum missed cleavages: 3, (d) variable modifications: acetylation (Protein N-term), oxidation (Met), carbamidomethylation (Cys), (e) peptide mass tolerance: 30 ppm, (f) MS/MS tolerance: 0.5 Da, (g) Instrument: ESI-Quad-TOF. To evaluate the false-positive rate of each analysis, we generated a reversed yeast database using a Pearl script (makeDecoyDB.pl, Bruker): the fusion of this decoy database with the original forward yeast database allowed Proteinscape to validate the data at FDR<1%. Moreover, redundancy was taken into account by Proteinscape (grouping of proteins sharing exactly the same set of peptides) and a manual inspection of the MS/MS fragmentation spectra was done for proteins identified with 1 peptide and a Mascot Score below 50. The same MS/MS raw data files were then submitted to a second database search algorithm, (Paragon, AB Sciex) with the following parameters: (a) database: the same Baker's Yeast taxonomy as described previously for Mascot, (b) enzyme: Trypsin, (c) Cys alkylation: Iodoacetamide, (d) ID focus: biological modifications, (e) search effort: thourough ID, (f) confidence p-value: 0.05, (g) Instrument: Triple-TOF 5600.

To facilitate the biological interpretation of the protein identification lists, the PANTHER classification system (http://pantherdb.org/) has been applied in a first instance on both nanoLC-MS/MS and CESI-MS/MS datasets. A second classification system named iLoc-Euk (http://www.jci-bioinfo.cn/iLoc-Euk) was then used to assess the subcellular localization of each identified protein. Combining both prediction softwares during the final data mining step allowed to strengthen the confidence related to each protein biological significance. Parameters used for both systems are detailed in the Supplemental Information section.

Results

Peptide metrics to highlight CESI-MS/MS and nanoLC-MS/MS approaches.

Consecutively to the development of the ESI ionization source and its application to study biological analytes in MS, technical developments have been achieved in order to perform a separation prior to the MS analysis. Performing a separation before the ionization process

allows to analyze samples with an increased complexity and tends to enhance the sensitivity of the signal by limiting the competition effect during ESI ionization [36]. Interfaces allowing CE-MS hyphenation were therefore developed as CE is well suited for the separation of biological analytes such as peptides or proteins while providing high efficiency separation due to the absence of stationary phase. In this work, a recently introduced sheathless CE-MS interface (CESI) has been used. A detailed description of the system has been provided in the literature by Haselberg et al.[13]. Recently, following the improvement of shotgun proteomics with recent mass spectrometer, different research groups have described shotgun proteomic experiments performed by sheathless CE-MS using commercially available or custom-made interfaces [37,38]. One of the major goals of this study was in a first instance to identify a great number of proteins from a relatively low amount of yeast mitochondrial extract. For that purpose, we used either the state-of-art shotgun proteomic configuration with a RP-nanoLC-MS/MS interface, or the promising CESI-MS/MS interface [39-41]. ESI parameters for NanoLC and CESI systems are defined in order to be compatible with nanoESI flowrate (300 nL/min for nanoLC and <100 nL/min for CESI) and totally comparable in terms of performance (See Experimental section). 100ng of material, a quantity which can be typically obtained after a sub-cellular fractionation process on a wide variety of organisms, were injected on both couplings. When considering studies published during the last 5 years, we can notice that starting sample amounts ranging from 1ng to 400ng were often considered, with an average at 100ng (Supplemental Table S1). Moreover, as stated by Sun et al in their review published in 2014 [37], the vast majority of proteomic studies using a nanoLC-MS/MS approach are typically starting with biological samples quantities in the milligram or microgram range. That's why CESI-MS/MS approach appears as a good alternative when the starting material is limited and falling into the midnanogram range.

To obtain a simple picture from the two complex MS datasets generated, heat maps were reconstructed using PeakView software. The mass/charge ratio of the peptides (Da) is plotted against the LC elution time or the CZE migration time (Figures 1-A and 1-B). A nanoLC-MS/MS heat map, like the one we obtained for the injection of 100ng of a yeast mitochondrial extract, is generally characterized by a wide cloud of dots, each of them representing a unique MS/MS spectrum related to the fragmentation of a given precursor ion (Figure 1-A). This cloud of dots means that LC conditions are regularly eluting peptides from the reverse-phase column, without introducing any additional phenomenon except the fact that hydrophobic peptides will elute latter than hydrophilic peptides. In contrast, the heat map obtained with the CESI-MS/MS interface displays a very different picture: distinct and successive lines or "strikes" of analyzed peptides can be drawn on the graph (Figure 1-B). This pattern can be explained by the CZE

migration process itself as it depends not only on the charge in solution of the analytes but also on their mass-to-charge ratios.

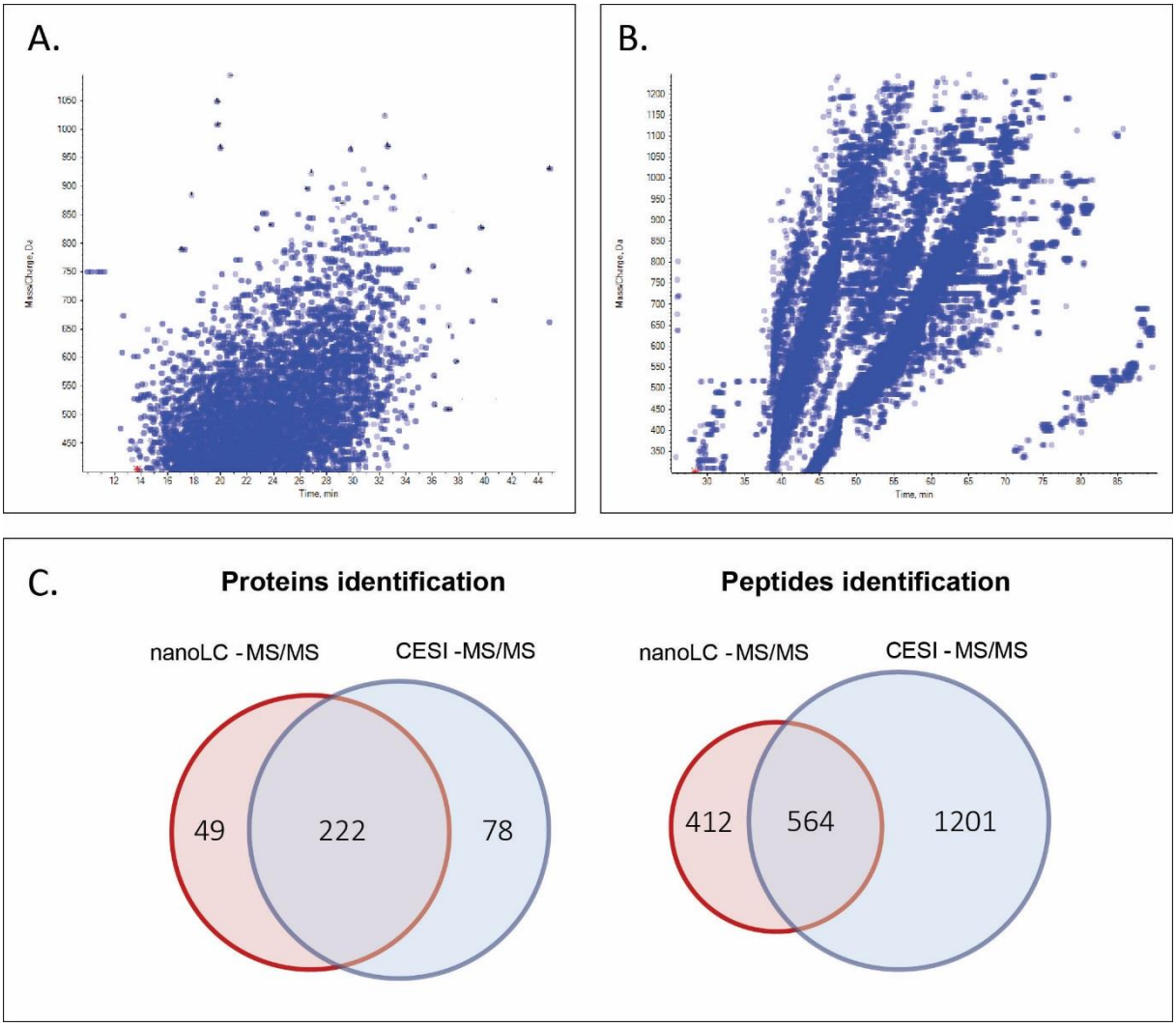

Figure 1. MS data obtained on 100ng of yeast mitochondrial extract. (A) NanoLC-MS/MS injection: heat map, where the mass-to-charge ratio of the peptides (Da) is plotted against the LC elution time (min). Each blue dot is related to a single MS/MS fragmentation spectrum. (B) CESI-MS/MS injection: heat map, where the mass-to-charge ratio of the peptides (Da) is plotted against the CE migration time (min). (C) Number of proteins and peptides identified in either the nanoLC-MS/MS (red) or the CESI-MS/MS (blue) injection.

A yeast UniProtKB database was used to identify peptides and proteins and a maximum false-positive rate of 1% was applied to validate both MS datasets. As shown in the Venn diagram (Figure 1-C), 222 proteins were identified with either CESI-MS/MS or nanoLC-MS/MS analysis. In addition, 78 proteins were only identified by CESI-MS/MS while 49 other proteins were only identified by nanoLC-MS/MS. Proteins identified by both approaches represent the majority of the proteins identified in the two nanoLC-MSMS and CESI-MSMS analyses, 82% and 74% respectively. This complementarity observed between both couplings also agrees with previous reports [37,42]. Interestingly, CESI-MS/MS coupling allowed to identify 3 times more additional peptides compared to the classical nanoLC-MS/MS configuration (412 versus 1201

peptides), providing already a positive insight into the power of CESI-MS/MS for obtaining good sequence coverages. It must be emphasized that the number of identified proteins for each approach was obtained with a single injection and is not the sum of multiple technical repeats. Among the 222 proteins identified commonly by both CESI-MS/MS and nanoLC-MS/MS, half of them are characterized by shared peptides, as well as additional peptides identified specifically by CE or LC (scenario #2 = EQUIV in Supplemental Figure S1): this scenario leads all the same to an equivalent result which is the identification of a given protein in the biological sample. Interestingly, this observation means that the two orthogonal methods of separation used in this study were identifying peptides bearing the same biophysical properties (564 common peptides, representing 26% of the whole dataset) but that there are more additional peptides only identified by the CE approach (1201) compared to the LC approach (412). This observation is also concordant with the fact that 41 proteins identified by LC are also identified by CE but with a higher number of peptides (scenario #3 = CE+ in Supplemental Figure S1). To note, 25% of the 222 common proteins are characterized by a completely different set of peptides (scenario #5 = DIFF in Supplemental Figure S1). This latest observation is nicely showing the usefulness of combining orthogonal separative methods before the mass spectrometry analysis.

To evaluate each type of coupling for identification of peptides bearing specific biophysical properties, we next represented the percentage of observed peptides obtained by each separation against their molecular weights (MW, Figure 2-A), mass-to-charge ratios (m/z, Figure 2-B), isoelectric points (pI, Figure 2-C) and Mascot scores (Figure 2-D). The most evident trend arises from the molecular weight distribution pattern: very large peptides were mostly detected with CESI-MS/MS rather than nanoLC-MS/MS. Indeed, only 10 peptides observed by nanoLC-MS/MS have a molecular weight above 2800 Da, whereas this number is reaching 466 peptides in the CESI-MS/MS injection: this might be explained by their irreversible adsorption to either the preconcentration or the chromatographic column, or the difficulty to correctly elute these large peptides. The detection of large peptides by CESI-MS/MS is also a phenomenon visible in the medium mass range: the apexes of both MW distributions are separated by 400 Da, with the CESI-MS/MS data displaying the highest apex and the wider distribution. For example, in the medium mass range from 751 to 1250 Da, 52% of peptides were detected with CESI-MS/MS, versus 10% with nanoLC-MS/MS. On the contrary, in the low mass range from 400 to 750 Da, 90% of peptides were detected with nanoLC-MS/MS, versus 48% with CESI-MS/MS (Figure 2-A). The same results are obtained when we look at mass-to-charge ratios of the peptides, which are more widely used instead of the global molecular weight (Figure 2-B). In this study, no enrichment towards low MW peptides was observed in the CESI-MS/MS dataset. Interestingly, some other groups also highlighted,

on another biological system, the fact that CESI-MS/MS allows to elute peptides of low MW [18]. These peptides are often composed of only a few amino acids, thereby giving them a hydrophilic property. This type of peptides is also more difficult to observe with nanoLC-MS/MS, as they are weakly bound or even unretained on reverse-phase stationary phases.

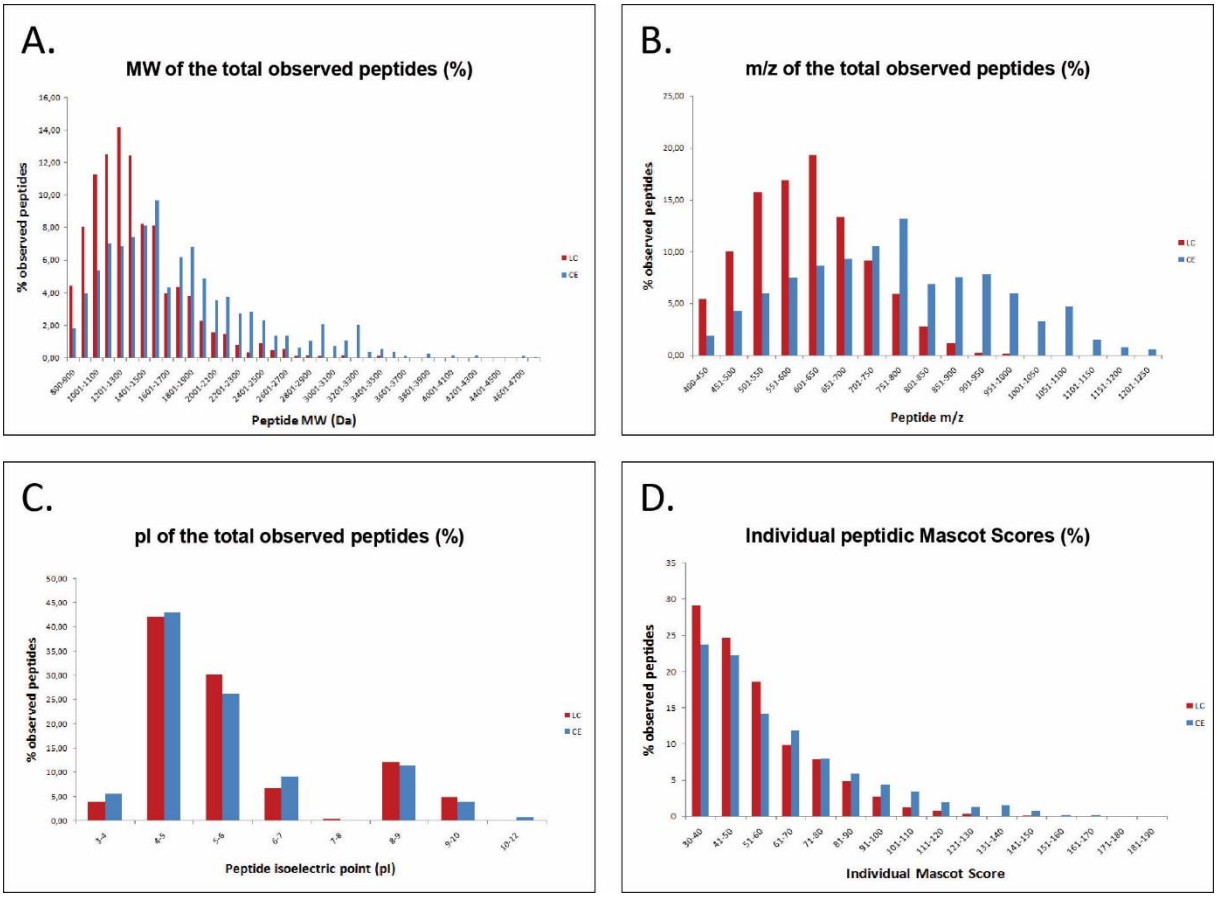

Figure 2. Comparison of peptides molecular weight (A), mass-to-charge (*m/z*) ratio (B), pI (C) and Mascot score (D) by means of CESI-MS/MS (red) and nanoLC-MS/MS (blue). Data are obtained from the analysis of 100ng of a mitochondrial yeast tryptic digest on both types of coupling.

Another important feature to monitor is the pI value from the identified peptides. Here, we used the « Compute MW/pI » tool from Expasy to obtain the isoelectric point values for all the observed peptides, resulting in a bi-modal distribution with unevenly distributed peptidic pI values across the pH scale, with a gap at pH 7-8, and the majority of peptides clustering at pH 4-6 (Figure 2-C). It can be noted that the pI distribution obtained on this S. cerevisiae mitochondrial extract is in agreement with previously published data like theoretical 2D-gels reconstructed by Knight et al. [43] on a wide variety of organisms including the total yeast proteome as well as sub-cellular yeast fractions. Results on pI distribution obtained in this study on a mitochondrial yeast extract are relatively similar. Slight differences can be observed all the same for the extreme pI values, especially extremely basic peptides (pI>10) or peptides with more acidic property (pI<4), as already shown by other research groups 37. This

phenomenon is potentially a result of the decreased ion suppression at very low flow rates and could be of particular interest for phosphopeptides detection without any enrichment techniques [42].

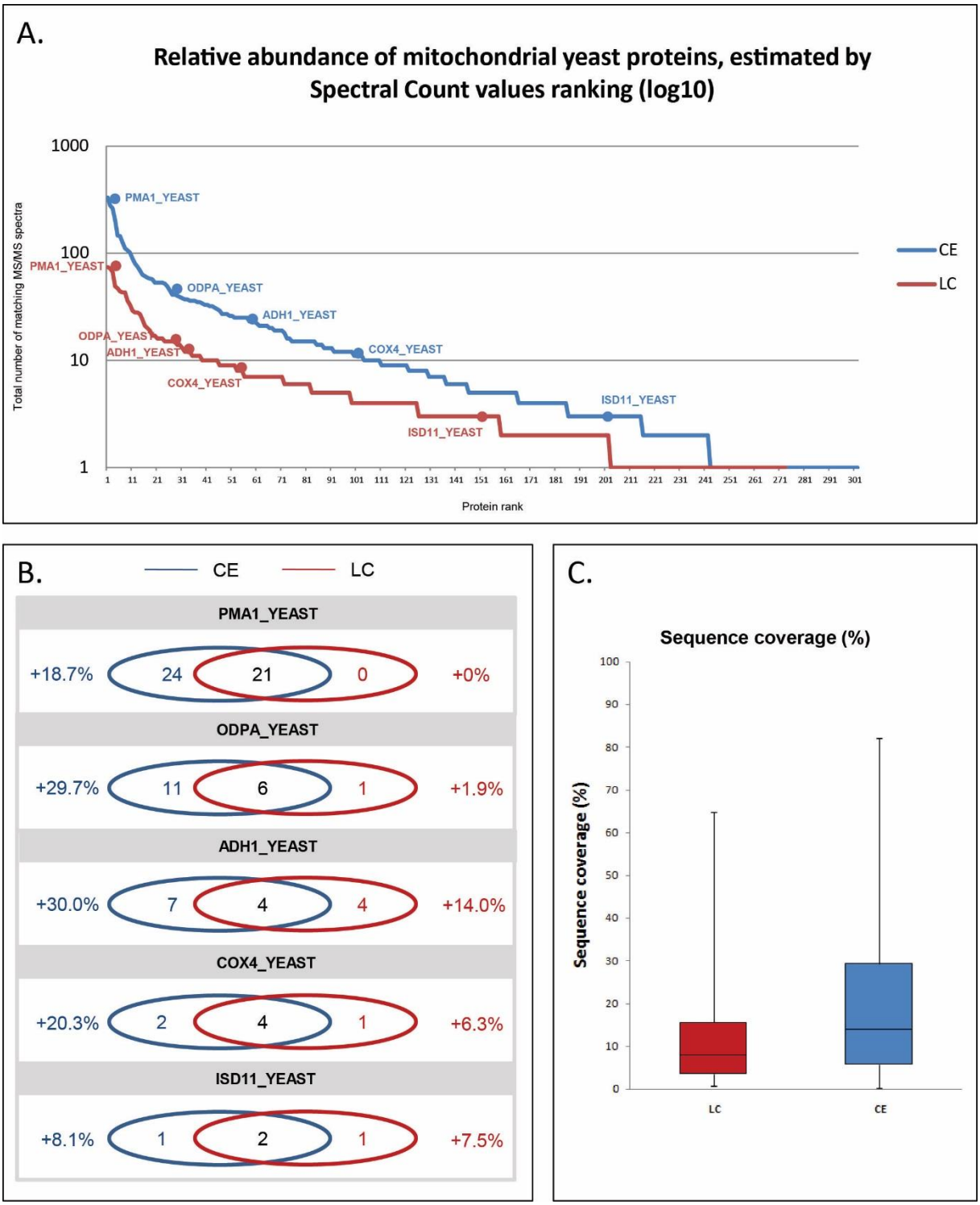

Figure 3. Effect on the global protein coverage when comparing the CESI-MS/MS (blue) and the nanoLC-MS/MS (red) datasets. (A) Proteins are ranked according to the total number of MS/MS spectra matching to each sequence (log scale) and 5 proteins, distributed all along the dynamic range, were further investigated. (B) Comparison of the sequence coverage from the 5 previously chosen proteins: shared and specific peptides are indicated in each Venn diagram, as well as the total sequence coverage increase (in %). (C) Boxplot showing the sequence coverage distributions for both datasets when considering all the identified proteins.

Discussions

Increasing sequence coverage versus spectrum quality.

Maximal sequence coverage is recommended to increase protein identification confidence but coverage should not be won to the detriment of the spectral quality. As this parameter can be evaluated with the Mascot Score obtained for each spectrum, we next represented the Mascot Score distributions obtained on the CESI-MS/MS and nanoLC-MS/MS injections of 100ng mitochondrial extract (Figure 2-D). Interestingly, there are more peptides with a score lower than 61 identified by the nanoLC-MS/MS approach. The CESI-MS/MS coupling provides more peptides with a score upper than 61: these high quality MS/MS spectra are representing 39.8% of the whole CESI dataset and 27.7% of the whole LC dataset. As for the median Mascot Score, it shows a slight increase from 48.3 to 52.8, for the LC or CE separations respectively (Supplemental Figure S4-A). A wider distribution is also observed for the CESI-MS/MS dataset, and this observation is more noticeable when considering the sub-set of peptides composed of the LC- and CE-specific features (Supplemental Figure S4-A). This Mascot Scores comparison is thus reflecting the overall good spectral quality obtained for both types of coupling.

As Wang et al [39] did previously while using the Sequest XCorr, we decided to use the Mascot score, which is probably the most widely-used database search algorithm by the proteomic community. Moreover, to strengthen these data, we also used a second dataset search algorithm, namely Paragon (ProteinPilot, AB Sciex): relevant observations can thus be conducted by combining 2 different algorithms, as recommended by proteomics guidelines. Paragon algorithm returns the percentage of MS/MS spectra above a given confidence threshold: we can observe that there are always more fragmentation spectra above this threshold in the CESI-MS/MS dataset compared to the nanoLC-MS/MS dataset, whether the confidence threshold is set (up to 99%, Supplemental Figure S4-C).

There is a particular interest in the development of alternative technologies that could improve sequence coverages, to secure more confident identifications, especially for low-abundant proteins often characterized by a single or a few peptides. To study the usefulness of CESI-MS/MS to increase sequence coverage, we first ranked the identified proteins according to the total number of MS/MS fragmentation spectra matching to each of them. While plotting the resulting Spectral Count value as a function of the protein rank, we were able to select a panel of 5 different proteins which were homogeneously distributed all along the dynamic range from both datasets (Figure 3-A). When specifically inspecting the peptides distribution from this sub-set of proteins, we can observe that sequence coverage is indeed increased in the CESI-MS/MS analysis compared to the nanoLC-MS/MS analysis (Figure 3-B). While the effect on

the sequence coverage increase is drastically observable on high-abundant proteins like PMA1, ODPA and ADH1, a more reasonable impact is observed for mid- and low-abundant proteins like COX4 and ISD11. We further considered the whole set of proteins and thus represented the distribution of the global protein sequence coverage depending on the type of coupling (Figure 3-C) and observed the same trend. The example of the yeast plasma membrane ATPase PMA1 can illustrate the sequence coverage increase that was observed in this yeast mitochondrial dataset (Supplemental Figure S3-A). 24.8% sequence coverage was reported for PMA1 by nanoLC-MS/MS with 21 identified peptides, while sequence coverage from CESI-MSMS was 43.5% with 45 identified peptides (Supplemental Figure S3-A). Interestingly, the 21 peptides identified by nanoLC-MS/MS were all identified by CESI-MS/MS (annotated in red on Supplemental Figure S3-A). Moreover, among the additional 24 peptides identified only by CESI-MS/MS, it can be noted that 7 of them have a large mass-to-charge ratio (m/z > 1000, highlighted in yellow on Supplemental Figure S3-A): these 5 unique peptides at [175-215], [216-252], [386-414], [482-508] and [584-615] are already representing 17.7% of the whole PMA1 sequence (17678 Da over 99619 Da for the full length protein). For the 21 peptides identified by both separative techniques, 17 of them have a CE Mascot Score higher than the corresponding LC Mascot Score (highlighted in grey in Supplemental Figure S3-A). To note, the MS/MS fragmentation spectra of the 5 largest peptides are of good quality as suggested by their average Mascot Score (84.6, Supplemental Figure S4-B) and have acidic pI values (average pI = 4.25). Such very large peptides are heavily contributing to increase the total sequence coverage. Interestingly, as suggested by other works, CESI-MS/MS is also able to identify small peptides that aren't by nanoLC-MS/MS. For example, the C-terminal part from PMA1 was covered in the CESI-MS/MS analysis with the doubly-charged peptide [910-918] being sequenced at m/z=529.7611, as well as 4 other small peptides (highlighted in grey on Supplemental Figure S3-A, peptides [272-278], [380-385], [429-435] and [436-442]). Similarly, when investigating the 4 other proteins distributed all along the dynamic range from the yeast mitochondrial sample (Figure 3-A), the same sequence coverage increase is observed (Supplemental Figures S3-B to S3-E).

S.cerevisiae mitochondrial proteins identification

Biological interpretation of the data is of paramount importance for assessing the strength of mass spectrometric and proteomic results, and validation using orthogonal techniques is often requested to publish MS data. In this study, a focus was established on the yeast mitochondrial sub-proteome. Indeed, the determination of a protein's localization is useful for biologists because it is linked to its cellular function [44]. It can also pinpoint some molecular functions to specific organelles [45]. Thus, in order to perform a high-throughput cellular localization analysis of all identified proteins, two protein classification systems were used: PANTHER [46]

and iLoc-Euk [47]. With the nanoLC-MS/MS approach (Figure 4-A), 55.3% of proteins have been identified and clustered as mitochondrial proteins while 53% of them were classified as mitochondrial proteins in the CESI-MS/MS dataset (Figure 4-B). In Figure 4-B, we can see that 127 mitochondrial proteins overlap between the two methods. Among the 349 proteins identified with both methods, 52.1% are known mitochondrial proteins, while the rest of them have been reported to be located into other subcellular compartments.

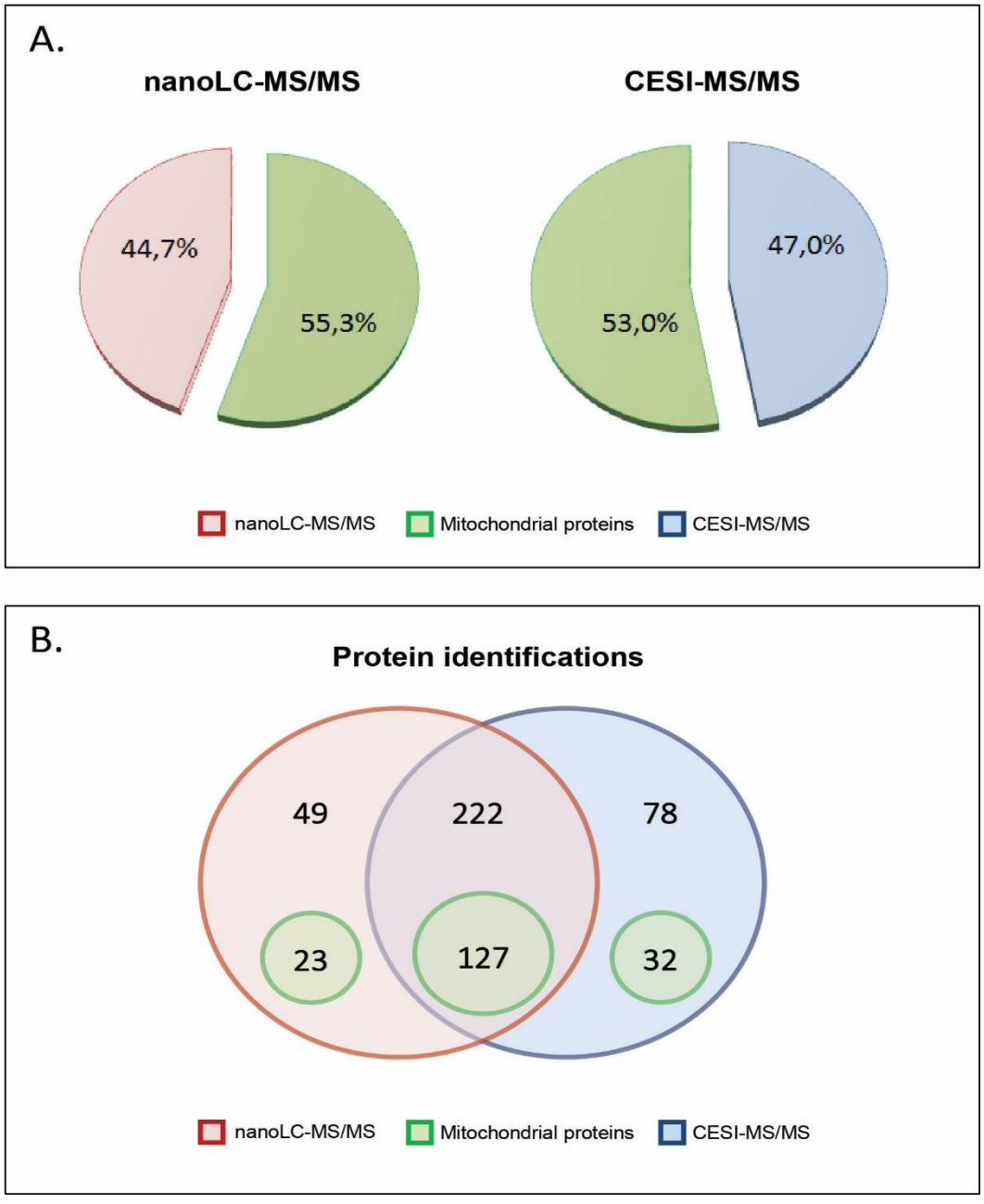

Figure 4. Venn diagrams indicating the distribution of mitochondrial proteins identified with Panther and iLoc-Euk databases using nanoLC-MS/MS and CESI-MS/MS approaches. Distribution of proteins between mitochondrion and other sub-cellular compartments using the nanoLC-MS/MS dataset (A) or the CESI-MS/MS dataset (B). (C) displays the overlap between the global protein dataset and the mitochondrial proteins using the two approaches.

Identifying 100% of the proteins of an entire organelle from only one proteomic experiment remains a big challenge. Some proteins may escape from detection because of the technology used to separate the analytes upstream from the mass spectrometer itself: this effect is often emphasized when analytes are of extreme biophysical properties (like hydrophilic or very small peptides). To evaluate the pertinence of our proteomic dataset compared to the set of mitochondrial proteins already seen by coworkers, we selected 16 studies dealing with yeast mitochondria published between 1997 and 2014 |23,33,35,48-60]. Accession numbers were homogeneously converted and aligned: many proteins thus appeared to be well-characterized by the community as seen by the frequency of observation, whereas a sub-set of 23 proteins was only identified in our study. Among them, an intriguing protein named ENO1 was further investigated because it was only identified in the CESI-MS/MS dataset with 41 MS/MS spectra. Enolases are essential glycolytic enzymes that catalyse the interconversion of 2-phosphoglycerate to phosphoenolpyruvate. They are present under two isoforms, named A and B, in *Saccharomyces ceravisiae* [61]. It is very important to describe a biological sample as exhaustively as possible, thus including isoforms characterization by mass spectrometry: this is possible by the identification of discriminant proteotypic peptides matching on each isoform sequence. Figure 5-A displays the number of distinct peptides identified for each enolase isoform using either the CE or LC approach: isoform B (ENO2) was covered by a higher number of peptides even if the 2 proteins have slightly the same molecular weight. 95% of identity is achieved on 495 residues when aligning the two protein sequences with Blast-P (98% of positive matches, Figure 5-B), rendering these isoforms difficult to distinguish. While investigating the detailed sequences of the matching peptides using the Sequence viewer from Proteinscape software, we can conclude that ENO1 (isoform A) was not validated in the LC dataset because its 6 peptides were an exact sub-set from the 8 peptides matching on the ENO2 sequence (ENO1 is a so-called sub-set protein). Whereas in the CE dataset, 2 proteotypic peptides were identified for the ENO1 sequence and 8 proteotypic peptides were specifically matching on the ENO2 sequence (Figure 5-C). To note, MS/MS fragmentation spectra of good quality were obtained for these proteotypic peptides (Figure 5-D) and allow to distinguish two peptide sequences with close amino acids composition. Likewise, the CE results also allow to observe two yeast glyceraldehyde-3-phosphate dehydrogenase isoforms (G3P2 and G3P3). The CESI-MS/MS injection was finally the only analysis in which we were able to distinguish the two yeast enolase isoforms with a good confidence. Moreover, we also clustered the protein identifications in terms of protein families, to see if specific families were identified by one of the two approaches. The CESI-MS/MS dataset was seen to increase the number of members from four protein families, related to mitochondrial protein import (TIM), ergosterol biosynthetic process (ERG), mitochondrial electron transport (QCR) and ATP synthesis (ATP) (Supplemental Figure S5-A). Proteins belonging to the TIM, QCR and ATP

families are located in the mitochondrial inner membrane whereas proteins from the QCR family are related to cell membrane, where sterols are targeted.

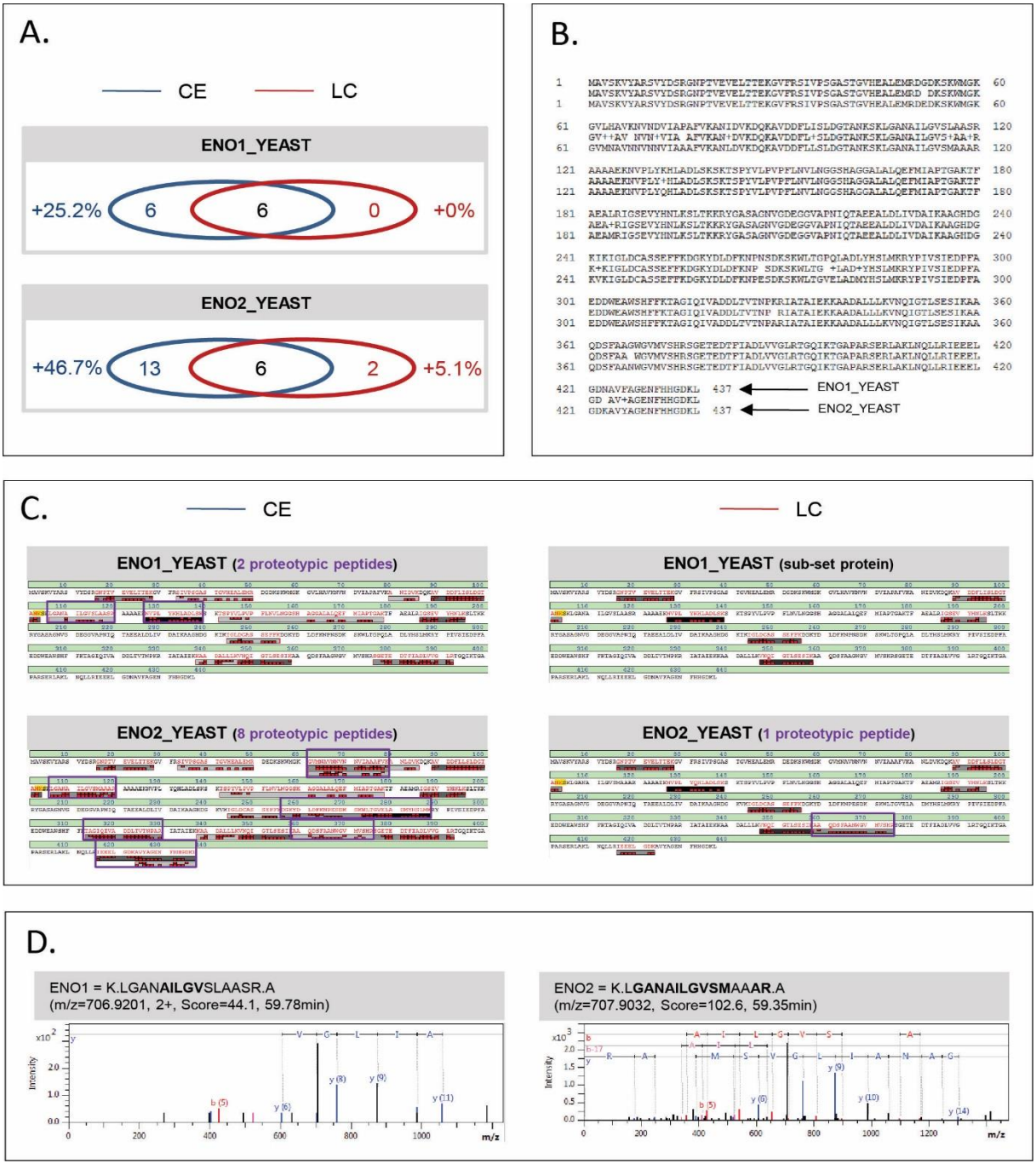

Figure 5. How is it possible to distinguish 2 protein isoforms using discriminant peptides and an increased sequence coverage: ENO1 and ENO2 proteins as an example. (A) Comparison of the sequence coverage for each isoform: shared and specific peptides are indicated in the Venn diagrams, as well as the total sequence coverage increase (in %). (B) BLAST-P alignment between both ENO isoforms. (C) Detailed sequence coverage on both isoforms obtained by either the CESI-MS/MS or the nanoLC-MS/MS analysis: identified amino acids are shown with a red label, whereas the grey box below the corresponding peptide is indicating the MS/MS spectrum quality (a red box is attributed under each amino acid if it was seen with either a b- or a y-ion, represented in the upper and lower section of the grey box respectively). Proteotypic peptides allowing the distinction between both isoforms in each dataset are indicated with purple boxes. (D) Focus on 2 proteotypic peptides matching on ENO1 and ENO2 sequences: a difference of 2 amino acids in the composition of a peptide is easily distinguishable when inspecting the MS/MS fragmentation spectra.

Given the importance of membrane proteins in various cellular processes, as well as their roles in diseases, it is important that this class of proteins be better studied. But identifying and characterizing proteins embedded in membranes still remains a challenge in proteomics due to difficulties during the solubilization step. With the complementary identifications obtained by combining orthogonal separative techniques, a better chance is given to membrane proteins and therefore hydrophobic related peptides to be detected and identified by MS. Exploring the 2 datasets with the same type of clusterization also allows to identify protein families that are covered by CESI-MS/MS and nanoLC-MS/MS in an equivalent manner, with members specifically identified by only 1 of the 2 techniques: this is the case for the 40S and 60S ribosomal proteins family (Supplemental Figure S5-B).

Conclusion

To compare correctly datasets and to assess an eventual complementarity between CESI-MS/MS and nanoLC-MS/MS techniques, we used a large panel of peptide properties: MW, m/z, pI, Mascot Score, % peptides per protein and % sequence coverage. Based on the results presented in this study, when a low quantity of yeast mitochondrial extract is analyzed, CESI-MS/MS enables the identification of more peptides than nanoLC-MS/MS, and thus a higher protein sequence coverage. This improvement is consistent with the actual proteomics guidelines, as drastic procedures must be employed to assess the maximum of confidency on protein identification. Moreover, one of the more visible trend is concerning the peptide metric related to the MW of the observed peptides: indeed, more peptides having a molecular weight above 2000 Da have been detected by CESI-MS/MS, whereas only some of them falling into the same mass range have been identified by nanoLC-MS/MS. In parallel to the increase of the number of large peptides, the peptide charaterizing metric concerning the pI distribution was also investigated: our results indicate that more extreme pI values (pI<4 and pI>10) are covered by CESI-MS/MS rather than nanoLC-MS/MS. These observations are leaving the door open for the detection of peptides carrying particular biophysical properties, like phosphopeptides, without the need of enrichment techniques. In this study, the evaluation of the overall spectral quality from the CESI-MS/MS and nanoLC-MS/MS data was performed by considering the Mascot score distributions and the Paragon confidence thresholds. Results are suggesting that increasing the sequence coverage has no detrimental impact on spectrum quality.

A challenging area of research in proteomics concerns isoforms charaterization. There are several explanations for protein isoforms, among which multiple gene copies (allele variation), alternative splicing, post-translational modifications (PTM) or degradation products. In this

study, we demonstrated that the 2 yeast enolase isoenzymes were both charaterized in the CESI-MS/MS dataset. The observation of discriminant proteotypic peptides is facilitated when a high number of precursors with high-quality MS/MS spectra are generated. Traditional approaches are combining different proteolytic enzymes or are benefiting from the integration of bottum-up proteomics with top-down and middle-down approaches. A combination of orthogonal separative techniques, coupled online to the mass spectrometer, also appears to be a good alternative to decipher a complex proteome.

The evaluation of both techniques using a large panel of peptide metrics allow us to say that they possess complementary properties for peptide and protein identification in mitochondria isolated from cultured S. cerevisiae. Moreover, the power of integrating two orthogonal separative technologies and two protein classification systems offers new promising opportunities to researchers working in the mitochondrial field of yeasts and other organisms, and more broadly in subcellular proteomics.

Acknowledgments

Authors would like to thank Sciex separations Inc. for lending a CESI8000, Dr. M. Anselme and Dr. Stephen Lock from Sciex Inc. for their support. This work was supported by the “Laboratoires d'excellence” (LABEX) NetRNA grant ANR-10-LABX-36 (PR) in the frame of “Programme d'Investissements d'Avenir.”

References

[1] Aebersold, R., Mann, M.; Mass spectrometry-based proteomics; Nature, (2003); 422: 198-207.

[2] Smith, R.D., Olivares, J.A., Nguyen, N.T., Udseth, H.R.; Capillary zone electrophoresis mass-spectrometry using an electrospray ionisation interface; Analytical Chemistry, (1988); 60: 436-441.

[3] Mann, M., Wilm, M.; Electrospray mass-spectrometry for protein characterization; Trends in Biochemical Science, (1995); 20: 219-224.

[4] Ramautar, R., Heemskerk, A.A.M., Hensbergen, P.J., Deelder, A.M., Busnel, J.M., Mayboroda, O.A.; CE-MS for proteomics: Advances in interface development and application; Journal of Proteomics, (2012); 75: 3814-3828.

[5] Hjerten, S.; High-performance electrophoresis - The elcectrophoretic conterpart of high-performance liquid-chromatography; Journal of Chromatography, (1983); 270: 1-6.

[6] Jorgenson, J.W., Lukacs, K.D.; Zone electrophoresis in open-tubular glass-capillaries; Analytical Chemistry, (1981); 53: 1298-1302.

[7] Busnel, J.M., Schoenmaker, B., Ramautar, R., Carrasco-Pancorbo, A., Ratnayake, C., Feitelson, J.S., Chapman, J.D., Deelder, A.M., Mayboroda, O.A.; High Capacity Capillary

Electrophoresis-Electrospray Ionization Mass Spectrometry: Coupling a Porous Sheathless Interface with Transient-Isotachophoresis; Analytical Chemistry, (2010); 82: 9476-9483.

[8] Gahoual, R., Busnel, J.M., Wolff, P., Francois, Y.N., Leize-Wagner, E.; Novel sheathless CE-MS interface as an original and powerful infusion platform for nanoESI study: from intact proteins to high molecular mass noncovalent complexes; Analytical and Bioanalytical Chemistry, (2014); 406: 1029-1038.

[9] Haselberg, R., de Jong, G.J., Somsen, G.W.; Low-Flow Sheathless Capillary Electrophoresis-Mass Spectrometry for Sensitive Glycoform Profiling of Intact Pharmaceutical Proteins; Analytical Chemistry, (2013); 85: 2289-2296.

[10] Heemskerk, A.A.M., Busnel, J.M., Schoenmaker, B., Derks, R.J.E., Klychnikov, O., Hensbergen, P.J., Deelder, A.M., Mayboroda, O.A.; Ultra-Low Flow Electrospray Ionization-Mass Spectrometry for Improved Ionization Efficiency in Phosphoproteomics; Analytical Chemistry, (2012); 84: 4552-4559.

[11] Geiger, M., Hogerton, A.L., Bowser, M.T.; Capillary Electrophoresis; Analytical Chemistry, (2012); 84: 577-596.

[12] Fonslow, B.R., Yates, J.R.; Capillary electrophoresis applied to proteomic analysis; Journal of Separation Science, (2009); 32: 1175-1188.

[13] Haselberg, R., de Jong, G.J., Somsen, G.W.; Capillary electrophoresis-mass spectrometry for the analysis of intact proteins 2007-2010; Electrophoresis, (2011); 32: 66-82.

[14] Gahoual, R., Burr, A., Busnel, J.M., Kuhn, L., Hammann, P., Beck, A., Franccois, Y.N., Leize-Wagner, E.; Rapid and multi-level characterization of trastuzumab using sheathless capillary electrophoresis-tandem mass spectrometry; Mabs, (2013); 5: 479-490.

[15] Wojcik, R., Dada, O.O., Sadilek, M., Dovichi, N.J.; Simplified capillary electrophoresis nanospray sheath-flow interface for high efficiency and sensitive peptide analysis; Rapid Communications in Mass Sepctrometry, (2010); 24: 2554-2560.

[16] Zhu, G.J., Sun, L.L., Yang, P., Dovichi, N.J.; On-line amino acid-based capillary isoelectric focusing-ESI-MS/MS for protein digests analysis; Analytica Chimica Acta, (2012); 750: 207-211.

[17] Sun, L.L., Zhu, G.J., Li, Y.H., Wojcik, R., Yang, P., Dovichi, N.J.; CZE-ESI-MS/MS system for analysis of subnanogram amounts of tryptic digests of a cellular homogenate; Proteomics, (2012); 12: 3013-3019.

[18] Faserl, K., Sarg, B., Kremser, L., Lindner, H.H.; Optimization and Evaluation of a Sheathless Capillary Electrophoresis-Electrospray Ionization Mass Spectrometry Platform for Peptide Analysis: Comparison to Liquid Chromatography-Electrospray Ionization Mass Spectrometry; Analytical Chemistry, (2011); 83: 7297-7305.

[19] Sun, L.L., Zhu, G.J., Zhao, Y.M., Yan, X.J., Mou, S., Dovichi, N.J.; Ultrasensitive and Fast Bottom-up Analysis of Femtogram Amounts of Complex Proteome Digests; Angewandte Chemie-Internation Edition, (2013); 52: 13661-13664.

[20] Sun, L., Hebert, A.S., Yan, X., Zhao, Y., Westphall, M.S., Rush, M.J.P., Zhu, G., Champion, M.M., Coon, J.J., Dovichi, N.J.; Over 10 000 Peptide Identifications from the HeLa Proteome by Using Single-Shot Capillary Zone Electrophoresis Combined with Tandem Mass Spectrometry; Angewandte Chemie-Internation Edition, (2014); 53: 13931-13933.

[21] Han, X., Wang, Y., Aslanian, A., Bern, M., Lavallee-Adam, M., Yates, J.R.; Sheathless Capillary Electrophoresis-Tandem Mass Spectrometry for Top-Down Characterization of Pyrococcus furiosus Proteins on a Proteome Scale; Analytical Chemistry, (2014); 86: 11006-11012.

[22] Faserl, K., Kremser, L., Müller, M., Teis, D., Lindner, H.H.; Quantitative Proteomics Using Ultralow Flow Capillary Electrophoresis–Mass Spectrometry; Analytical Chemistrty, (2015); 87: 4633-4640.

[23] Sickmann, A., Reinders, J., Wagner, Y., Joppich, C., Zahedi, R., Meyer, H.E., Schonfisch, B., Perschil, I., Chacinska, A., Guiard, B., Rehling, P., Pfanner, N., Meisinger, C.; The proteome of Saccharomyces cerevisiae mitochondria; Proceedings of the National Academy of Sciences of the United States of America, (2003); 100: 13207-13212.

[24] Muller, B., Grossniklaus, U.; Model organisms - A historical perspective; Journal of Proteomics, (2010); 73: 2054-2063.

[25] Castrillo, J.O., Oliver, S.G.; Yeast as a touchstone in post-genomic research: Strategies for integrative analysis in functional genomics; Journal of Biochemistry and Molecular Biology, (2004); 37: 93-106.

[26] Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J.D., Jacq, C., Johnston, M., Louis, E.J., Mewes, H.W., Murakami, Y., Philippsen, P., Tettelin, H., Oliver, S.G.; Life with 6000 genes; Science, (1996); 274: 546-552.

[27] Pereira, C.V., Moreira, A.C., Pereira, S.P., Machado, N.G., Carvalho, F.S., Sardao, V.A., Oliveira, P.J.; Investigating drug-induced mitochondrial toxicity: a biosensor to increase drug savety?; Current Drug Safety, (2009); 4: 34-54.

[28] Benard, G., Rossignol, R.; Ultrastructure of the mitochondrion and its bearing on function and bioenergetics; Antioxidants & Redox Signaling, (2008); 10: 1313-1342.

[29] Yi, M.Q., Weaver, D., Hajnoczky, G.; Control of mitochondrial motility and distribution by the calcium signal: a homeostatic circuit; Journal of Cell Biology (2004); 167: 661-672.

[30] Zamsami, N., Kroemer, G.; Apoptosis-condensed matter in cell death; Nature, (1999); 401: 127-128.

[31] Moyle, J., Mitchell, P.; Active-inactive state transitions of mithochondrial atpase molecules influenced by mg2+, anion and aurovertin; Febs Letters, (1975); 56: 55-61.

[32] Gaucher, S.P., Taylor, S.W., Fahy, E., Zhang, B., Warnock, D.E., Ghosh, S.S., Gibson, B.W.; Expanded coverage of the human heart mitochondrial proteome using multidimensional liquid chromatography coupled with tandem mass spectrometry; Journal of Proteome Research, (2004); 3: 495-505.

[33] Prokisch, H., Scharfe, C., Camp, D.G., Xiao, W.Z., David, L., Andreoli, C., Monroe, M.E., Moore, R.J., Gritsenko, M.A., Kozany, C,. Hixson, K.K., Mottaz, H.M., Zischka, H., Ueffing, M., Herman, Z.S., Davis, R.W., Meitinger, T., Oefner, P.J., Smith, R.D., Steinmetz, L.M.; Integrative analysis of the mitochondrial proteome in yeast; PLoS Biology, (2004); 2: 795-804.

[34] Meisinger, C., Sickmann, A., Pfanner, N.; The mitochondrial proteom: From inventory to function; Cell, (2008); 134: 22-24.

[35] Reinders, J., Zahedi, R.P., Pfanner, N., Meisinger, C., Sickmann, A.; The complete yeast mitochondrial proteome: Multidimensional separation techniques for mitochondrial proteomics; Journal of Proteome Research, (2006); 5: 1543-1554.

[36] Tang, K.Q., Page, J.S., Smith, R.D.; Charge competition and the linear dynamic range of detection in electrospray ionization mass spectrometry; Journal of the American Society for Mass Spectrometry, (2004); 15: 1416-1423.

[37] Zhu, G.J., Sun, L.L., Yan, X.J., Dovichi, N.J.; Single-Shot Proteomics Using Capillary Zone Electrophoresis-Electrospray Ionization-Tandem Mass Spectrometry with Production of More than 1 250 Escherichia coli Peptide Identifications in a 50 min Separation; Analytical Chemistry, (2013); 85: 2569-2573.

[38] Moini, M.; Simplifying CE-MS operation. 2. Interfacing low-flow separation techniques to mass spectrometry using a porous tip; Analytical Chemistry (2007); 79: 4241-4246.

[39] Wang, Y., Fonslow, B.R., Wong, C.C.L., Nakorchevsky, A., Yates, J.R.; Improving the Comprehensiveness and Sensitivity of Sheath less Capillary Electrophoresis-Tandem Mass Spectrometry for Proteomic Analysis; Analytical Chemistry (2012); 84: 8505-8513.

[40] Gahoual, R., Busnel, J.M., Beck, A., François, Y.N., Leize-Wagner, E.; Full Antibody Primary Structure and Microvariant Characterization in a Single Injection Using Transient Isotachophoresis and Sheathless Capillary Electrophoresis–Tandem Mass Spectrometry; Analytical Chemistry, (2014); 86: 9074-9081.

[41] Gahoual, R., Biacchi, M., Chicher, J., Kuhn, L., Hammann, P., Beck, A., Leize-Wagner, E., Francois, Y.N.; Monoclonal antibodies biosimilarity assessment using transient isotachophoresis capillary zone electrophoresis-tandem mass spectrometry; Mabs, (2014); 6: 1464-1473.

[42] Sarg, B., Faserl, K., Kremser, L., Halfinger, B., Sebastiano, R., Lindner, H.H.; Comparing and Combining Capillary Electrophoresis Electrospray Ionization Mass Spectrometry and Nano-Liquid Chromatography Electrospray Ionization Mass Spectrometry for the Characterization of Post-translationally Modified Histones; Molecural & Cellular Proteomics, (2013); 12: 2640-2656.

[43] Knight, C.G., Kassen, R., Hebestreit, H., Rainey, P.B.; Global analysis of predicted proteomes: Functional adaptation of physical properties; Proceedings of the National Academy of Sciences of the United States of America, (2004); 101: 8390-8395.

[44] Dreger, M.; Subcellular proteomics Mass Spectrometry Review, (2003); 22: 27-56.

[45] L Lilley, K.S., Dupree, P.; Mant organelle proteomics; Current Opinion in Plant Biology, (2007); 10: 594-599.

[46] Mi, H.Y., Muruganujan, A., Thomas, P.D.; PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees; Nucleic Acids Research, (2013); 41: D377-D386.

[47] Chou, K.C., Wu, Z.C., Xiao, X.A.; iLoc-Euk: A Multi-Label Classifier for Predicting the Subcellular Localization of Singleplex and Multiplex Eukaryotic Proteins; Plos One, (2011); 6: 146-153.

[48] Renvoise, M., Bonhomme, L., Davanture, M., Valot, B., Zivy, M., Lemaire, C.; Quantitative variations of the mitochondrial proteome and phosphoproteome during fermentative and respiratory growth in Saccharomyces cerevisiae; Journal of Proteomics, (2014); 106: 140-150.

[49] Ben-Menachem, R., Tal, M., Shadur, T., Pines, O.; A third of the yeast mitochondrial proteome is dual localized: A question of evolution; Proteomics, (2011); 11: 4468-4476.

[50] Zahedi, R.P., Sickmann, A., Boehm, A.M., Winkler, C., Zufall, N., Schonfisch, B., Guiard, B., Pfanner, N., Meisinger, C.; Proteomic analysis of the yeast mitochondrial outer membrane reveals accumulation of a subclass of preproteins; Molecular Biology of the Cell, (2006); 17: 1436-1450.

[51] Huh, W.K., Falvo, J.V., Gerke, L.C., Carroll, A.S., Howson, R.W., Weissman, J.S., O'Shea, E.K.; Global analysis of protein localization in budding yeast; Nature, (2003); 425: 686-691.

[52] Ohlmeier, S., Kastaniotis, A.J., Hiltunen, J.K., Bergmann, U.; The yeast mitochondrial proteome, a study of fermentative and respiratory growth; Journal of Biological Chemistry, (2004); 279: 3956-3979.

[53] Reijans, M., Lascaris, R., Groeneger, A.O., Wittenberg, A., Wesselink, E., van Oeveren, J., de Wit, E., Boorsma, A., Voetdijk, B., van der Spek, H., Grivell, L.A., Simons, G.; Quantitative comparison of cDNA-AFLP, microarrays, and GeneChip expression data in Saccharomyces cerevisiae; Genomics, (2003); 82: 606-618.

[54] Marc, P., Margeot, A., Devaux, F., Blugeon, C., Corral-Debrinski, M., Jacq, C.; Genome-wide analysis of mRNAs targeted to yeast mitochondria; Embo Reports, (2002); 3: 159-164.

[55] Snyder, M., Kumar, A.; Yeast genomics: past, present, and future promise; Functional & Integrative Genomics, (2002); 2: 135-137.

[56] Dimmer, K.S., Fritz, S., Fuchs, F., Messerschmitt, M., Weinbach, N., Neupert, W., Westermann, B.; Genetic basis of mitochondrial function and morphology in Saccharomyces cerevisiae; Molecular Biology of the Cell, (2002); 13: 847-853.

[57] Steinmetz, L.M., Scharfe, C., Deutschbauer, A.M., Mokranjac, D., Herman, Z.S., Jones, T., Chu, A.M., Giaever, G., Prokisch, H., Oefner, P.J., Davis, R.W.; Systematic screen for human disease genes in yeast; Nature Genetics, (2002); 31: 400-404.

[58] von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S.G., Fields, S., Bork, P.; Comparative assessment of large-scale data sets of protein-protein interactions; Nature, (2002); 417: 399-403.

[59] Pflieger, D., Le Caer, J.P., Lemaire, C., Bernard, B.A., Dujardin, G., Rossier, J.; Systematic identification of mitochondrial proteins by LC-MS/MS; Analytical Chemistry (2002); 74: 2400-2406.

[60] DeRisi, J.L., Iyer V.R., Brown, P.O.; Exploring the metabolic and genetic control of gene expression on a genomic scale; Science, (1997); 278: 680-686.

[61] Brandina, I., Graham, J., Lemaitre-Guillier, C., Entelis, N., Krasheninnkov, I., Sweetlove, L., Tarassov, I., Martin, R.P.; Enolase takes part in a macromolecular complex associated to mitochondria in yeast; Biochimica et Biophysica Acta, (2006); 1757: 1217-1228.

3. Conclusion

Le but de l'étude décrite dans ce chapitre, a été de démontrer l'amélioration en termes d'identification protéique d'un extrait mitochondrial de *Saccharomyces cerevisiae* en combinant l'analyse par CZE-MS/MS et NanoLC-MS/MS à l'aide d'une approche protéomique. Afin de comparer correctement les données de ces méthodes, nous les avons analysés suivant différentes propriétés physico-chimiques des peptides identifiés : le poids moléculaire, le ratio charge sur masse, le point isoélectrique, la couverture de séquence ainsi que la qualité spectrale MS/MS. Les résultats présentés dans cette étude portant sur l'analyse d'une faible quantité d'échantillon d'extrait mitochondriale de levure (100 ng), ont montré que la CESI-MS/MS permet d'obtenir une augmentation du nombre de peptides uniques identifiés par rapport à la NanoLC-MS/MS et ainsi une amélioration du recouvrement de séquence des protéines. Cette amélioration est en adéquation avec l'évolution constante des lignes directrices de la protéomique actuelle notamment en termes de confiance sur l'identification des protéines. L'analyse plus précise des résultats obtenus sur les propriétés des peptides, montre qu'une des tendances les plus visibles concerne leur poids moléculaires. En effet, plus les peptides ont un poids moléculaire supérieur à 2000 Da, plus ils seront identifiés en CESI-MS, alors que seulement certain d'entre eux seront détectés en NanoLC-MS/MS. Concernant l'analyse des résultats suivant le pI, nos résultats indiquent que les peptides dont les valeurs de pI sont extrêmes (pI<4 et pI>10) sont détectés principalement par le CESI-MS/MS. Ces observations ouvrent la voie à l'utilisation de cette méthode pour la détection de peptides portant des modifications influant sur la valeur du pI tel que les phosphopeptides, sans nécessité de techniques d'enrichissement. Enfin, l'analyse des résultats en termes de qualités spectrales avec la prise en considérations des scores Mascot et Paragon, a montré que l'augmentation de la couverture de séquence en CESI-MS/MS ne se faisant pas au détriment de la qualité spectrale.

L'un des grands défis de l'analyse protéomique concerne la caractérisation d'isoformes de protéines. L'existence d'isoformes de protéine peut être expliquée par divers phénomènes parmi lesquels les multiples copies de gêne (variation allélique), les modifications post-traductionnelles ou les produits de dégradation. Dans cette étude, nous avons démontré que deux isoenzymes enolase de levure ont été caractérisées par CESI-MS/MS. En effet, la détection de peptide protéolitique discriminant est facilitée lorsqu'un grand nombre de précurseurs avec une haute qualité spectrale est généré. Traditionnellement, l'identification d'isoformes de protéines est obtenue en réalisant la combinaison de différentes digestions enzymatique ou à l'aide d'approches complémentaires à la protémique classique tel que le

top-down ou le middle-down. L'apport de la CESI-MS/MS à une étude en NanoLC-MS/MS apparait donc comme une bonne alternative pour déchiffrer un protéome complexe. L'évaluation de ces deux méthodes séparatives couplées directement à la MS a permis de mettre en évidence une forte complémentarité pour l'identification de peptides et de protéines ouvrant la voie à l'analyse d'autres organismes et plus largement en protéomique subcellulaire.

III - Développement et évaluation du couplage CE-UV/MALDI-MS pour l'analyse de protéines intactes et d'applications protéomiques.

1. Introduction

Les protéines et les peptides jouent un rôle primordial dans les voies de régulation biologiques. Ces dernières années, la caractérisation de ce type de biomolécules représente un des plus grands défis dans de nombreux domaines de recherche tels que la biologie ou la pharmacie. Depuis 1986, et l'approbation du Murumomab-CD3 par la « Food and Drug Administration » (FDA), les anticorps monoclonaux (mAbs) ont pris une part majeur sur le marché des produits développés par l'industrie biopharmaceutique. Ce type de composés, en constante croissance en termes de développement, sont particulièrement intéressant car ils présentent des propriétés thérapeutiques favorables aussi bien d'un point de vue pharmacocinétique que pharmacodynamique et peuvent de plus permettre la réduction d'effets secondaires. Les mAbs sont des glycoprotéines très complexes présentant potentiellement de nombreuses micro-hétérogénéités naturelles. Le besoin constant de caractérisations fines, précises et rapides pour cette famille de protéines est un défi permanant pour les méthodes d'analyse.

Aujourd'hui, la méthode de choix pour l'identification et la caractérisation de protéines est la MS généralement couplée avec une méthode séparative comme l'HPLC ou de l'électrophorèse en gel. Techniquement, la facilité de mise en œuvre de l'HPLC/MS ajouté à la grande robustesse du système fait de ce couplage une référence en termes de caractérisation protéique. Cependant, une méthode ne permet pas de répondre aux besoins infinis d'une telle caractérisation. Une alternative majeure au couplage HPLC/MS est l'utilisation de la CE en tant que méthode séparative en remplacement de l'HPLC. La CE est une méthode électrophorétique en veine liquide permettant de réaliser des séparations rapides, de très hautes efficacités et hautes résolutions. Cependant, les propriétés intrinsèques de la CE comme des électrolytes supports (BGE) de fortes forces ioniques ainsi que la nécessité de maintenir le courant tout au long de la séparation, font du couplage CE/MS un système compliqué à mettre en place. En effet, le détecteur MS est très sensible à la présence de sels qui peut provoquer des phénomènes de suppressions d'ions. Néanmoins, malgré ces aspects techniques, la CE/MS apparaît comme un couplage très attrayant

permettant de réaliser des avancées significatives dans de nombreux domaines de recherche tels que la protéomique et la caractérisation de protéines entières.

Depuis la fin des années 1980, un grand nombre d'interfaces CE/MS ont été développées. Dû à son caractère direct, c'est-à-dire intégré à la source MS, le couplage CE/ESI-MS a été fortement implanté. Afin de permettre le maintien du courant tout au long de la séparation, trois types d'interfaces basées sur des principes différents ont été décrits : l'interface coaxiale à liquide additionnel dite « Sheath liquid », l'interface à jonction liquide dite « Junction liquid » et l'interfaces sans liquide additionnel dite « Sheathless ». Cependant, l'une des limites majeures du couplage CE/ESI-MS est que la nature du BGE (présence de sel, volatilité) peut affecter fortement la stabilité du spray et donc limiter le nombre de milieu de séparations compatibles avec l'ESI-MS. La source MALDI est une technologie MS qui présente une meilleure tolérance en présence de sels ou de détergents. Le couplage CE/MALDI-MS représente donc une alternative intéressante au couplage CE/ESI-MS. Tout comme le couplage avec l'ESI, une grande variété d'interfaces ont été développées. Concernant les procédés permettant de maintenir le courant au long de la séparation, la quasi-totalité des interfaces reprennent les principes physiques du couplage de la CE avec une source ESI, à savoir l'utilisation d'un liquide additionnel coaxial, la jonction liquide ou l'utilisation d'un capillaire conducteur. En ce qui concerne la réalisation de dépôt solide, principe fondamentale du MALDI, cet impératif a conduit au développement de deux types de couplage distincts : le couplage direct avec un dépôt à l'intérieur de la source suivi d'une étape de cristallisation intégrée et le couplage indirect avec une étape de dépôt sur cible hors de la source et une étape de passage manuelle ou automatisé de la cible dans la source. Quelques équipes de recherche ont réalisé le couplage direct en faisant un lourd développement instrumental, notamment en modifiant la source MALDI. Ce paramètre n'a pas aidé au développement de ce type de couplage. Parallèlement au couplage direct, le couplage indirect a été réalisé. Cette stratégie, plus facile à mettre en œuvre, ne fait pas intervenir de modifications instrumentales sur le MALDI-MS. De plus, le caractère indirect ouvre de nouvelles voies dans le domaine du traitement de l'échantillon. En effet, le fait de fractionner l'échantillon permet de le retraiter (enrichissement, digestion enzymatique, dérivation chimique…) indépendamment avant analyse MALDI ou même de le déposer sur différent supports : plaque ELISA, capillaire creux, membrane greffée ou cible MALDI. Cependant, malgré le fait que la quasi-totalité des interfaces CE/MALDI-MS aient été évaluées en termes de d'efficacité et de robustesse, aucune d'entre elles n'est automatisable et commercialisée ce qui rend ce couplage confidentiel par rapport au couplage LC/MS. Malgré cela, un grand nombre de publications ont été recensées montrant l'intérêt grandissant pour ce couplage particulièrement pour l'étude de protéines entières ou digérées.

Dans ce travail, nous avons développé le premier couplage CE-UV/MALDI-MS/MS avec système de dépôt de matrice totalement automatisé. En effet, plusieurs évolutions ont été réalisées pour rendre ce couplage robuste, autonome et polyvalent. Pour ce faire, nous nous sommes basés sur la modification de deux systèmes commerciaux. Afin de contrôler la séparation électrophorétique, la cellule UV originelle de l'appareil de CE a été déportée pour permettre simultanément la détection UV et la collection de fraction. Ceci permet d'éviter l'ajout d'un spectrophotomètre UV externe. De plus, basé sur la propriété de l'appareil de CE à permettre l'application simultanée d'une différence de potentiel et de pression et dans le but de développer un système permettant de déposer la matrice directement après le dépôt de l'échantillon, une modification de la cartouche CE a été optimisée afin de rendre le couplage universel en termes de mode de dépôt MALDI. Ainsi, la séparation, le suivi UV, la collecte de fraction et l'adjonction de matrice sont automatisés et pilotés par seulement deux logiciels. Le caractère polyvalent du système est conféré par la possibilité de déposer sur tout type de support et d'élaborer tout type de dépôt quel que soit la stratégie analytique.

Dans un premier temps, les performances du système ont été évaluées avec la séparation d'un mélange de 5 protéines entières. Les résultats obtenus montrent la haute résolution et la grande répétabilité de séparation conservés en utilisant le couplage CE-UV/MALDI-MS. Dans un second temps, le système a été évalué sur un digestat trypsique de 9 protéines. Un grand nombre de peptides ont été identifiés à l'aide des spectres MS/MS permettant la caractérisation des protéines avec une couverture de séquence > 60%. Une comparaison avec le couplage référence HPLC/MS a été réalisé sur l'approche de peptide mapping. Cela a permis de montré une très grande complémentarité entre les deux approches. Enfin, dans le but de confirmer les performances du système pour la séparation de protéines entières de haut poids moléculaire, l'application à la séparation de mAbs intacts et digérés a été réalisée à l'aide du couplage CE-UV/MALDI-MS. Ce travail représente la première analyse de mAb intactes à l'aide d'un couplage de la CE avec une détection MS. Concernant l'étude du digestat trypsique de mAb, l'étude a permis d'obtenir la quasi-totalité de la couverture de séquence ainsi que la caractérisation de plusieurs formes glycosylées en une seule analyse.

Les résultats présentés dans cette partie ont fait l'objet d'une publication dans le journal scientifique *Electrophoresis* (Wiley 2014).

2. Analysis of monoclonal antibody by a novel CE-UV/MALDI-MS interface

Michael Biacchi, Ricky Bhajun, Nassur Saïd, Alain Beck, Yannis Nicolas François, Emmanuelle Leize-Wagner ***Electrophoresis*** 2014, 35(20), 2986-2995.

Abstract

Monoclonal antibodies (mAb) are highly complex proteins that advertise a wide range of microheterogeneity that requires multiple analytical methods for full structure assessment and quality control. As a consequence, the characterization of mAbs on different levels is particularly product - and time - consuming. Capillary Electrophoresis-Mass Spectrometry (CE-MS) couplings, especially to matrix-assisted laser desorption/ionization (MALDI) appear really attractive methods for the characterization of biological samples. In this work, we report on the last instrumental development and performance of the first totally automated off-line CE-UV/MALDI-MS/MS. This interface is based on the removal of the original UV cell of the CE apparatus, the modification of the spotting device geometry and the creation of an integrated delivery matrix system. The performance of the method was evaluated with separations of 5intact proteins and a tryptic digest mixture of 9 proteins. Intact proteins application shows the acquisition of electropherograms with high resolution and high repeatability. In the peptide mapping approach, a total number of 154 unique identified peptides were characterized using MS/MS spectra corresponding to average sequence coverage of 64.1%. Comparison with NanoLC/MALDI-MS/MS showed complementarity at the peptide level with an increase by 42% when using CE/MALDI-MS coupling. Finally, this work represents the first analysis of intact mAb charge variants by capillary zone electrophoresis using a MS detection. Moreover, using a peptide mapping approach CE-UV/MALDI-MS/MS fragmentation allowed 100% sequence coverage of the light chain and 92 % on the heavy chain and the separation of the 4 major glycosylated peptides and their structural characterization.

Introduction

The characterization of complex protein and peptide mixtures represents one of the biggest challenge in many research fields such as biological [1-3]or biopharmaceutical sciences [4-5].Since 1986 and the approbation of Muromonab-CD3 by the Food and Drug Administration (FDA), monoclonal antibodies (mAbs) have taken a major market share in the pharmaceutical industry and their development is constantly increasing [6-7]. mAbs are particularly interesting because they have good therapeutic efficiency, favorable pharmacokinetic and pharmacodynamics, and lead to reduction of side-effects [8]. mAbs are highly complex glycoproteins potentially displaying many naturally-occurring molecular micro-heterogeneities combined with imperfect processing. There is a continuous need for analytical methods improvement to be ultimately able to provide a fast and accurate characterization. Today, the method of choice for the identification of proteins and peptides is the mass spectrometry (MS) [9] generally coupled with separation techniques such as high-performance liquid chromatography (HPLC) or gel electrophoresis. Technically, HPLC-MS is the most widely used coupling thanks to its high reproducibility and its great resolution but also especially for the relative ease of coupling with a fraction collector device or directly in the source of a mass spectrometer. As an alternative to HPLC, capillary electrophoresis (CE) has been demonstrated to be a useful and powerful separation method for the characterization of charged and neutral molecules [10-12]. Advantages of CE are fast separation, high efficiency and high resolution. However, properties of CE such as an ultralow flow rate, a background electrolyte (BGE) that can be highly salted, and the necessity of maintaining the electric field during the separation make the CE-MS coupling more difficult to implement [13-14]. Nevertheless, despite these technical aspects, the hyphenation of CE to MS appears as a very attractive coupling allowing to perform significant advances in many research areas such as proteomics [15-16], metabolomics [17] and intact proteins characterization [18-19]. Since the end of the 80s, a large number of CE-MS interfaces has been developed. Due to the direct coupling property, most innovation have been realized on the coupling between CE and electrospray ionization (ESI) [20] by sheath-flow [21-22], liquid-junction [23-24], or sheathless [25-26] interfaces. However, the nature of the BGE (presence of salts, volatility) can affect the stability of the spray and then limits the number of BGE which does not prevent the performances of MS. Matrix-assisted laser desorption/ionization time of flight MS (MALDI-TOF-MS) is a MS technology which presents a greater tolerance to the presence of salts or surfactants [22]. Moreover, hyphenation of electrophoretic methods with MALDI-MS allows easily to detect in positive and negative modes in MS. Then, as an interesting alternative to CE/ESI-MS, a great number of CE/MALDI-MS interfaces have been developed [27-43]. The first strategy to hyphenate electrophoretic method with MALDI-MS has consisted to develop

on-line interface [30-32].However, because of sample is first dried on a solid surface before insertion into the MALDI source, heavy instrumental development appeared as one of the most significant issue for the implementation of this interface. The second option consisted to promote the off-line coupling using CE collection fractions on a MALDI plate. This strategy opens the way to sample treatment directly on the plate such as enzymatic digestion [44] or enrichment [45]. Concerning the deposit process, majority of the existing interfaces used a T-junction with sheath-flow [33-39]. More recently sheathless interfaces have been developed. In 2008, Wang et al. have realized an interface based on the use of a junction liquid which allows to eliminate the addition of the sheath liquid [42]. In 2009, Busnel et al. have designed an iontophoretic fraction collection approach to avoid the sheath-flow and the possible suction effect [43]. However, one of the interface drawbacks is having to fill little reservoir on the plate before deposit process. This prevents automation and high-throughput analysis. Concerning fraction collection devices, only one commercial system is compatible with CE/MALDI-MS [37]. However, this interface is originally designed for HPLC/MALDI-MS coupling and presents some geometry drawbacks for CE/MALDI-MS such as the minimum length of the capillary. At the same time, homemade robotic systems have been developed, such as electrospray [40] or inkjet spotting [41]. Meanwhile, since 2001 interfaces between MALDI-MS and micro fluidic devices have been developed [46]. The easiest strategy was reported to carry out CE separation in an open channel, where the separated analytes were directly ionized from the open channel to a mass spectrometer [47-48]. However, the open channel was filled by gel to suppress electroosmotic flow (EOF) and reduce analyte dispersion after the CE separation. In that case, only capillary gel electrophoresis is applicable. Then this strategy is not possible to regular free-solution CE modes including CZE.

The efficiency and the robustness of these interfaces have been reported, but none of these systems allows the automation. Most off-line CE/MALDI-MS coupling do not possess a second detection systems to control the separation. Few interfaces used an additional external UV system which made the methodology more expensive and made difficult the automation with the development of synchronized software [35-36].

In this work we have developed the first totally automated off-line CE-UV/MALDI-MS/MS with integrated delivery matrix system. Based on different homemade modifications performed on a commercial CE device and a fraction collection system, we introduced a totally automated approach to make protein separation and peptide mapping strategy. In order to control the CE separation, the original UV cell of the CE apparatus was deported to allow the simultaneous UV detection and fraction collection without external UV system. Modification of the spotting device geometry allowed to reduce the total capillary length. Based on the properties of the CE apparatus to apply simultaneously a pressure and a power supply, and the modification on

the external cartridge, an integrated homemade delivery matrix system was developed to be universally compatible in terms of MALDI deposition types. The performance of the method was evaluated with the separations of five intact proteins and a tryptic digest mixture of nine proteins. Finally, in order to confirm the performances of the system for the characterization of mAbs, application to the separation of intact mAb and bottom up approach have been realized using the CE-UV/MALDI-MS interface.

Materials and methods

Chemicals and Materials. All chemicals used were of analytical grade or high purity grade and purchased from Sigma-Aldrich (Saint Louis, MO, USA). Water used to prepare buffers and sample solutions was obtained using an ELGA purelab UHQ PS water purification system (Bucks, UK). Cytochrome c (Cyt c), ribonuclease A ((RNase A), α-lactalbumin (αLac), insulin (Ins), myoglobin (Myo), bovine serum albumin (BSA), lysozyme (Lys), β-Casein (βCas), carbonic anhydrase II (CAII) were obtained from Sigma-Aldrich (Saint Louis, MO, USA). Trypsin (modified, sequencing grade) was purchased from Promega AG (Dübendorf, Switzerland).

Standard protein mixtures. For intact protein study, a standard protein solution of αLac (14.2kDa), RNase A (13.6kDa), Cyt c (12.3 kDa), Lys (14.3 kDa) and Myo (16.9 kDa) were prepared in Water at 120nM, 50nM, 40 nM, 40nM and 50nMrespectively.

Tryptic digestion procedure. Tryptic digests of protein solution were prepared using sequence-grade trypsin (1:10 w/w). Before digestion, each protein was diluted to a concentration of 3µM in 1% formic acid. Dithiothreitol (DTT) was then added to the protein solution at a final concentration of 10 mM.The mixture was incubated for 5 minutes at 95°C, then cooled to ambient temperature. Then, Iodoacetamide (IAM) 100 mM was added at a final concentration of 20 mMand the mixture was placed in the dark for 20 minutes. 10 µL of trypsin was added and the mixture was incubated at room temperature for 3 hours; additional 10 µL of trypsin were added to the sample for overnight digestion at room temperature. This solution was then analyzed directly by CE/MALDI-MS without any further purification step. Amounts of protein tryptic digests with 300 fmol of each proteins were injected.

Capillary Electrophoresis. The CE experiments were carried out on a P/ACE MDQTMCE system from Beckman Coulter (Brea, CA) equipped with a temperature controlled autosampler and a power supply able to deliver up to 30 kV. A 32 KaratTM (Beckman Coulter, Brea, CA) was used for instrument control, data acquisition and data handling. Polymicro bare fused-silica capillaries of 50 µm i.d. were obtained from Photonlines (St-Germain-en-Laye, France).

Concerning modified capillaries, capillaries were coated in laboratory with hydroxypropylcellulose (HPC) following the protocol described by Shen et al. [49]. For model protein separation, 83.3mM ionic strength ammonium acetate (pH 4.0) has been used as BGE. Samples were introduced into the capillary by electrokinetic injections (3 kV, 8 min), and the separation of proteins was performed by applying 20 kV across the capillary. For intact mAbs separation, 400 mM ε-Amino-caproic acid (EACA)-acetic acid, 0.05% m/v, pH 5.7 has been used as BGE. For peptide mapping approach, solution of 1% formic acid and ammonium acetate (pH 9.3, 935 m Mionic strength) were used in CE as background electrolyte (BGE) and leading electrolyte (LE), respectively [50]. Except for the model proteins separation, injection volumes have been calculated using CEToolbox application (Pansanel, Google Play).

Nano-HPLC. Protein digests were analyzed using an UltiMate 3000 nano-HPLC system (Dionex, Germering, Germany) coupled to a Proteineer FC (Bruker Daltonics, Bremen, Germany). A Pepmap100 C18 column (Dionex, Germering, Germany) was used. The gradient (solvent A, 0.1% trifluofoacetic acid (TFA); solvent B, 0.1% TFA in 80% acetonitrile) was as follows: 0-4 min, 4% B; 4-32 min, 4-55% B; 32-34 min, 55-90% B; 34-40 min 90% B; 40-45 min, 90-4% B; 45-65 min, 4% B. A flow rate of 300nL/min was applied. The sample injection volume was 1 µL.

CE/MALDI-MS Interface. Automated off-line coupling of CE to MALDI-MS was performed by using a homemade modified automatic spotting device, Proteineer FC (BrukerDaltonics, Bremen, Germany) for the sheath flow-assisted spotting from the CE capillary end onto a MALDI target. The capillary outlet was positioned inside a fixed steel needle, being a modified part of a robotic x-y-z axis motion system. The original set-up delivered from Bruker Daltonics was modified by us in order to reduce the distance between CE output and the MALDI target. This change has consisted to shift the original steel needle close to CE output and has reduced the minimum total capillary length from110cm to 60cm. At the upper end of the needle a cross connection was fitted. On one side of the cross, the sheath liquid, which is the separation BGE, was supplied by means of the syringe-pump included in the Proteineer FC (Bruker Daltonics, Bremen, Germany) and on the other side, the ground was connected. The steel needle functioned as an electrode for the CE and the current was maintained by a liquid junction between the steel needle and the end of the capillary (Fig. 1). During CE separation, the emerging droplet of CE effluent and sheath liquid made contact with the MALDI target. The original set up of the UV cell in the P/ACE MDQTM (Beckman Coulter, Brea, CA) was modified in order to allow the simultaneous UV detection near the outlet of the capillary and fraction collection (Fig. 1). A homemade UV-window compatible with P/ACE MDQ system (Beckman Coulter, Brea, CA) was mounted at the upper of the cross connection. Using three optical fibers with one couple connected in pairs, UV-detection was deported out of the CE instrument

without instrumental modification and without additional UV spectrometer. Therefore, for CE with fraction collection, a capillary with the UV detection window positioned 10 cm from the CE capillary end was used.

In order to be universally compatible in terms of MALDI deposition types, a homemade delivery matrix system was developed to add the matrix just after sample deposition allowing dried droplet fraction collection directly onto the MALDI target. The matrix interface consisted of a second steel needle surrounding the matrix capillary (100cm total length, 100µm i.d.) in the form of a coaxial tube. Matrix capillary end was placed at a distance of 5 millimeters from CE capillary end corresponding to the distance between two consecutive spots of a MALDI target provided by Bruker Daltonics. To avoid the use of an additional syringe pump to deposit the matrix, a modification was made on the external cartridge delivered by Beckman Coulter. The external extremity of the cartridge was pierced to allow the setup of a second capillary. Separation and matrix capillaries were then placed in the external cartridge and positioned in the CE inlet for matrix capillary and the CE outlet for separation capillary. Properties of P/ACE MDQTM(Beckman Coulter, Brea, CA) allowing to applied at the same time a pressure and a power supply allowed performing simultaneously CE separation while delivery matrix. The interface is also compatible with a thin layer deposition types. In that case, matrix capillary was removed.Hystar 3.2 (BrukerDaltonics, Bremen, Germany) was used for Proteineer FC control.

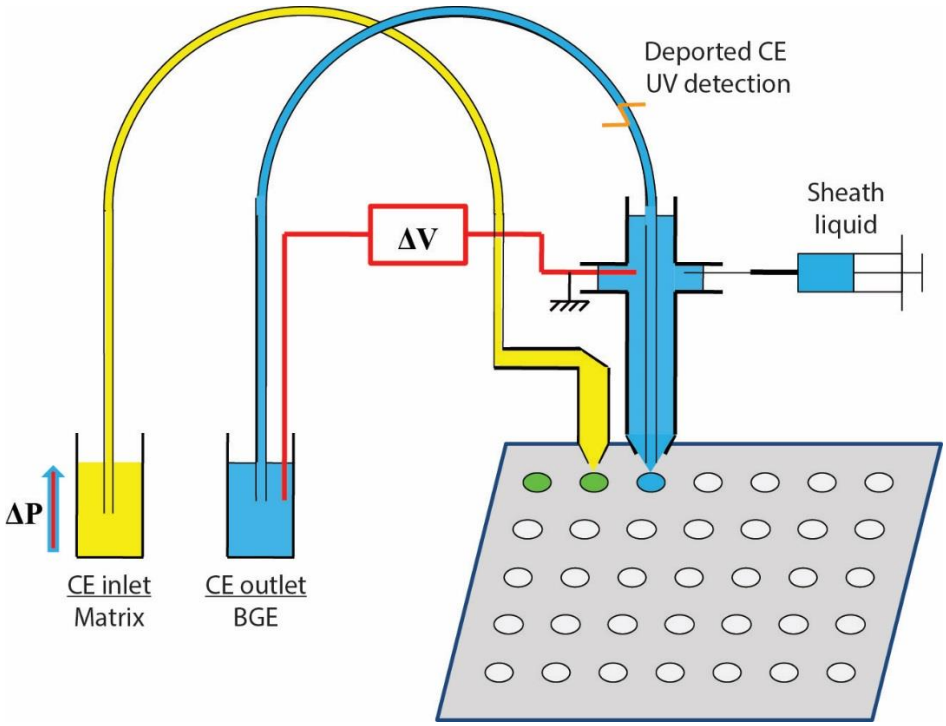

Figure 1. Schematic representation of the experimental set up of the CE-UV/MALDI-MS interface including the deported UV detection and the integrated delivery matrix system.

MALDI-TOF-MS and MALDI-TOF-TOF MS/MS. For intact protein and intact mAb separations, the matrix was prepared by dissolving 2,5-dihydroxybenzoïque (DHB) (2g.L-1) in 0.1% trifluoroacetic acid/acetonitrile (TFA/ACN) (50/50). Fraction collection fraction was realized using Ground Steel MALDI target (Bruker Daltonics, Bremen, Germany). For peptide mapping approach, α-cyano-4-hydroxycinnamic acid (CHCA)(5g/L) matrix was prepared following the classical AnchorChip protocol provided by Bruker Daltonics. Fraction collection was realized using AnchorChip MALDI target (Bruker Daltonics, Bremen, Germany). When the samples were directly collected on a MALDI target with 0.6µL sheath liquid, 0.6 µL of the matrix solution were automatically deposited on the same spot. Mass spectra of the CE fractions were recorded using an Autoflex III MALDI-TOF (Bruker Daltonics, Bremen, Germany), operating in reflector mode and with Flex Control software. Positively charged ions were detected and sums of 1500 single-shot spectra were acquired automatically from each sample by using the AutoXecute software. For peptide mapping study WarpLC 1.2 software was used. Data processing was performed with Flex Analysis 3.0 and Biotools 3.2 provided by the mass spectrometer manufacturer. All spectra were calibrated according an external calibration using Protein calibration standard I (Bruker Daltonics, Bremen, Germany) for intact protein separation and Peptide calibration standard II (Bruker Daltonics, Bremen, Germany) for peptide mapping study. Data obtained from CE-MALDI-MS/MS experiments were processed using Mascot search algorithm developed by Matrix Science (Boston, MA, USA).

Results and discussion

Evaluation of the CE-UV/MALDI-MS interface.

CE-UV/MALDI-MS interface has been evaluated in terms of repeatability and robustness following the optimized separation conditions described by the team of Prof. Girault [51]. As a first step, we have evaluated the external UV detection (Fig. 1) by studying the repeatability in terms of migration time, peak area and resolution, of a standard protein separation in CE/MALDI-MS configuration. This is a significant parameter to determine if the developed system is adapted to CE collection. Indeed, especially for the coupling between CE and MALDI-MS, the off-line nature requires a second detection system to control the separation and then to approve the fraction collection. Moreover, an electropherograms obtained with the external UV detection as a control allowed optimizing the spotting process in terms of deposit window and deposit time interval. A neutral coated capillary has been used with a standard protein mixture, containing αLac, RNase A, Cyt c, Lys and Myo at 120nM, 40nM, 50 nM, 50nM and 40nM respectively. Samples were introduced into the capillary by electrokinetic injections (3 kV, 8 min), and the separation of proteins was performed by applying 20 kV across the capillary. No fraction collection has been made (Fig. 2a); CE capillary tip was positioned in the

external BGE vial placed on the Proteineer FC (Bruker Daltonics, Bremen, Germany). In order to avoid Joule heating effect during the separation, the length of the cartridge tubing was optimized to perform thermostated capillary until the UV cell. The UV detection window is located at 10 cm from the outlet when the total length of the capillary is 60 cm. In term of migration time, peak area and resolution, obtained RSD (n=4) are less than 0.5%, 2.5% and 3% respectively. In addition, we have calculated signal-to-noise ratio ranging from 6 to 52. These results confirm that the externalization of the UV detection allows acquiring electropherograms with good quality, high repeatability and a good robustness. In addition this confirms the good stability of the HPC coating during the separation. Information concerning migration times allows defining the window of fraction collection, and assessing time interval for each deposit. These results confirm the possibility of obtaining CE-UV/MALDI-MS coupling without heavy instrumental development and without using an additional external UV detector. In addition, due to the use of single software to control CE apparatus, this UV modification contributes to strengthening the totally automated aspect of the coupling.

As a second step, in order to evaluate the impact of the spotting process on the separation, we have studied with the external UV detection the evolution of resolution and efficiency of the same standard protein mixture separation with fraction collection. For direct deposit on a MALDI plate, fraction collection window has been defined between 15 and 36 minutes and a spotting time interval of 30 s has been chosen.

To prove the automated spotting process, Fig. 2 shows the UV electropherograms with (Fig.2b) and without (Fig. 2a) fraction collection and in the case of collection the representation of the experimental fraction using MS spectra of each spotting position.

Concerning the deposition process, UV electropherogram (Fig.2b) shows a preservation of the resolution and efficiency as compared to the separation without fraction collection (Fig. 2a). The steel needle served as outlet electrode for the CZE. Liquid junction between steel needle and capillary tip maintained electrical current. Then, even if the steel needle moves on the *z* axis to reposition the capillary tip from a spotting position to the next, no current breakdown takes place throughout the separation and then electro migration never stops during the analysis. The localization of each protein, represented in the Fig. 2c, shows a time shift between the UV detection and the spotting process. This is totally in agreement by the fact that UV detection is located at 10 cm from the capillary outlet. Moreover, offsets in time, between 1 minute for Cyt c and 4-5 minutes for αLac, can be explained by the difference of protein effective mobilities described by the equation:

$$\mu_{eff} = \frac{L\,l}{t\,V}$$

with μ_{eff} the effective mobilities, L and l the total capillary length and length to detection window, respectively, V the applied voltage and t the migration time.

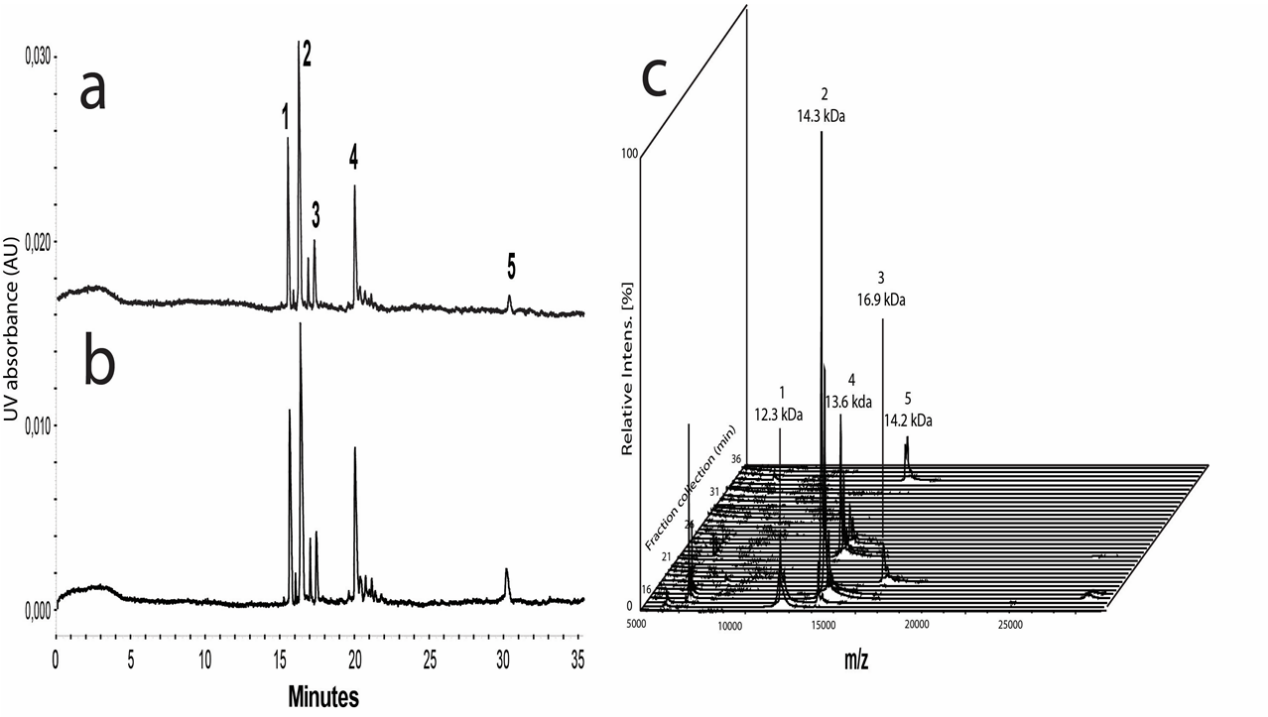

Figure 2. UV Electropherogram without (a) and with fraction collection (b) and the corresponding MS spectra of each fraction (c) of a five protein sample by CE-UV/MALDI-MS. Experimental conditions: HPC coated capillary, 50 µm d.i. x 60 cm (detection cell, 60 cm); BGE: 83.3mM ionic strength ammonium acetate (pH 4.0); Voltage: 20 kV; Temperature: 25°C; UV Detection: 200 nm; Injection: 3 kV, 8 min; Sample: (1) 40 nMCyt c, (2) 40 nM Lys, (3) 50 nMMyo, (4) 50 nMRNase A and (5) 120 nM αLac in water, MS Experimental conditions: See Experimental.

For example for RNase A, a time shift of 3 min between detection time (20 min) and spotting time (23 min) correspond to a value of µeff = 12.9.10-5 cm2. V-1.s-1.Good agreement between time shift values and effective mobilities showed that the interface does not induce any significant suction effect. This confirms the use of a sheath liquid junction. Under these experimental conditions, it appears that the automated interface supports intact protein separations. In addition, flexibility of spotting parameters (spotting interval, sheath liquid flow rate…) combine with UV control electropherograms allow high possibility of optimization as function of the sample complexity.

Evaluation of the CE-UV/MALDI-MS/MS for peptide mapping approach

Peptide mapping is essential for the characterization of a complex mixture of digested proteins. In comparison to peptide mass fingerprinting (PMF), MS/MS peptide identification is not only based on digested peptide m/z ratio measure, but also on the fragmentation of those peptides,

increasing the confidence of the identification. WarpLC1.2 (Bruker Daltonics, Bremen, Germany) is a software initially dedicated to perform peptide mapping strategy using off-line coupling between HPLC and MALDI-MS/MS. This software is fully compatible and could be easily adapted to the CE-UV/MALDI-MS/MS approach. Due to the complexity of biological samples, a low carry over may involve the presence of the same m/z ratio in two or three adjacent spots. In that case, the software selects the spot with the most intense m/z parent ion to perform MS/MS. After MS/MS spectra acquisition, WarpLC1.2 proceeds to the exclusion of the same m/z parent ion present in the adjacent spots. The software allows a drastic reduction in the possible redundant peptides and improves the number of selected m/z ratios which increase the number of possible unique peptides.

To further characterize the fraction collection interface, 100 nL of a peptide mixture sample containing the tryptic digests of βCas, RNase A, CAII, BSA, Lys, Ins, Cyt c, αLac and Myo (each at 3 µM) has been analyzed. Amounts of protein tryptic digests with300 fmol of each proteins were injected. In this study, a neutral coated capillary has been used and a transient isotachophoresis (t-ITP) preconcentration step has been implemented. For its compatibility with MALDI-MS, we used ammonium acetate (pH 9.3; 935mM ionic strength) as the leading electrolyte (LE) as in the Busnel et al. study [50]. Prior to the sample injection, a short zone of LE has been injected to allow t-ITP phenomena. A first separation without collection fraction allowed with UV electropherogram to fix the spotting parameters. At the beginning of the separation, for 25 minutes, the capillary tip was positioned in the external BGE vial placed on the Proteineer FC (Bruker Daltonics, Bremen, Germany).

According to electropherogram, during this time no peptide has been detected by the UV. Fraction collection began after 25 minutes with spotting intervals of 30 s. In order to perform MS/MS, a dried droplet deposition protocol has been involved. Using the homemade delivery matrix system, CHCA matrix was placed in the inlet position of P/ACE MDQTM system (Beckman Coulter, Brea, CA) and delivered by a second capillary with a flow rate of 1.2 µL/min in accordance to deposit 0.6 µL per spot. To determinate the exact value of the pressure to perform 1.2 µL/min, we followed the protocol described by Gahoual et al [3]. The separation was stopped after 65 min. A total of 80 fractions have been collected.

	CE-UV/MALDI-MS/MS		NanoLC/MALDI-MS/MS		Direct MALDI-MS		Combination of the three deposition modes	
	identified peptides	Sequencecoverage (%)	identified peptides	Sequencecoverage (%)	identified peptides	Sequencecoverage (%)	identified peptides	Sequencecoverage (%)
βCas	3	22.3	4	20.5	3	14.7	7	34.4
CAII	20	58.8	11	34.2	5	37.7	28	59.6
Cyt C	13	58.1	7	43.8	1	10.5	16	74.3
αLac	9	40.8	12	51.4	1	7.0	18	53.5
Lys	28	79.6	24	83.0	12	58.5	41	85.7
BSA	55	73.0	68	77.3	13	23.6	113	90.1
Myo	15	85.1	6	44.2	0	0	19	85.1
Ins	1	58.8	4	100	2	43.1	5	100
RNase A	10	100	25	93.5	8	54.4	32	100
	Total of id. pep.	**AverageSeq. Cov. (%)**	**Total of id. pep.**	**AverageSeq. Cov. (%)**	**Total of id. pep.**	**AverageSeq. Cov. (%)**	**Total of id. pep.**	**AverageSeq. Cov. (%)**
	154	64.1	161	60.8	45	27.7	279	75.9

Table 1. Number of Unique Identified Peptides and Sequence Coverage for a nine protein Tryptic digest using CE-UV/MALDI-MS-MS, NanoLC/MALDI-MS/MS, direct MALDI-MS and the combination of the three deposition modes. CE Experimental conditions: HPC coated capillary, 50 µm d.i. x 80 cm (detection cell, 70 cm); BGE: 1% formic acidacid, pH 2.1; Voltage: - 25 kV; Temperature: 25°C; UV Detection: 200 nm; Injection: 1 psi, 76 s (100nL injected); NanoLC and MS Experimental conditions; See Experimental; Sample: digests of βCas, RNase A, CAII, BSA, Lys, Ins, Cyt c, αLac and Myo

In order to evaluate the strength of the present CE-UV/MALDI-MS/MS coupling, the same 9 proteins digest loading amounts (300 fmol of each protein) were applied for both NanoLC/MALDI-MS/MS and direct MALDI-MS. For LC-MS/MS analysis, NanoLC coupled with the same spotting device was used. Sample were separated onto a reverse phase C18 analytical column for a 65 min gradient and deposited on the MALDI target with a 30s spotting intervals. Dried droplet deposition protocol has been performed in order to be similar to CE-MS/MS analysis. For direct MALDI-MS deposition, a 9 proteins mixture (each at 300 fmol)

without separation process has been deposited in the same MALDI target using the same dried droplet protocol. Table 1 summarizes the number of unique identified peptides and the sequence coverage for each protein using CE-UV/MALDI-MS/MS, NanoLC/MALDI-MS/MS, direct MALDI-MS and the combination of the three deposition modes.

When the number of peptides identified with the three techniques is compared, CE-UV/MALDI-MS/MS and NanoLC/MALDI-MS/MS appeared to obtain identical performances and are superior to direct MALDI-MS. While the direct MALDI-MS analysis allowed the identification of 45 peptides, CE-MS/MS analysis and NanoLC-MS/MS analysis resulted respectively in the identification of a total number of 154 different peptides ranging from 1 (Ins) to 55 (BSA) corresponding to average sequence coverage of 64.1%, and in the identification of a total number of 161 different peptides ranging from 1 (Cytc) to 68 (BSA) corresponding to average sequence coverage of 60.8%, with a mass accuracy of 20 ppm in MS and 0.5 Da in MS/MS. These results confirmed the usefulness of integrating a separation step prior to MS analysis. The results of the combination of the three deposition modes (Table 1) show a significant improvement of the total number of identified peptides, and of the average sequence coverage. Furthermore, we represented the merger of peptides from CE-MS/MS, NanoLC-MS/MS and direct MALDI-MS analysis (Fig.3). Results from the Venn comparison clearly show that CE-UV/MALDI-MS/MS and NanoLC/MALDI-MS/MS were complementarity methods at the peptide level with an increaseof 42% in terms of unique identified peptides from 161 (NanoLC analysis only) to 276 (NanoLCand CE analysis). This complementarity also was useful for improvement of sequence coverage (Table 1). Using the combination of both NanoLC and CE analysis, we obtained an average sequence coverage of 75.9 % as compared to 60.8% for single NanoLC analysis. This represented an increase of 20% in terms of average sequence coverage. This resulted in more confident protein identification.

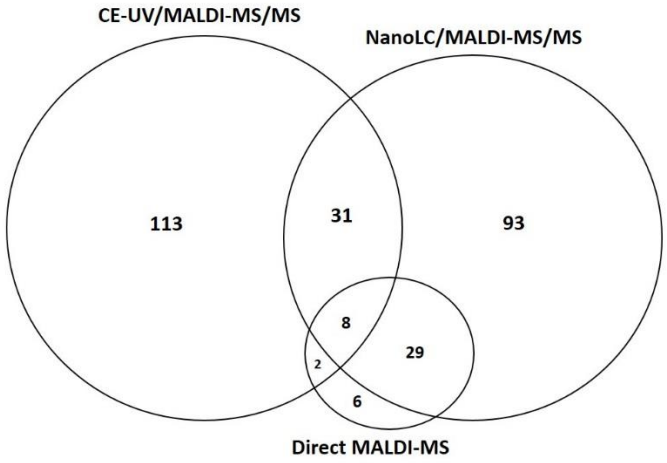

Figure 3. Venn comparisons representing the overlap of identified peptides by CE-UV/MALDI-MS/MS, NanoLC/MALDI-MS/MS and direct MALDI-MS. Data originate from Table 1.

Due to the ability of CE to separate a wide variety of peptides regardless of their chemical nature, these results unambiguously show the possibility of using CE-UV/MALDI-MS/MS to perform MS/MS peptide mapping characterization. Moreover, the automated setting up of a dried droplet deposition protocol which is essential in MALDI for MS/MS acquisition in the case of complex protein mixture, demonstrate the fully compatibility of the system for peptide mapping approach.

Analysis of mAb using CE-UV/MALDI-MS/MS.

mAbs are tetrameric glycoproteins having molecular mass of approximately 150 kDa. They are composed of two heavy chains (HC) and two light chains (LC) linked to each other by several disulfide bonds. This type of compounds presents a large number of micro heterogeneities commonly found in proteins such as PTMs including glycosylations and small chemical modifications [52]. In this work we performed the analysis of a model commercially marketed humanized mAb following two strategies: the separation of intact mAb and the peptide mapping approach both using CE-UV/MALDI-MS interface.

The use of capillary zone electrophoresis (CZE) as a tool for the analysis of charge heterogeneity of intact therapeutic mAbs was investigated by several groups [53-55]. However, due to the drastic condition in terms of salt concentration, no MS detection have been investigated. Indeed, CZE separation of intact proteins is highly affected by the pH and the ionic strength of the BGE. This represent a major drawback especially for the ESI-MS detection. Due to a greater tolerance to the presence of salts [22], MALDI-MS represent a good alternative. To demonstrate the CE-UV/MALDI-MS/MS fraction collection approach, an intact model humanized mAb has been considered. These experiments have been performed following the optimized condition described by the group of Dr. Ruesch [53].BGE contained 400 mM EACA-acetic acid, 0.05% m/v, pH 5.7 has been used. He et al. demonstrated the influence of the pH and the EACA concentration on the mAb separation. As charge variants of mAbs have different acid-base dissociation constants but similar Stokes' radii, pH buffer has to be selected such that the analytes exhibit significant relative difference in charge, increasing relative difference in electrophoretic mobility. Concerning EACA concentration, they demonstrated that increasing EACA concentration may improve separation efficiency and resolution. To avoid any protein adsorption to the capillary wall, HPC neutral capillary coated has been performed [55].In addition, concentrated EACA buffer also was used to suppress protein adsorption to the capillary wall and to interact with residual silanol groups that are not covered by HPC [53]. According to electropherograms represented Fig.4a, tip was positioned in the external BGE vial placed on the Proteineer FC (Bruker Daltonics, Bremen, Germany) for 45 min then fraction collection began with spotting intervals of 30 s. Fig.4b represent the reconstruction of the separation using MS spectra of each fractions.

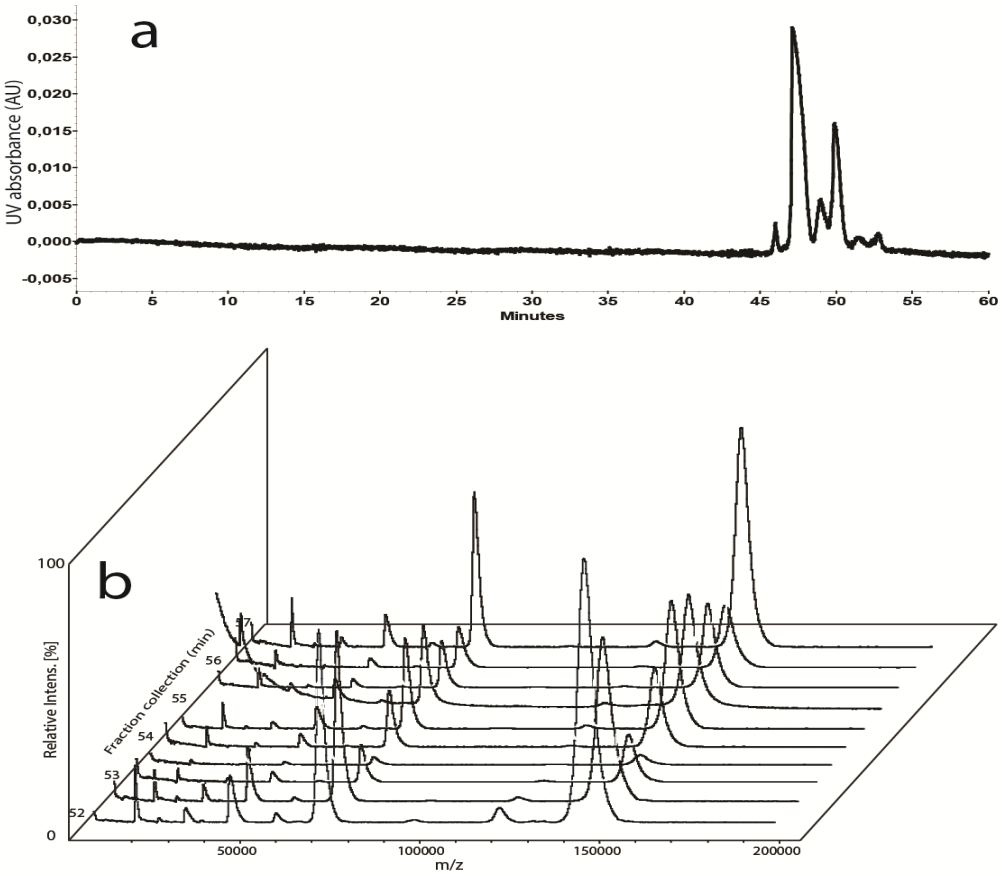

Figure 4.UV Electropherogram (a) and corresponding MS spectra of each fraction (b) of intact mAb charge variants by CE-UV/MALDI-MS.Experimental conditions: HPC coated capillary, 50 µm d.i. x 80 cm (detection cell, 70 cm); BGE: 400 mM EACA-acetic acid, 0.05% m/v, pH 5.7; Voltage: 30 kV; Temperature: 25°C; UV Detection: 200 nm; Injection: 2 psi, 30 s; Sample: mAb at 4 g/L, MS Experimental conditions: See Experimental.

First of all, these experiments showed the availability of the MALDI-MS to detect intact mAb using highly salted conditions. Moreover, the CE/MALDI-MS profile (Fig.4b) confirms the good agreement between the UV detection and the deposition time. This confirm the absence of carryover effect and diffusion phenomenon. Unfortunately, limitation of MALDI-MS resolution for molecules up to 100 kDa does not allow to measure the exact mass of the charge variants and then to characterize these glycoproteins. But this study is the proof of principle of the first analysis of intact mAb charge variants by CZE using a MS detection. This opens new pathways on the characterization of intact mAbs by CE-MS such as fraction collection, enrichment or Top-down approach.

The second strategy consisted to characterize mAb by bottom-up approach. As discussed previously, peptide mapping is commonly used to determine protein amino acid sequence. Moreover, this method allows to locate and/or quantify post translational modification (PTM). This methodology is quite important in early development of therapeutic antibodies as well as during long term life cycle management of the biopharmaceutical products. As part of this work, one of the objective was to investigate the peptide mapping of the same model humanized mAb by CE-UV/MALDI-MS/MS. A sample of mAbs was digested with trypsin enzyme following

the same protocol previously described. Fig.5a shows a separation obtained for the tryptic digest of mAb (500 fmol injected). The total duration of the separation was less than 50 minutes with a detection window of 20 minutes between 25 and 45 minutes.

In this work, peptide identification was automated and based on MS/MS data meaning that identification was performed on parent ion mass measure and fragment identification. Results of Mascot search showed sequence coverage of 92% for the HC and 100% for the LC of mAb corresponding to a total of 92 unique identified peptides (Fig. 5b). Some peptides were selected and fragmented several times during the analysis leading to the same identification. These phenomenon reinforce the confidence of the identification by enabling peptide overlapping.

mAbs are heterogeneous by nature due to different structural modification which may experience the protein during the lifetime. Those modifications can be Critical Quality Attributes (CQA) as they can change the protein conformation and therefore influence the antibody activity. mAb glycosylations represent a class of PTM which can significantly influence its structure, immunogenicity and stability [56]. Results obtained on the same set of data allowed to characterize the 4 major N-Glycosylation carried by the mAb (G0F, G1F, G2F and Man5). As expected by the work of Gahoual et al. [5], we detected glycans still linked to the corresponding peptide backbone of the HC and we confirmed the ability of CE migration mechanism to separate glycopeptides having a difference of only one galactose (Fig. 5a). The glycopeptide were detected between 34.1 and 36.1 min. Fig. 5c represents the MS/MS fragmentation spectra of the glycopeptide bearing G1F. As expected, pro duct ions observed in MS/MS mainly correspond to the fragmentation of the glycan moiety. Study of the fragmentation spectra allowed to deduce the structure of the glycan G1F. Such advanced characterization of digested mAb peptide mixture in a 50 min analysis shows the potential of CE-UV/MALDI-MS/MS as a fast and sensitive method in mAb characterization.

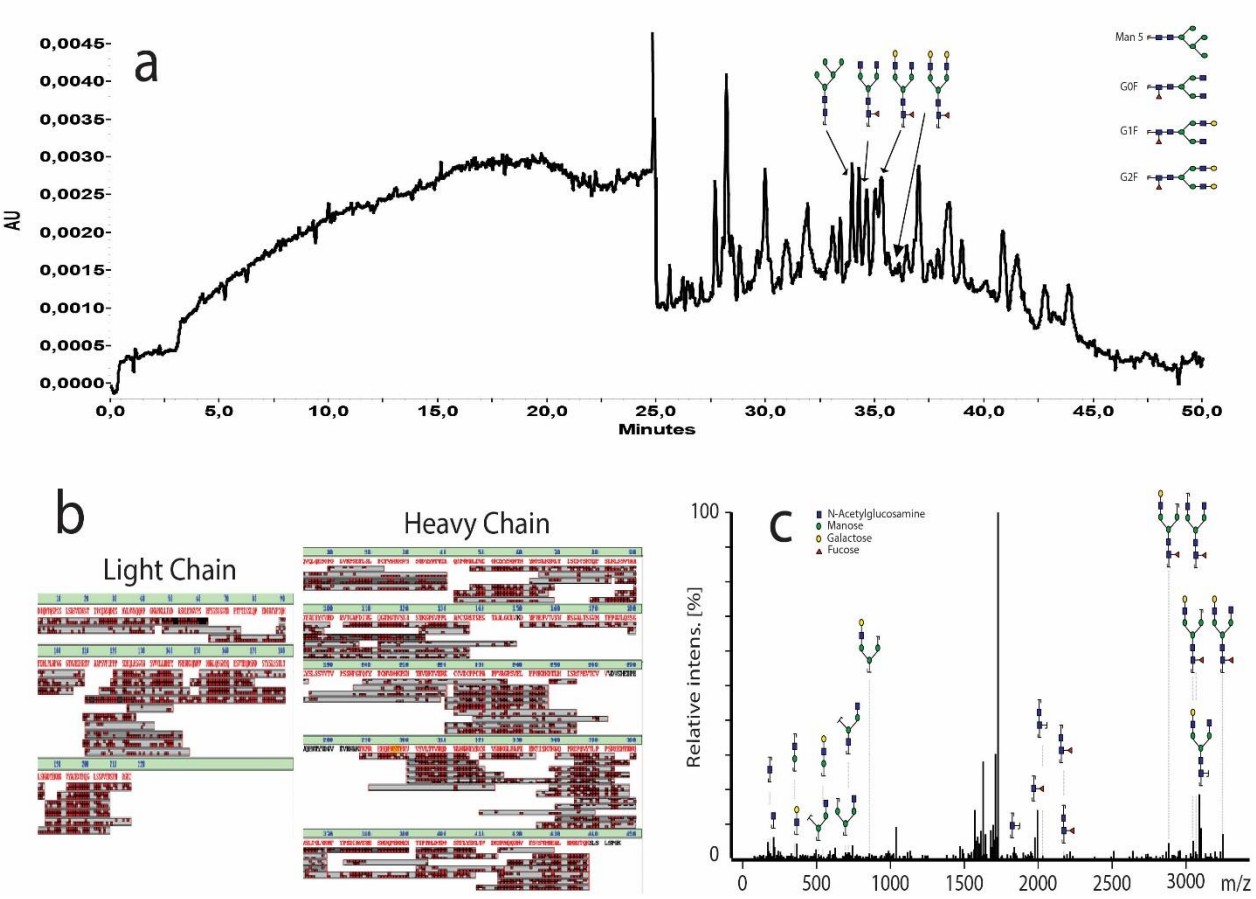

Figure 5.UV Electropherogram corresponding to the analysis by CE-UV/MALDI-MS/MS of mAbtryptic digest (a), Sequence coverage obtained for mAb HC and LC (b) and MS/MS fragmentation spectra of G1F glycopeptide (c). CE Experimental conditions and MS Experimental conditions: See Table 1 and Experimental.

Conclusion

On this work, we have developed the first automated off-line CE-UV/MALDI-MS/MS with integrated delivery matrix system. The system is based on different homemade modifications performed on a P/ACE MDQ^{TM} CE system (Beckman Coulter, Brea, CA) and a Proteineer FC automatic spotting device (Bruker Daltonics, Bremen, Germany). In order to control the CE separation, the original set up of the UV cell in the CE apparatus was deported to allow the simultaneous UV detection near the outlet of the capillary and fraction collection. In order to reduce the total capillary length, the spotting device was modified with a shift of the original steel needle close to the CE output. Minimum total capillary length can be 60 cm. Finally, in order to be universally compatible in terms of MALDI deposition modes, a homemade delivery matrix system was developed. Neutral coated capillary can therefore be used for the achievement of intact protein or peptide separation. The evaluation of the system on the separation of five proteins showed the acquisition of electropherograms with high resolution and high repeatability. Migration times allow defining the window of fraction collection, and

assessing time interval for each spot. Concerning peptide mapping approach on a nine digested proteins mixture, 154 unique identified peptides were characterized using MS/MS spectra corresponding to average sequence coverage of 64.8%. On the basis of the results obtained and as compared to NanoLC/MALDI-MS/MS, CE-UV/MALDI-MS/MS can be considered as a complementarity methods to conventional LC-MALDI/MS especially to improve the confidence in protein identification. Finally, this system has been used for the characterization of a humanized mAb following intact protein separation and peptide mapping strategies. In spite of the lack of resolution of the MALDI-MS, this work represents the first analysis of intact mAb charge variants by CZE using a MS detection and opens new ways of research for the characterization of intact mAbs. Moreover, the CE-UV/MALDI-MS/MS fragmentation allowed 100% sequence coverage of the LC and 92 % on the HC and the separation of the 4 major glycosylated peptides and their structural characterization in the same analysis data set. These results analysis demonstrates the power of CE-UV/MALDI-MS/MS as a sensitive method in mAb characterization and more generally in bottom-up proteomic approach.

Acknowledgments

Michael Biacchi would like to thank the ANFH for funding his Ph.D work. LSMIS would like to thank Dr. Bernard Capolaghi (Hospital Bel Air, Thionville) for his support. LSMIS would like to thank Paape Rainer, Pierre-Olivier Schmit and Anja Resemann from Bruker for their help. The authors would like also to express their gratitude Dr. E. Wagner-Rousset, Dr. D. Ayoub, MC. Janin-Bussat, O. Colas (Centre d'immunologie Pierre Fabre, St Julien en Genevois, France) and Rabah Gahoual (LSMIS, Strasbourg, France) for helpful discussions around antibody structural characterization. Dr. Yannis François would like to thank Dr. Jean-Marc Busnel for very fruitful discussions.

References

[1] Skalnikova, H.K.*Biochimie*2013, *95 (12)*, 2196-2211.

[2] Ma, N., Ling, R., Zaidah, L., Su, S.*Int.J.Mol.Sci.*2013, *14 (4)*, 7515-7541.

[3] Gahoual, R., Busnel, J.M., Wolff, P., Francois, Y.N., Leize-Wagner, E. *Anal.Bioanal.Chem.*2014, *406*, 1029-1038.

[4] Beck, A., Sanglier-Cianférani, S., Van Dorsselaer, A.*Anal. Chem.*2012, *84*, 4637-4646.

[5] Gahoual, R., Burr, A., Busnel, J.M., Kuhn, L., Hammann, P., Beck, A., François,Y.N., Leize-Wagner, E.*mAbs*2013,*5 (3)*, 479-490.

[6] Beck, A., Wurch, T., Bailly, C., Corvaia, N. Nat Rev *Immunol.* 2010, *10*, 345-352.

[7] Reichert, J.M.*mAbs* 2012,*4(3)*, 413-415.

[8] Wang, W., Singh, S., Zeng, D. L., King, K., Nema, S.*J.Pharm. Sci*. 2007, *96*, 1–26.

[9] Aebersold, R., Mann, M.*Nature*2003,*422*, 198-207.

[10] Krylov, S.N., Dovichi, N.*Anal. Chem.*2000, *72*, 111R-128R.

[11] Kasicka, V.*Electrophoresis*2010, *31*, 122-146.

[12] El Rassi, Z. *Electrophoresis*2010, *31*, 174-191.

[13] Smith, R.D., Udseth, H.R.*Nature*1988, *331*, 639-640.

[14] Rose, D.J., Jorgenson, J.W.*J.Chromatogr.A*1988, *438*, 23-34.

[15] Ahmed, F.E.*J.Chromatogr.B*2009, *877*, 1963-1981.

[16] Faserl, K., Sarg, B., Kremser, L., Lindner, H.*Anal. Chem.*2011,*83*, 7297-7305.

[17] Ramautar, R.,Mayboroda, O.A., Somsen,G.W., de Jong, G.J.*Electrophoresis*2011,*32*, 52-65.

[18] Haselberg, R, de Jong, G.J., Somsen, G.W. *Journal of Chromatography A* 2007, *1159*, 81-109.

[19] Haselberg, R., de Jong,G.J.,Somsen, G.W.*Electrophoresis*2013,*34*, 99-112.

[20] Ramautar, R., Heemskerk, A.A.M., Hensbergen, P.J., Deelder, A.M., Busnel, J.M., Mayboroda, O.A.*J.Proteomics*2012,*75*, 3814-3828.

[21] Smith, R.D., Olivares, J.A., Nguyen, N.T.,Udseth, H.R.*Anal. Chem.*1988,*60*, 436-441.

[22] Hommerson, P., Khan, A.M., de Jong, G.J.,Somsen, G.W.*Mass.Spectrom.Rev.*2011,*30*, 1096-1120.

[23] Wachs, T., Sheppard,R.L., Henion, J.*J.Chromatogr.BBiomed.Appl*.1996,*685*, 335-342.

[24] Zhu, G., Sun, L., Yan, X.,Dovichi, N.*Anal. Chem.*2013,*85*, 2569-2573.

[25] Moini, M.*Anal. Chem.*2007,*79*, 4241-4246.

[26] Busnel, J.M., Schoenmaker, B., Ramautar, R., Carrasco-Pancorbo, A., Ratnayake, C., Feitelson, J.S., Chapmann, J.D., Deelder, A.M., Mayboroda, O.A. *Anal. Chem.*2010,*82*, 9476-9483.

[27] Chang, S.Y., Yeung, E.S., *Anal. Chem.* 1997, *69*, 2251-2257.

[28] Murray, K.K., *Mass. Spectr. Rev.* 1997, 16, 283-299.

[29] Mechref, Y., *Capillary Electrophoresis and Microchip Capillary Electrophoresis: Principles, Applications, and Limitations*, Wiley, 2013, 367-380.

[30] Preisler, J., Hu, P., Rejtar, T., Karger, B.L. *Anal. Chem.* 2000, *72*, 4785-4795.

[31] Preisler, J., Foret, F., Karger, B.L. *Anal. Chem.* 1998, *70*, 5278-5287.

[32] Musyimi, H.K., Narcisse, D.A., Zhang, X., Stryjewski, W., Soper, S.A., Murray, K.K. *Anal. Chem.* 2004, *76*, 5968-5973.

[33] Amon, S., Plematl, A., Rizzi, A. *Electrophoresis* 2006, *27*, 1209-1219.

[34] Lechner, M., Seifner, A., Rizzi, A.*Electrophoresis*2008, *29*, 1974-1984.

[35] Silvertand, L.H.H., SastreToraño, J., de Jong, G.J., van Bennekom, W.P.*Electrophoresis*2009, *30*, 1828-1835.

[36] Weiss, N.G., Zwick, N.L., Hayes, M.A.*J.Chromatogr.A*2010, *1217*, 179-182.

[37] Johnson, T., Bergquist, J., Ekman, R.,Nordhoff, E., Schürenberg, M., Klöppel, K.D., Müller, M., Lehrach, H.,Gobom, J.*Anal. Chem.*2001,*73*, 1670-1675.

[38] Wang, J., Ye, H., Zhang, Z., Xiang, F., Girdaukas, G., Li, L.*Anal. Chem.*2011,*83*, 3462-3469.

[39] Bachmann, S., Bakry, R., Huck, C.W., Polato, F., Corradini, D., Bonn, G.K.*Electrophoresis*2011,*32*, 2830-383.

[40] Hanton, S.D.,Hyder, I.Z., Stets, J.R., Owens, K.G., Blair, W.R., Guttman, C.M.,Giuseppetti,A.A.*J.Am. Soc. Mass.Spectr.*2004, *15*, 168-179.

[41] Vannatta, M.W., Whitmore, C.D., Dovichi, N.J.*Electrophoresis*2009, *30*, 4071-4074.

[42] Wang, J., Ma, M., Chen, R., Li, L., *Anal*. *Chem*. 2008, *80*, 6168-6177.

[43] Busnel, J.M., Josserand, J., Lion, N., Girault, H.H.*Anal. Chem.*2009, *81*, 3867-3872.

[44] Li, Y., Yun, B., Deng, C., Tang, J., Liu, J., Zhang, X.*Proteomics*2007, *7*, 3661-3671.

[45] Bi, H., Qiao, L., Busnel, J.M., Devaud, V., Liu, B., Girault, H.H.*Anal. Chem.*2009, *80*, 1177-1183.

[46] Luo, Y.Q., Xu, S.Y., Schiling, J.W., Lau, K.H., Whitin, J.C., Yu, T.T.S., Cohen, H.H. *JALA*, 2009, *14*, 252-261.

[47] Liu, J., Tseng, K., Garcia, B., Lebrilla, C. B., Mukerjee, E., Collins, S., Smith, R. *Anal. Chem.* 2001, *73*, 2147-2151.

[48] Xu, Y., Little, M.W., Murray, K.K. *J. Am. Soc. Mass Spectrom.* 2006, *17*, 469-474.

[49] Shen, Y., Smith, R.D.*J.Microcolumn.Sep.*2000,*12*, 135-141.

[50] Busnel, J.M., Lion, N., Girault, H.H.*Electrophoresis*2008,*29*, 1565-1572.

[51] Pourhaghighi, M.R., Busnel, J.M., Girault, H.H.*Electrophoresis*2011, *32*, 1795-1803.

[52] Walsh, G., Jefferis, R. *Nat.Biotech.* 2006, *24*, 1241-1252.

[53] He, Y, Lacher, N.A., Hou, W., Wang, Q., Isele, C., Starkey, J., Ruesch, M. *Anal. Chem.* 2010, *82*, 3222-3230.

[54] He, Y., Isele, C., Hou, W., Ruesch, M. *J. Sep. Sci*. 2011, *34*, 548-555.

[55] Gassner, A.L., Rudaz, S., Schappler, J. *Electrophoresis*2013, *34*, 2718-2724.

[56] Jefferis, R.*Biotechnol.Prog.* 2005,*21*, 11-16.

3. Conclusion

Lors ce travail, nous avons développé le premier couplage CE-UV/MALDI-MS/MS avec système de dépôt de matrice totalement automatisé. En effet, plusieurs évolutions ont été réalisées pour rendre ce couplage robuste, autonome et polyvalent. Pour ce faire, nous nous sommes basés sur la modification de deux systèmes commerciaux. Afin de contrôler la séparation électrophorétique, la cellule UV originelle de l'appareil de CE a été déportée pour permettre simultanément la détection UV et la collection de fraction. Ceci permet d'éviter l'ajout d'un spectrophotomètre UV externe. De plus, basé sur la propriété de l'appareil de CE à permettre l'application simultanée d'une différence de potentiel et de pression et dans le but de développer un système permettant de déposer la matrice directement après le dépôt de l'échantillon, une modification de la cartouche CE a été optimisée afin de rendre le couplage universel en termes de mode de dépôt MALDI. Ainsi, la séparation, le suivi UV, la collecte de fraction et l'adjonction de matrice sont automatisés et pilotés par seulement deux logiciels. Le caractère polyvalent du système est conféré par la possibilité de déposer sur tout type de support et d'élaborer tout type de dépôt quel que soit la stratégie analytique.

Les performances du système ont été évaluées avec la séparation d'un mélange de 5 protéines entières ainsi qu'un digestat tryspique de 9 protéines. La séparation de protéines entières montre la robustesse du système en termes d'efficacité, de résolution et de temps de migrations. Dans l'approche protéomique, un nombre total de 154 peptides uniques ont été caractérisés par MS/MS correspondant à une couverture de séquence moyenne de 64,1%. Une comparaison avec le couplage de référence HPLC/MS a été réalisé et a montré une très grande complémentarité entre les deux approches représentées par l'augmentation de 42% au niveau du nombre de peptides identifiés. Enfin, dans le but de confirmer les performances du système pour la séparation de protéines entières de haut poids moléculaire, l'application à la séparation de mAbs intacts et digérés ont été réalisées à l'aide du couplage CE-UV/MALDI-MS. En dépit de l'absence de résolution du MALDI-MS, ce travail représente la première analyse des variants de charge de mAbs intactes à l'aide d'un couplage de la CE avec une détection MS et ouvre de nouvelles voies de recherche pour la caractérisation des mAbs intactes tel que l'approche Top Down. Concernant l'analyse du digestat trypsique de mAbs, l'étude a permis d'obtenir 100 % de couverture de séquence de la chaine légère et 92 % de la chaine lourde ainsi que la séparation et la caractérisation des quatre formes majoritaires de glycosylation en une seules analyse. Ces résultats démontrent la puissance du couplage CE-UV/MALDI-MS/MS comme une méthode sensible et robuste pour la caractérisation de mAbs.

PARTIE II : Caractérisation de la structure primaire d'anticorps monoclonaux et évaluation de biosimilarité à l'aide du couplage

CESI-MS

I - Caractérisation rapide et multi-niveaux d'anticorps monoclonaux par CESI-MS/MS

1. Introduction

L'avènement des anticorps monoclonaux (mAbs) a débuté en 1986 avec l'autorisation de mise sur le marché de muromomab-CD3 par la Food and Drug Administration (FDA) comme traitement contre les rejets de greffes du rein. Les mAbs thérapeutiques ont depuis pris une part de marché significative dans l'industrie pharmaceutique et leurs développements sont en constante augmentation. Ce succès s'explique en partie par la grande spécificité de l'anticorps pour son antigène, ce qui ouvre de nouvelles possibilités en termes de voie de traitement, en particulier en oncologie. Les mAbs possèdent également une pharmacodynamique (PD) et pharmacocinétique (PK) favorable pour une utilisation en tant que traitement thérapeutique (leur demi-vie dans le plasma est de 21 jours).

Les mAbs sont des glycoprotéines d'une grande complexité qui peuvent abriter un nombre conséquent de micro-hétérogénéités dont les sources sont variées (procédés de production inadapté, conservation prolongée). La complexité de ces protéines et leur utilisation en tant que traitement thérapeutique justifient la nécessité d'améliorer les méthodologies analytiques de caractérisation. Ceci de manière à obtenir une caractérisation rapide et précise de ce type de molécules. La spectrométrie de masse (MS) a su rapidement s'imposer comme une méthode de choix pour la caractérisation des mAbs, notamment en raison de son excellente spécificité, sensibilité mais aussi car elle permet d'obtenir des informations structurales dans certain cas. La MS peut ainsi être utilisée dans le développement de nouveaux mAbs ou encore pour suivre différents lots d'un mAb afin de distinguer des différences mineures. L'enjeu dans un tel contexte est de fournir rapidement des informations structurales permettant en amont d'orienter les activités de R&D (optimisation du procédé de production, élimination d'un candidat non satisfaisant) et de production (libération de lots). Les méthodes analytiques utilisées doivent ainsi être capables de couvrir un large champ de la structure des mAbs : séquence d'acides aminés, glycosylation, modifications post-traductionnelles (PTMs), ponts disulfures. En amont de l'analyse par MS, il convient d'utiliser un certain nombre de méthodes orthogonales destinées à caractériser la protéine sur les différents niveaux qui définissent sa structure. Ainsi la MS est généralement couplée à la chromatographie liquide en phase inverse (RP-LC), la chromatographie d'échange d'ions (IEC) ou encore la chromatographie d'exclusion stérique (SEC).

La caractérisation de la séquence d'acides aminés (*peptide mapping*) des mAbs est communément réalisée en utilisant la dégradation d'Edman ou encore une stratégie de type protéomique *bottom-up* car en plus d'obtenir des informations importantes concernant la séquence d'acides aminés, elle permet dans une certaine mesure d'avoir accès aux PTMs.

Les travaux présentés dans cette partie décrivent le développement d'une méthode de CE-ESI-MS/MS, utilisant le système CESI-MS, permettant la caractérisation rapide et précise d'un digestat de mAb. Le mAb étudié dans le cadre de ces travaux est trastuzumab (Herceptin®, Roche). Ce mAb, dont la commercialisation a été approuvée entre 1998 et 2000, est utilisé comme traitement contre le cancer du sein HER positif. Trastuzumab est une immunoglobulin gamma 1 humanisée (HzIgG1) dirigé contre le récepteur HER2/neu. II est composé de deux chaînes lourdes (HC) et deux chaînes légères (LC) identiques qui comptent respectivement 449 et 214 acides aminés. La HC du mAb porte un site de N-glycosylation situé sur l'acide aminé asparagine 300 (Asn^{300}). L'ensemble formant une glycoprotéine tétramérique dont la masse se situe aux alentours de 148 kDa. Trastuzumab est un mAb largement décrit dans la littérature et souvent considéré comme représentatif de ce type de molécules thérapeutiques.

Afin de réaliser la digestion trypsique des échantillons d'anticorps, un protocole de digestion en solution a été mis en place. Le digestat a ensuite été séparé et analysé par CESI-MS/MS afin d'obtenir la caractérisation de la protéine simultanément sur différents niveaux. Tout d'abord la caractérisation de la séquence d'acides aminés a été réalisée par l'intermédiaire des données CESI-MS/MS. L'identification des spectres MS/MS a été effectuée par l'utilisation d'un algorithme de recherche afin de diminuer le temps nécessaire au traitement des données. Les résultats obtenus ont permis, à partir des données issues d'une seule injection, d'obtenir le recouvrement complet de la séquence d'acides aminés de trastuzumab. L'ionisation des analytes délivrée par le système CESI-MS s'est avérée, comme évoquer dans la précédente partie, d'une grande efficacité. Cette caractéristique du couplage se répercute directement sur la qualité des spectres notamment en MS/MS : les spectres générés ont permis d'obtenir un grand nombre des fragments y et b, ce qui améliore de manière significative la confiance des identifications mais également fourni de précieuses informations sur l'enchainement précis des acides aminés le long du squelette peptidique.

Les PTMs ont également été étudiées par l'intermédiaire des données CESI-MS/MS. En particulier, les sites dits PTMs *hot-spots* ont été étudiés. Ces sites sont décrits dans la littérature comme étant induit par une dégradation de la protéine, influencent l'activité thérapeutique du mAb, et peuvent par conséquent être utilisés pour estimer l'état de conservation d'un échantillon de mAb. Par l'intermédiaire de la même analyse CESI-MS/MS que celle utilisée pour la caractérisation de la séquence d'acides aminés, il a été possible de

caractériser 6 PTMs d'intérêt de nature variées (cyclisation de l'acide glutamique en N-terminal, déamidation d'asparagine, oxidation de methionine). La qualité des données générées par la méthode développée permet en autre d'obtenir le spectre MS/MS du peptide modifié mais également de son homologue n'ayant pas subi la modification. Ceci permet d'en déduire à une modification partielle de l'échantillon ce qui informe sur la dégradation partielle de l'échantillon. De plus, les spectres MS/MS permettent de caractériser précisément l'acide aminé portant la modification.

Enfin les données de CESI-MS/MS ont été utilisées pour réaliser la caractérisation structurale de glycosylations les plus abondantes portées par trastuzumab. En utilisant le même set de données que précédemment, l'identification des 5 glycosylations les plus abondantes portées par trastuzumab a pu être obtenue. En particulier les spectres MS/MS montrent la fragmentation du glycanne, ce qui permet d'en déduire la structure de la glycosylation.

Les résultats présentés dans cette partie ont fait l'objet d'une publication dans le journal scientifique *mAbs (*Landes Bioscience 2013).

2. Rapid and multi-level characterization of trastuzumab using sheathless capillary electrophoresis-tandem mass spectrometry

Rabah Gahoual, Alicia Burr, Jean-Marc Busnel, Lauriane Kuhn, Phillipe Hammann, Alain Beck, Yannis Nicolas François, Emmanuelle Leize-Wagner. ***mAbs*** 2013, 5(3), 479-490.

Abstract

Monoclonal antibodies (mAbs) are highly complex proteins that display a wide range of microheterogeneity that requires multiple analytical methods for full structure assessment and quality control. As a consequence, the characterization of mAbs on different levels is particularly product - and time - consuming. This work presents the characterization of trastuzumab sequence using sheathless capillary electrophoresis (referred as CESI) – tandem mass spectrometry (CESI-MS/MS). Using this bottom-up proteomic-like approach, CESI-MS/MS provided 100% sequence coverage for both heavy and light chain via peptide fragment fingerprinting (PFF) identification. The result was accomplished in a single shot, corresponding to the analysis of 100 fmoles of digest. The same analysis also enabled precise characterization of the post-translational hot spots of trastuzumab, used as a representative widely marketed therapeutic mAb, including the structural confirmation of the five major N-glycoforms.

Introduction

Since 1986 and the approbation of muromonab-CD3 by the US Food and Drug Administration (FDA), monoclonal antibodies (mAbs) have taken a major market share in the pharmaceutical industry and their development is constantly increasing [1,2]. MAbs are highly complex glycoproteins potentially displaying many naturally-occurring molecular micro-heterogeneities and those that are introduced through imperfect processing, physico-chemical and enzymatic changes that occur during their production and long-term conservation. As a consequence, there is a continuous need for improvement of analytical methods to enable fast and accurate characterization. Mass spectrometry (MS) plays a major role in the characterization of therapeutic mAbs because it provides specificity, sensitivity and may also provide structural information. Characterization using MS has helped product and process development [3] and batch consistency assessment [4,5]. Furthermore, recent raising importance of biosimilar and biobetter mAbs represents a major stake for characterization methodologies. Biosimilar mAbs are produced from different clones and processes [6]. The European Medicines Agency (EMA) and the FDA have published draft guidelines concerning the approval pathway of biosimilar mAbs. These documents discuss the physicochemical characterization steps required to assess biosimilarity. For example, EMA guidelines suggest that the biosimilar candidate should have the same amino acid sequence as the reference mAb and a comparable glycosylation profile [7]. The selection of a highly similar candidate to the innovator product is mandatory before performing pre-clinical and clinical biosimilarity studies. Obtaining suitable structural information rapidly, which allows quick elimination of candidates dissimilar to the reference protein, is therefore crucial. The methods used should cover a wide range of structural information, e.g., amino acid sequence, disulfide bonds, glycosylation, posttranslational modifications (PTM) [8]. Additional to MS, mAbs characterization requires a panel of orthogonal methods to cover a full range of structural information [9], including reverse phase liquid chromatography (RP-LC) [10-12], ion exchange chromatography (IEC), size exclusion chromatography (SEC) [13] capillary isoelectric focusing (cIEF) [14] or micellar electrokinetic chromatography (MEKC) [15,16].

Peptide mapping of mAbs by MS is commonly performed using a bottom-up proteomic approach because it allows acquisition of information on the amino acid sequence and PTMs. In this approach, the protein undergoes enzymatic digestion, resulting in a peptide digest most often separated by liquid chromatography and detected by UV spectrophotometry and MS.

A novel sheathless CE-ESI-MS platform, referred as CESI, allowing hyphenation of capillary electrophoresis (CE) to electrospray ionization mass spectrometry (ESI-MS) has been developed by Beckman Coulter based on a design previously described by Moini [17]. This

CESI prototype (also known as sheathless capillary electrophoresis) uses a bare fused silica capillary whose outlet has been etched using hydrofluoric acid, which makes it porous to the electrical transport of small ions without permitting significant matter transfer through the pores. The porous tip provides electrical contact via a second capillary. Detailed description of this interface has been given by Haselberg et al [18]. The sheathless interface has already been used to perform successful CE-ESI-MS experiments in proteomics [19,20], metabolomics [21], intact proteins [22,23] demonstrating drastically increased sensitivity compared with sheath-liquid CE-MS due to operating flow rates of well below 100 nL/min. Whitmore and Gennaro recently used the CESI prototype to performed the peptide mass fingerprinting (PMF) of a mAb. They demonstrated the successful use of PMF to cover the amino sequence of a mAb, including the hydrophilic peptides that are eluted in the void volume in RP-LC-MS [23].

In this work, we used the CESI-MS system to develop a CE-MS/MS method that allowed the fast and precise characterization of a mAb digest. The mAb selected was trastuzumab which was approved by FDA and EMA, in 1998 and 2000 respectively. Trastuzumab has been intensively studied and can be considered as a highly representative therapeutic mAb [24]. It is a humanized immunoglobulin gamma 1 (HzIgG-1) directed against the HER2/neu receptor which is overexpressed in about 20–30% of invasive breast cancer patients [25,26]. Trastuzumab is composed of two heavy chains (HC) having 449 amino acids and two light chains (LC) having 214 amino acids [27]. The HC has an N-glycosylation consensus site on its asparagine (Asn) in position 300, and the glycoprotein has a total molecular mass of about 148 kDa [4]. A tryptic digest of trastuzumab was prepared using a conventional in-solution digestion protocol [28]. The digest was then studied at different levels. First a peptide mapping of the mAb was performed using a CESI-MS/MS analysis. The MS/MS spectra were automatically analyzed to identify the peptides using a search engine widely used in proteomic studies in order to hasten peptide identifications. PTMs were also characterized using CESI-MS/MS, particularly those known as degradation hot spots. These PTMs were identified by Mascot search algorithm and manually confirmed. At a final level, glycosylation were also investigated using CESI-MS/MS. Although glycosylation is a type of PTM, they are treated in this article separately from amino acid PTMs because they are not degradation hot spots. The glycans could be located on the amino acid sequence and the structure of the major glycoforms could be deduced from MS/MS data. Remarkably, the characterization of the mAb over these different described levels was realized using only one CESI-MS/MS analysis of the digest.

Results

Peptide mapping of trastuzumab using CESI-MS/MS. Peptide mapping is commonly used to determine protein amino acid sequence and to locate or quantify PTMs. This methodology is quite important in early development of therapeutic antibodies as well as during long-term life cycle management of the biopharmaceutical products. As part of this work, one of the objective was to investigate the peptide mapping of trastuzumab by CESI-MS/MS. Tandem MS data interpretation and peptide identification was done automatically using Mascot to obtain fast and accurate data treatment.

A sample of trastuzumab was digested with trypsin enzyme. Tryptic digestion was performed using an in-solution digestion protocol applied on a routine basis in the laboratory. It includes reduction and alkylation of disulfide bonds. Prior to digestion, an unfolding step was performed using guanidine as a chaotropic agent. Prior to MS analysis, peptides were separated using the CESI prototype which allows direct coupling of CE with the nanospray source of the MS without using an additional liquid like in sheath-liquid CE-MS [29].

Figure 1 illustrates an example of the separation obtained for the tryptic digest of trastuzumab for 100 fmol of injected protein. The total duration of the separation was less than 50 min and the resulting electropherogram showed that all peptides were detected between 19 and 47 min. Separation was performed at a voltage of +20 kV to maintain a balance between peptide migration speed and separation efficiency.

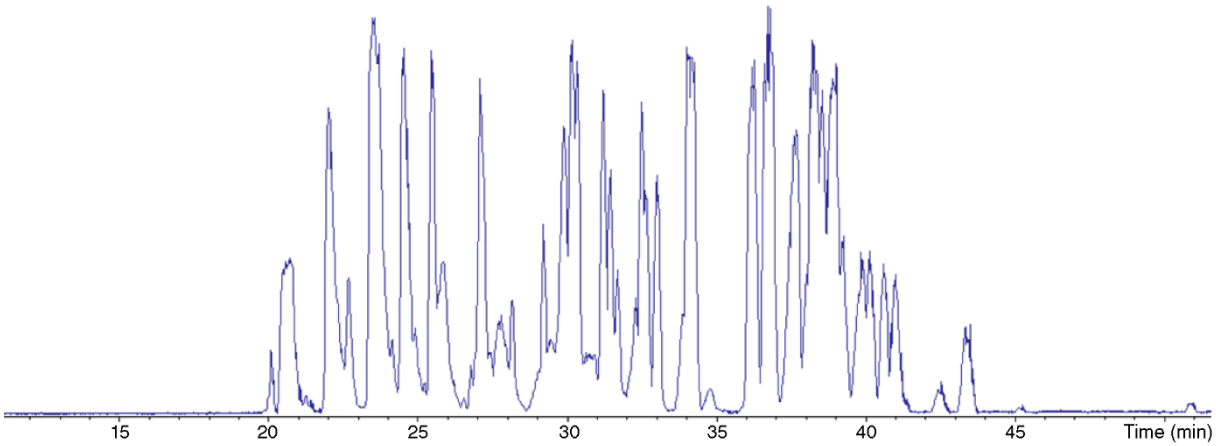

Figure 1. Base Peak Electropherogram corresponding to the analysis by CESI-MS/MS of trastuzumab tryptic digest. Experimental conditions: bare fused silica capillary with porous tip, total length 95 cm (30 µm i.d., 150 µm o.d.); CE voltage +20 kV; BGE 10% acetic acid; sample trastuzumab tryptic digest 2.5 mg/mL in BGE (11 nL injected). MS capillary voltage: -1.3 kV, m/z range: 50–3000.

Digested peptides were identified from MS/MS data acquired using a Mascot search algorithm, although 100% sequence coverage of a mAb has already been reported using the prototype CESI-MS platform [23]. In this work, peptide identification was automated and based on

MS/MS data, meaning that identification was made on precursor ion mass measure and fragment identification. Results of Mascot search showed sequence coverage of 100% for both the HC and LC of the trastuzumab. The detailed sequence coverages for both chains are shown in **Figure 2**. Some peptides were selected and fragmented several times during the analysis leading to the same identification. Similarly, missed cleavages could be promptly detected for several peptides reinforcing the confidence of the identification by enabling peptide overlapping.

Figure 2. Sequence coverage obtained by CESI-MS/MS for trastuzumab HC (left-hand side) and LC (right-hand side) based on the identified peptides. Experimental conditions: see materials and methods.

Tryptic digestion produces a heterogeneous mixture of peptides, especially with regard to their number of amino acids. Results show the identification during the analysis of small peptides having 4–5 amino acids such as HT05 or LT09 (**Fig. 3**) as well as a 63 amino acids peptide having a molecular mass of 6715.26 Da (HT21). During its elution, the ion corresponding to HT21 peptide was selected several time for fragmentation and resulted in the same identification, which reinforced the confidence of this important identification.

It should be noted that such heavy peptides usually can attain higher charge state; therefore, the search algorithm must consider high charge state as well. In the CESI-MS/MS analysis of the mAb, peptides having charge state up to 6+ could be detected. Fragmentation of the peptides enables an important part of the amino acid sequence to be retraced reinforcing the high confidence in peptide identification.

Heavy chain

	Peptide sequence	Position	Measured m/z
HT01	EVQLVESGGGLVQPGGSLR	1~19	941.501
HT02	LSCAASGFNIKDTYIHWVR	20~38	746.703
HT03	QAPGKGLEWVAR	39~50	656.355
HT04	IYPTNGYTR	51~59	543.265
HT05	YADSVK	60~65	341.670
HT06	YADSVKGR	60~67	448.229
HT07	YADSVKGRFTISADTSK	60~76	615.976
HT08	YADSVKGRFTISADTSKNTAYLQMNSLR	60~87	785.391
HT09	GRFTISADTSK	66~76	591.802
HT10	GRFTISADTSKNTAYLQMNSLR	66~87	825.411
HT11	GRFTISADTSKNTAYLQMNSLRAEDTAVYYCSR	66~98	948.442
HT12	FTISADTSKNTAYLQMNSLR	68~87	754.377
HT13	NTAYLQMNSLR	77~87	655.827
HT14	NTAYLQMNSLRAEDTAVYYCSR	77~98	876.066
HT15	AEDTAVYYCSR	88~98	667.784
HT16	WGGDGFYAMDYWGQGTLVTVSSASTK	99~124	929.089
HT17	WGGDGFYAMDYWGQGTLVTVSSASTKGPSVFPLAPSSK	99~136	988.975
HT18	GPSVFPLAPSSK	125~136	593.822
HT19	GPSVFPLAPSSKSTSGGTAALGCLVK	125~150	830.436
HT20	STSGGTAALGCLVK	137~150	661.34
HT21	DYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPS SSLGTQTYICNVNHKPSNTK	151~213	1343.868
HT22	DYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPS SSLGTQTYICNVNHKPSNTKVDK	151~216	1177.085
HT23	VDKKVEP	214~221	471.778
HT24	SCDKTHTCPPCPAPELLGGPSVFLFPPKPK	222~251	834.404
HT25	SCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISR	222~258	692.675
HT26	TPEVTCVVVDVSHEDPEVK	259~277	713.676
HT27	FNWYVDGVEVHNAK	278~291	559.932
HT28	FNWYVDGVEVHNAKTKPR	278~295	540.771
HT29	TKPREEQYNSTYR	292~304	557.934
HT30	TKPREEQYNSTYRVVSVLTVLHQDWLNGKEYK	292~323	777.198
HT31	VVSVLTVLHQDWLNGKEYK	305~323	743.4
HT32	VVSVLTVLHQDWLNGKEYKCK	305~325	839.44
HT33	VVSVLTVLHQDWLNGKEYKCKVSNK	305~329	737.139
HT34	CKVSNKALPAPIEK	324~337	518.953
HT35	VSNKALPAPIEK	326~337	422.914
HT36	TISKAKGQPREPQVYTLPPSR	338~358	589.073
HT37	AKGQPREPQVYTLPPSREEMTK	342~363	636.327
HT38	GQPREPQVYTLPPSREEMTK	344~363	781.722
HT39	GQPREPQVYTLPPSREEMTKNQVSLTCLVK	344~373	872.441
HT40	NQVSLTCLVK	364~373	581.315
HT41	GFYPSDIAVEWESNGQPENNYK	374~395	848.712
HT42	GFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSK	374~412	1467.351
HT43	GFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDK	374~417	1240.085
HT44	TTPPVLDSDGSFFLYSK	396~412	625.312
HT45	TTPPVLDSDGSFFLYSKLTVDK	396~419	810.75
HT46	LTVDKSR	413~419	409.735
HT47	LTVDKSRWQQGNVFSCSVMHEALHNHYTQK	413~442	901.175
HT48	SRWQQGNVFSCSVMHEALHNHYTQK	418~442	609.877
HT49	GQPREPQVYTLPPSR	344~358	1723.910
HT50	GQPREPQVYTLPPSREEMTK	344~363	2342.179
HT51	EPQVYTLPPSREEMTK	348~363	1903.943
HT52	EPQVYTLPPSREEMTKNQVSLTCLVK	348~373	3046.555
HT53	EEMTKNQVSLTCLVK	359~373	1778.900
HT54	EEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYK	359~395	4304.050
HT55	WQQGNVFSCSVMHEALHNHYTQK	420~442	934.752
HT56	SLSLSPG	443~449	660.355

Light chain

	Peptide sequence	Position	Measured m/z
LT01	DIQMTQSPSSLSASVGDR	1~18	626,963
LT02	DIQMTQSPSSLSASVGDRVTITCR	1~24	870,42
LT03	ASQDVNTAVAWYQQKPGK	25~42	763,056
LT04	LLIYSASFLYSGVPSR	46~61	886,779
LT05	LLIYSASFLYSGVPSRFSGSR	46~66	769,738
LT06	SGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTK	67~103	1047,983
LT07	SGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIK	67~107	1165,553
LT08	SGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIKR	67~108	963,859
LT09	VEIKR	104~108	322,703
LT10	RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPR	108~142	971,253
LT11	TVAAPSVFIFPPSDEQLK	109~126	973,513
LT12	TVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPR	109~142	745,982
LT13	TVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAK	109~145	811,817
LT14	VQWKVDNALQSGNSQESVTEQDSK	146~169	893,089
LT15	VQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSK	146~183	833,201
LT16	VDNALQSGNSQESVTEQDSK	150~169	712,656
LT17	VDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSK	150~183	905,928
LT18	VDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEK	150~188	1057,488
LT19	DSTYSLSSTLTLSK	170~183	751,879
LT20	ADYEKHKVYACEVTHQGLSSPVTK	184~207	916,441
LT21	HKVYACEVTHQGLSSPVTK	189~207	714,354
LT22	HKVYACEVTHQGLSSPVTKSFNRGEC	191~214	748,847
LT23	SFNRGEC	208~214	435,178

Figure 3. List of trastuzumab digested peptides identified by the CESI-MS/MS analysis. Experimental conditions: see material and methods.

As seen in **Figure 4**, the MS/MS spectra quality, has allowed the amino acids sequence composing the variable domain of the mAb HC for 109 of 120 amino acids to be retraced. Similarly 98 of the 107 amino acids sequence composing the variable domain of the LC could

be retraced using the MS/MS data. Thus, in addition to allowing the identification of the digested peptides of the mAb, the MS/MS spectra enabled characterization, without ambiguity, of most amino acids composing such a sensitive part of the mAb, i.e., the variable domain.

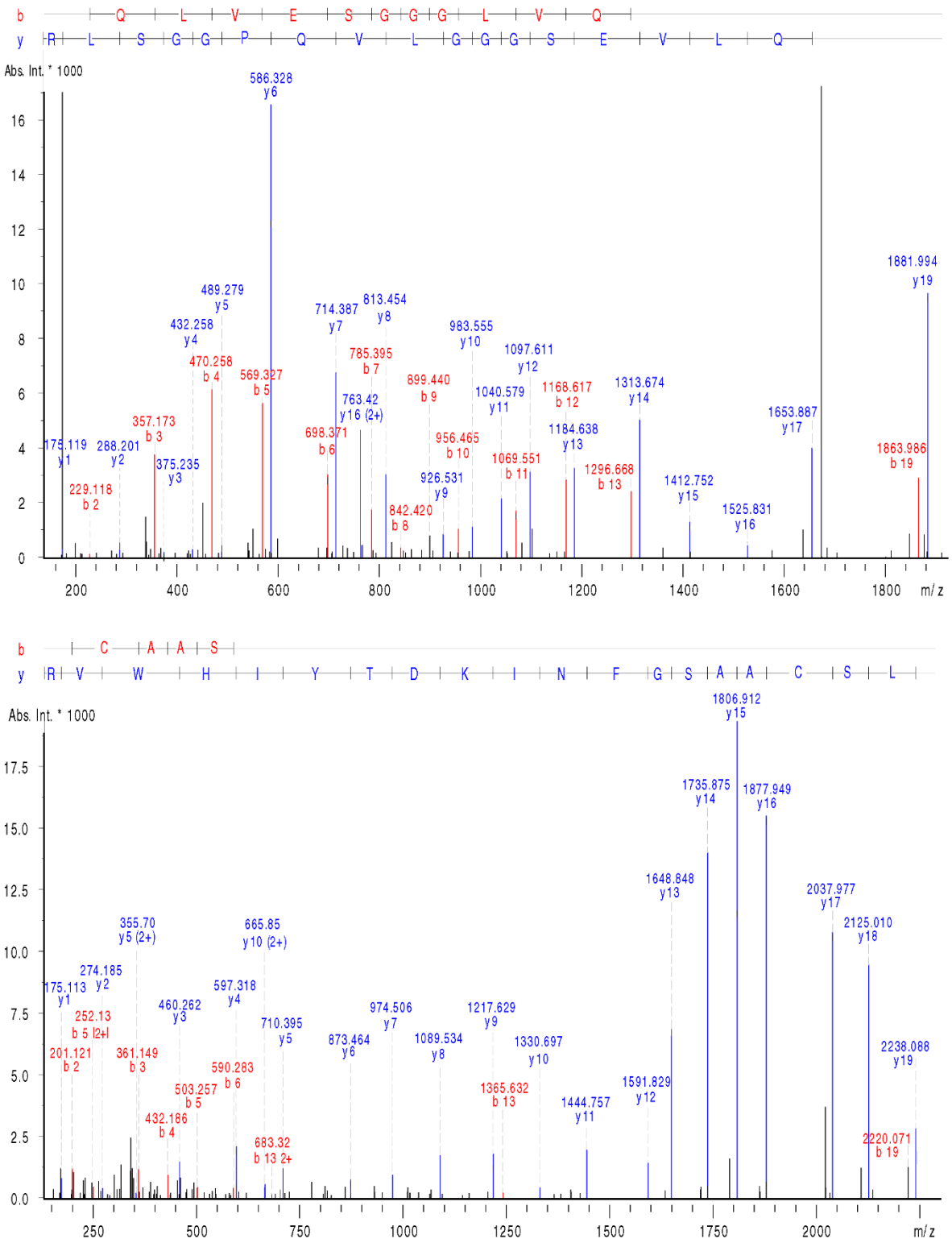

Figure 4. MS/MS spectra of peptide HT01 and HT02 illustrating the amino acid sequence retracing in the variable domain. Experimental conditions: see material and methods.

To evaluate the robustness of the method, the CESI-MS/MS analysis of the mAb digest was repeated six times consecutively. Results of these analysis demonstrated 100% sequence coverage for each run and the same detected peptides in each analysis.

Characterization of amino acids PTMs hot-spots. Mabs are heterogeneous by nature because of structural modifications that the proteins may undergo during their lifetime. Those modifications may be Critical Quality Attributes (CQA) because they can modify the conformation of the protein and therefore influence the antibody activity. Posttranslational modifications and other physico-chemical modifications can be induced by intra or extra cellular processes, buffer, purification process, storage or stress conditions [31, 32]. Among those PTMs, several referred to as “hot spots” must be carefully monitored to determine the stability of the mAb during pharmaceutical development. Known degradation hot spots potentially modified on the HC are cyclization of N-terminal glutamic acid (Glu) into pyroglutamic acid and deamidation of Asn^{55} and Asn^{387} but also oxidation of Met^{255} and Met^{431}. The main hot spot on the LC is the deamidation Asp^{30}. For the characterization of trastuzumab PTMs, the same CESI-MS/MS data presented in the previous section to obtain the mAb peptide mapping was used. In this case again, the PTMs were automatically identified by the Mascot search algorithm. Mass spectra were further manually studied to confirm the presence of the modifications.

Results obtained showed glutamine cyclization of the N-terminal extremity of the HC. Also, the same peptide without the modification was identified, which suggests a partial modification of the protein. Use of the extracted ion electropherograms (XIE) for m/z ratio of 932.500 (2+) and 941.501 (2+) corresponding to the peptide HT01 with both N-terminal extremity forms allowed the determination of a migration time of 34.5 min for the intact form of HT01 and 45.6 min for the modified HT01. This charge difference influences the electrophoretic mobility of the modified peptide, leading to its separation.

In the same manner, several deamidation hot spots were also identified. Results presented in **Figure 5** shows the correct identification of the deamidation of Asn^{55} (HT04). In this case again the parent peptide could be identified without the modification as well, and the fragmentation allowed confirmation of the deamidation of this specific amino acid. Another deamidation hot spot present on the HC on Asn^{387} (HT41) was also characterized with the method. In this case again, the MS/MS spectra obtained for HT41 allowed the identification of the deamidation on this specific amino acid which is crucial because the digested peptide is composed of several Asn. Other micro-variants, such as methionine (Met) oxidation hot spots, were also identified by the algorithm using the CESI-MS/MS data. Results allowed the identification of the oxidations of Met255 (HT25) and Met431 (HT48) as a confirmation of the trend of a partial

modification/degradation; both peptides could also be observed without any oxidation. The peptide HT25 was detected at a time of 23.8 min for the intact form and 24.0 min for the oxidized HT25. A similar results could be observed for peptide HT48 as the intact peptide could be detected at 25.2 min and respectively at a time of 25.5 min for the oxidized HT25. Concerning the LC, the main deamidation hot spot located on Asn30 (LT03) was correctly identified by the algorithm. As for the other hot spots, the manual confirmation suggested a partial modification/degradation of the protein.

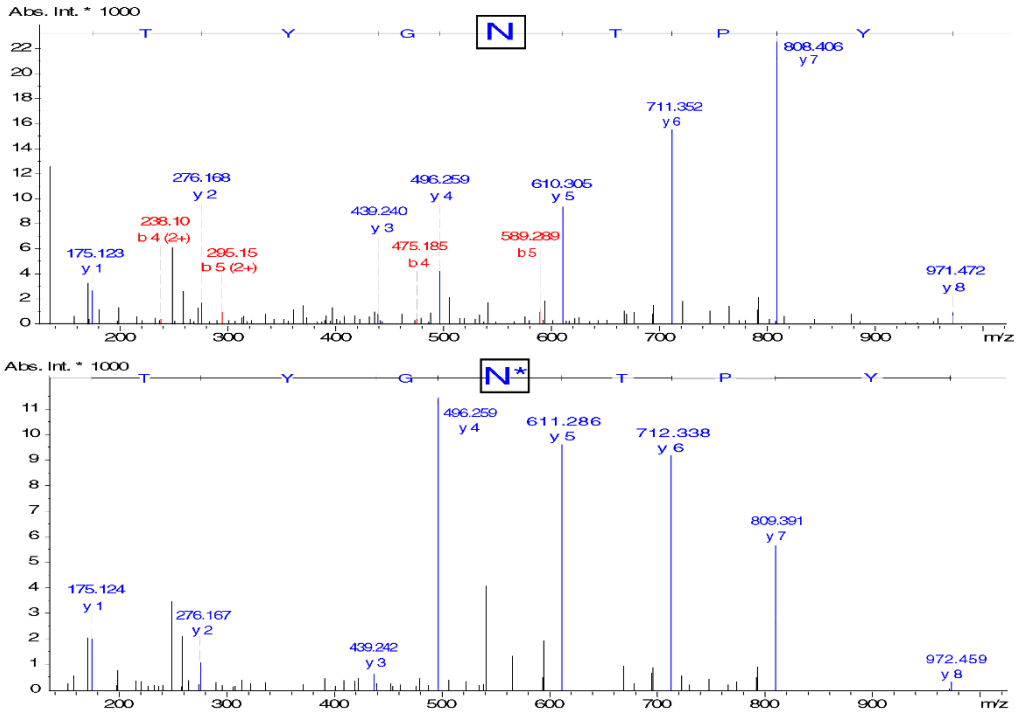

Figure 5. MS/MS spectra of peptide HT04 (IYPTNGYTR) (A) without modification and (B) with a deamidation on Asn^{55} illustrating the partial modification of the mAb and the characterization of both forms by CESI-MS/MS. Experimental conditions: see Materials and Methods.

These results unambiguously demonstrate the possibility of using CESI-MS to perform simultaneous MS/MS peptide mapping and amino acid PTM characterization. The separation of the digested peptide mixture allowed the MS to fragment a large number of ions, which enables the detection of PTMs. The mass spectra quality obtained using the CESI interface allowed the algorithm to successfully identify each studied modification.

Characterization of major N-glycoforms. For most IgGs, glycosylation represents only 2 to 3% of the total mass, but it can significantly influence the protein's structure, immunogenicity and stability [33, 34] Glycosylation is also a class of mAb PTMs: here their study was described separately from amino acid PTMs because they are not considered degradation hot spots. To perform an advanced characterization of trastuzumab, an estimation of the ratio of the main glycoforms, as well as a structural determination is necessary. Trastuzumab is N-glycosylated on the Asn^{300} in the CH2 region [35, 36]. G0F and G1F, the major glycoforms found on human

and recombinant IgG produced in CHO, NS0 and SP2/0 cells, represent more than 80% of the glycoforms [30, 37]. Considering that Mascot search algorithm is only intended to identify the peptide sequences it was necessary to manually study the CESI-MS/MS data to identify the glycosylations and determine their structures. Interestingly, in this case as well, no extra experiments needed to be performed to focus on glycopeptides. The same data sets used for peptide mapping and hot spots characterization were also used here.

In conventional protocols, a mAb is commonly treated with PNGase F before analysis to release the glycans and allow study of them while not attached to the peptide backbone [37]. In this work, the mAb was digested by trypsin only, and therefore we expected to detect glycans still linked to the corresponding peptide backbone of the HC. The extracted ion electropherograms (XIE) of the m/z ratio matching to the peptide TKPREEQYN^{300}STYR bearing respectively the glycosylation G0F, G1F, G0F-GlcNac, G2F and G0 are shown in Figure 6A–E, respectively. All the detected glycopeptides have a charge state of 3.

The glycopeptides were detected between 26.3 and 28.5 min. The electrophoretic resolution calculated between the peaks corresponding to the glycopeptides bearing G0F and G1F gave a value of 0.88, while the resolution between the peaks of glycopeptides G1F and G2F gave a value of 1.08. These results demonstrate the separation of the glycopeptides G0F, G1F and G2F glycan having only one galactose difference in each case. The glycopeptide bearing G0 could also be observed on its corresponding XIE; however, this glycopeptide co-migrated with glycopeptides G0F. Moreover, the sensitivity provided by the CESI prototype allows detection of these glycopeptides with significant intensities and the signal was sufficient to generate good quality

MS/MS spectra of each glycopeptides. More generally, these results allowed us to accurately locate the position of the glycosylation on Asn^{300} because it is the only potential glycosylation site on this peptide. The nature of the glycosylation, however, had to be confirmed using MS/MS data. MS/MS spectra (**Fig. 7**) corresponding to the fragmentation of the different glycopeptides were subsequently extracted from the data. The precursor ion matches the mass of the glycopeptide being glycosylated by G0F with a precision of 8.66 ppm. As expected with low energy collision induced dissociation (CID), product ions observed in MS/MS mainly correspond to the fragmentation of the glycan moiety and the peptide backbone is usually preserved. Study of the fragmentation spectra allowed us to deduce the structure of the glycan G0F. In the same manner, the MS/MS spectra of the glycopeptides bearing G1F, G2F and G0 glycans were extracted from the data; in each case the mass measure precision on the precursor ion was inferior to 9.69 ppm. Similarly to the previous results, every ions

corresponding to the different fragmentation of the glycan could be observed while the peptide backbone remained intact, allowing the structures of the studied glycans to be deduced.

Results presented here demonstrate the possibility of using CE to separate and identify glycopeptides obtained from the trastuzumab enzymatic digestion mainly due to the baseline separation efficiency of the five major glycopeptides provided by the CE. As an additional benefit, the sensitivity provided by the CESI interface and the microTOFQ-II MS allowed us to obtain the complete fragmentation of the glycan and thereby deduce the structure of the five major glycans (G0F, G1F, G2F, G0F-GlcNac and G0). The same data was used for MS/MS peptide mapping and hot-spot characterization without requiring additional experiments.

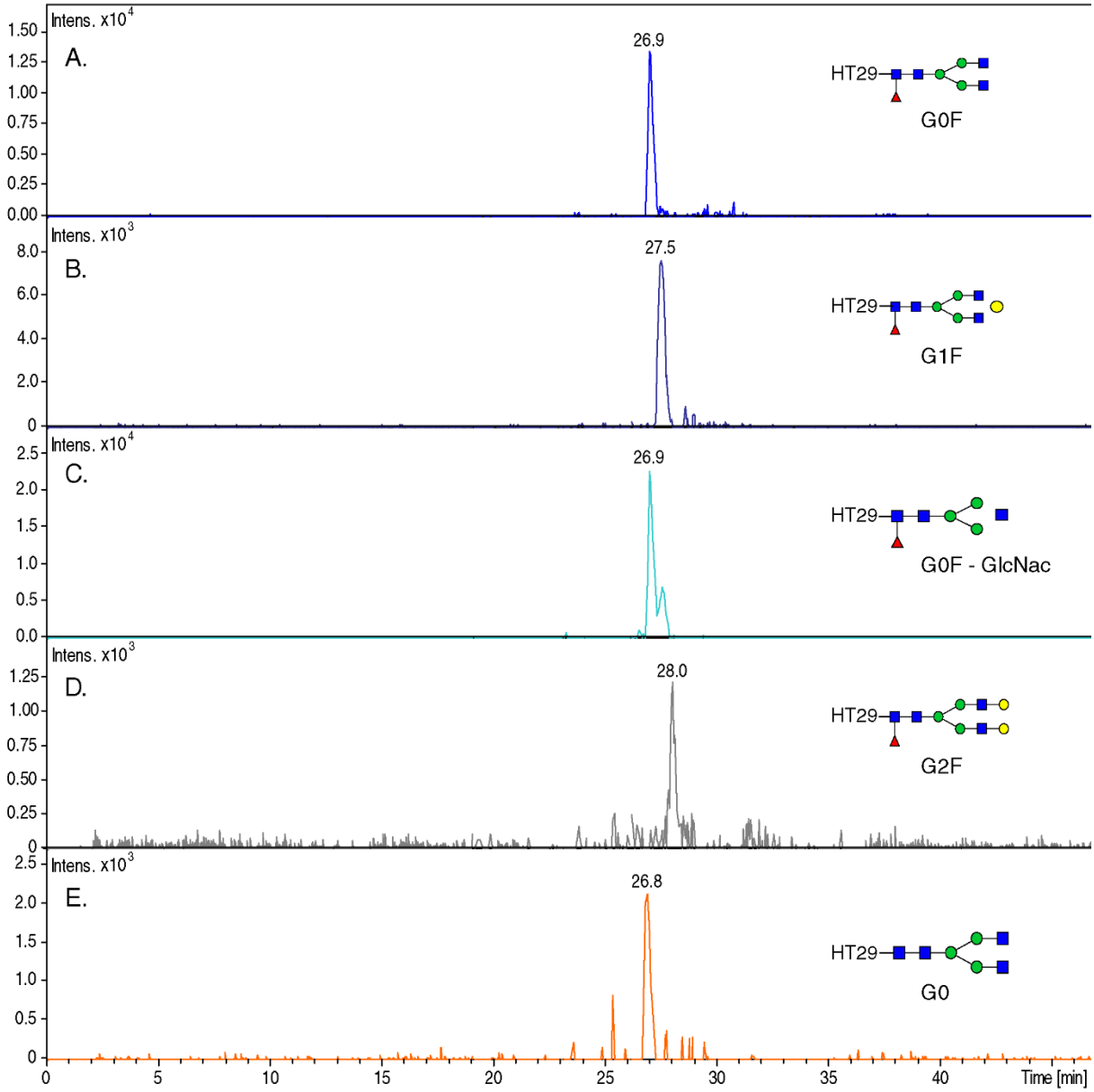

Figure 6. Extracted Ion Electropherogram of m/z ratios (A) 1039.80 (HT29 + G0F), (B) 1093.78 (HT29 + G1F), (C) 972.09 (HT29 + G0), (D) 1147.82 (HT29 + G2F), (E) 991.10 (HT29 + G0). Experimental conditions: see Material and Methods.

Discussion

Peptide mapping is essential for the characterization of therapeutic mAbs, which are complex glycoproteins that contain a plethora of micro-heterogeneities due to the production process. In comparison to PMF, MS/MS peptide identification is not only based on digested peptide m/z ratio measure, but also on the fragmentation of those peptides, increasing the confidence of the identification. Capillary zone electrophoresis (CZE) separation is based on the migration of species present in a background electrolyte (BGE) under an electrical field. Separation is obtained by differences in electrophoretic mobility among analytes depending on size and charge of the molecule at the BGE pH.

Peptide mapping results showed the identification of wide variety of peptides regarding their chemical composition (**Fig. 2**). This result indicates clearly that CE separation driven by peptide migration phenomena is able to separate and transfer successfully to the MS every digested peptide regardless of its nature. In RP-LC, small peptides (from one to two amino acids [23]) and more generally hydrophilic peptides, are often not detected because they are not retained on the column and elute in the void volume [19] The detection and identification of peptide such as LT09 or HT05 demonstrate the ability of CE to conserve and detect small peptides. The identification of the 63 amino acids peptide HT21 also demonstrates one of the advantages of using CE hyphenated to MS for peptide characterization. The presence of such large peptides is not surprising as trastuzumab HC does not have a tryptic cleavage site from amino acid 151 to 213. The elution and detection of such an imposing peptide may be difficult in RPLC-MS/MS because it can irreversibly interact with the stationary phase. Identification of very large peptides having no enzymatic digestion sites or corresponding to peptides or with several missed cleavages is also a very important asset of this method because 100% sequence coverage is obtainable in various cases. Indeed, some digestion sites in mAbs may not be accessible to trypsin due to the quaternary structure of the protein, which may cause missed cleavages during the digestion even when guanidine is present as a chaotropic reagent. In addition to the complete sequence coverage obtained of the HC and the LC, the sensitivity provided by the CESI prototype supported by the high resolution of the mass spectrometer allowed generation of high quality MS/MS spectra. This result is explained by a remarkable ionization efficiency of the peptides. The low BGE flow rate during the separation indeed enhances the ionization process and reduces ion suppression [20, 38]. The spectral quality is reflected by the possibility of retracing the major part of the amino acid sequence in the variable domain. As the CESI-MS system is able to transfer efficiently an important number of ions, MS/MS spectra were of an excellent quality in terms of resolution and intensities (**Fig. 3**). Therefore, a significant number of y/b fragments could be attributed, allowing retracing on the variable domain of the HC 109 of 120 amino acids and 98 of 107 on the LC. Understanding

the variable domain of the mAb is crucial because it is directly involved in the antibody-antigen binding. It is noteworthy that such level of characterization could be attained with only one run of the CESI-MS/MS system.

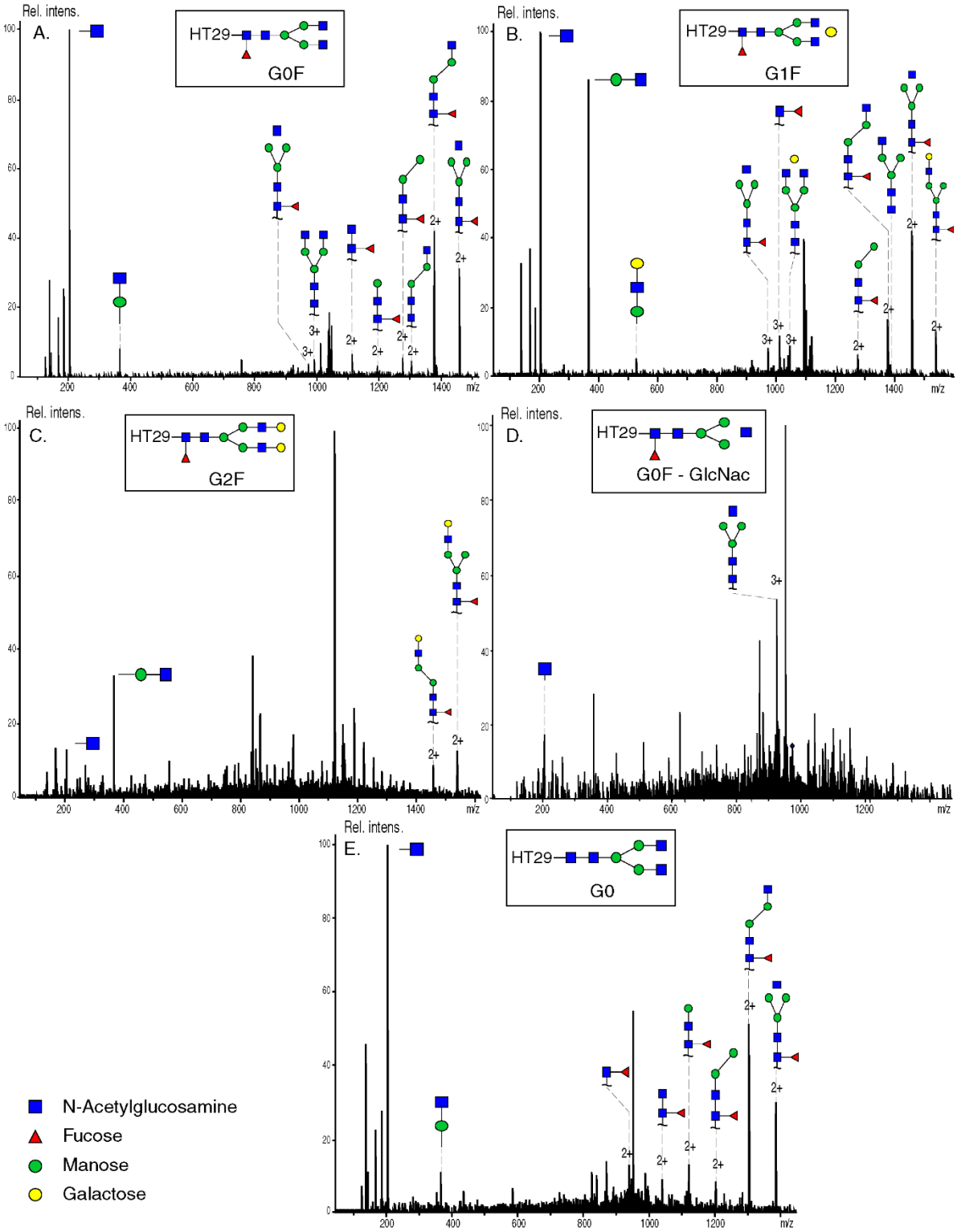

Figure 7. MS/MS fragmentation spectra of (A) HT29 - G0F (precursor m/z 1039.78; charge state 3+), (B) HT29 - G1F (precursor ion 1093.80; charge state 3+), (C) HT29 - G2F (precursor ion 1147.82; charge state 3+), (D) HT29 - G0F-GlcNac (precursor ion 972.10; 3+), (E) HT29 - G0 (precursor ion 991.09; charge state 3+). Experimental conditions described in Materials and Methods section.

Our results demonstrated the successful characterization of every single degradation hot spots on HC and LC and we observed in each case the partial modification of the protein because the modified and the intact form of the peptide could be identified (**Fig. 5**). In the case of the glutamine cyclization of the N-terminal extremity on the HC, the CE separation mechanism allowed the separation of the modified and intact forms of the peptide based on conformational changes induced by glutamine cyclization. Such separation is explained by the fact that even if the mass difference between the parent and the modified peptide is only 17.02 Da, Glu cyclization involves the loss of the amine function on its lateral chain; therefore, there is a difference of one charge between glutamic acid and pyroglutamic acid because the amine is protonated in 10% acetic acid (pH 2.2). The possibility of separate the peptides prior to MS analysis, allows the identification of both forms. Indeed as the number of precursor ions selected for fragmentation per cycle is fairly limited on the microTOF-Q II, and taking into account the complexity of the sample, it would have been difficult for the MS to automatically select the m/z ratio corresponding to both forms of this peptide during their detection window if they had co-migrated. In the case of the deamidation hot spots present on Asn^{55} and Asn^{387} on the HC and Asn^{30} on the LC, the MS was able to fragment both the modified and the intact form for each peptide. Unlike in the glutamic acid cyclization, the XIE corresponding to those peptides (HT04, HT41 and LT03) showed in each case the complete co-migration between the intact peptide and its deamidated homolog. This is explained by the fact that the mass difference between asparagine and aspartic acid is only 0.98 Da. Regarding deamidation, the amide functional group located on asparagine lateral chain, is replaced by an alcohol function. Most importantly, there is thus no significant charge difference between the deamidated peptide and the intact peptide because those functions are protonated at BGE pH. In this case of oxidized Met^{255} (HT25) and Met^{431} (HT48), even if the detection times were not exactly the same for the intact peptide and its oxidized homologous, differences in electrophoretic mobilities could not lead to the baseline separation of those peptides. The sensitivity provided by the system allowed unambiguous characterization of each studied degradation hot spot, in addition to the low represented to PTMs and degradations present in small amounts.

Using the CESI-MS/MS data, the complete structural characterization of the glycosylation G0F, G1F, G2F, G0F-GlcNac and G0 could be achieved (**Fig. 7**). Because trypsin is not able to cleave glycosylation from the peptide, the glycan was observed on the corresponding digested peptide. The XIE corresponding to the glycopeptides bearing different glycans (Fig. 6) demonstrated the separation of several glycopeptides in a reduced time. This result demonstrated the ability of the CE migration mechanism to completely separate glycopeptides having a difference of only one galactose (e.g., between G0F and G1F). Such levels of separation could be explained by the fact that one more galactose leads to a difference

between the two glycopeptides in term of size (+162 Da). This difference affects their relative electrophoretic mobilities leading to their separation. The efficiency of CE in the glycopeptides separation has allowed the separation to be obtained in ~27 min. A difference of a fucose, however, did not lead to the separation of the glycopeptides G0F and G0, suggesting that the size of the glycopeptides is not significantly affected by the fucosylation at the BGE pH.

To evaluate the robustness of this method, the analysis was consecutively repeated six times using the same mAb digest sample. Each analysis resulted in 100% sequence coverage in term of MS/MS peptide mapping. Similarly, the degradation hot spots identified were the same in each case and the five same major N-glycans could also be structurally characterized in the six analyses. Intensity ratios between the different glycopeptides were calculated for each analysis; the results exhibited a standard deviations ranging from 6 to 13%. Based on this results, glycopeptide intensities appear to be related to glycan distribution. It is therefore it is possible to consider profiling the mAb glycan using this very same CESI-MS/MS method, which would enriching the amount of information obtain from only one analysis.

In conclusion, this work presents the development of a rapid and reliable method for performing a multi-level characterization of trastuzumab, which was used as reference therapeutic antibody in a single run. The mAb was digested using a conventional digestion protocol and the digest was analyzed by CESI hyphenated to a high resolution tandem MS. The CESI-MS/MS fragmentation obtained in only one acquisition allowed the complete amino acid sequence peptide map to be recorded. A Mascot search engine was used for peptide identification to provide a faster protocol. Also, the characterization of every PTM hot spot could be obtained within the same single run. Finally, separation of the major glycosylated peptides and their structural characterization could be obtained using the same CESI-MS/MS analysis data set. Such advanced characterization using only a single injection of 100 fmol of digested mAb peptide mixture in a 45 min analysis demonstrates the potential of CESI-MS as a fast and sensitive method in mAb characterization and more generally in bottom-up proteomic approach.

Materials and methods

Materials. All chemicals used were of analytical grade or high purity grade. Acetic acid was provided by VWR. Bicarbonate ammonium, guanidine hydrochloride, iodoacetamide (IAM), dithiothreitol (DTT), trifluoroacetic acid, acetonitrile (ACN) and trypsin enzyme were purchased from Sigma-Aldrich. Water used to prepare buffers and sample solutions was obtained using an ELGA purelab UHQ PS water purification system. Desalting steps were performed using C8 spin tips purchased from Protea Biosciences.

Tryptic digestion procedure. Tryptic digests of antibodies were prepared using sequence-grade trypsin (1:10 w/w). Before digestion, sample was diluted to a concentration of 10 mg/mL. One mg of protein was mixed with 155 µL of denaturing reagent (guanidine HCl 6 M in 50 mM ammonium bicarbonate) and 15 µL of DTT 100 mM were added. The mixture was incubated for 5 min at 95°C, then cooled to ambient temperature. Then, 30 µL of IAM 100 mM was added and the mixture was placed in the dark for 20 min. Ten µL of trypsin was added and the mixture was incubated at room temperature for 3 h; additional 10 µL of trypsin were added to the sample for overnight digestion at room temperature. Buffer exchange and concentration were performed using a C8 spin tips. Four hundred µL of mAb tryptic digest in BGE could be obtained. A volume of ten µL of the obtained digest was used to perform all CESI-MS/MS experiments.

Capillary electrophoresis. The CE experiments were performed with a PA 800 plus capillary electrophoresis (CE) system from Beckman Coulter equipped with a temperature controlled autosampler and a power supply able to deliver up to 30 kV. Hyphenation was realized using a CESI prototype made available by Beckman Coulter. Prototype of bare fused-silica capillaries (total length 100 cm; 30 µm i.d.) with characteristic porous tip on its final 3 cm supplied by Beckman Coulter, a second capillary (total length 80 cm; 50 µm i.d.) filled during experiments with BGE allows electric contact. New capillaries were flushed for 10 min at 75 psi (5.17 bar) with methanol, then 10 min with 0.1 M sodium hydroxide, followed 10 min with 0.1 M hydrochloric acid and water for 20 min also at 75 psi. Finally, the capillary was flushed 10 min at 75 psi with BGEn which was acetic acid 10%. Hydrodynamic injection (69 mbar for 1 min) corresponding to a total volume of 11 nL of sample injected was used. Separations were performed using a voltage of +20 kV.

Mass spectrometry. For antibody characterization, the CESI system was coupled to a microTOF-Q II mass spectrometer (Bruker Daltonics). The microTOF-Q II MS is equipped with a hybrid analyzer composed of a quadrupole followed by a time-of-flight (TOF) analyzer. Positive mode acquisition was used to detect precursor ions (MS) and fragmented product ions (MS/MS). Concerning the ESI source parameters, capillary voltage was set to -1.3 kV. Nebullizer gas was deactivated, the dry gas was set to 1.5 L/min and temperature of the source was set at 160°C. Spectra were collected at a data acquisition frequency of 2 Hz; for fragmentation spectra, collision energy ranged from 0 to 45 V depending on the m/z ratio and charge state of the precursor ion. For each MS scan, 3 precursor ions were selected for fragmentation, and total duty cycle was therefore 2 sec. Mass range was 50–3000 for MS as well as MS/MS scans.

MS/MS data analysis. Data obtained from CESI-MS/MS experiments were processed using Mascot search algorithm developed by Matrix Science. Tryptic cleavage rules were applied for both HC and LC sequences of the mAb. Carbamidomethylation of cysteine (+57.02 Da) was selected as a fixed modification, N-deamidation of aspartic/isoaspartic acid (+0.985 Da) or succinimide intermediate (-17.03 Da) were selected as a variable modifications. Methionine oxidation (+15.99 Da) and N-terminal glutamic acid cyclization (-17.02 Da) were also selected as variable modifications. The mass tolerance for precursor ions was set to ± 25 ppm and to ± 0.5 Da for fragments. A maximum of 3 missed cleavages was tolerated. MS/MS N-glycan identification and structural characterization were done manually.

Acknowledgments

Rabah Gahoual would like to thank the MRT for funding his Ph.D. work. LSMIS would like to thank Beckman Coulter Inc. for lending a CESI prototype, Bruker Daltonics lending the microTOF-Q II and M. Anselme from Beckman Coulter Inc. for his help. The authors would also thanks Dr. E. Wagner-Rousset, Dr. D. Ayoub, MC. Janin-Bussat and O. Colas (Centre d'Immunologie Pierre Fabre, Saint-Julien en Genevois, France) for discussions around sample preparation and antibody LC-MS characterization.

References

[1] **Beck A, Wurch T, Bailly C, Corvaia N**. Strategies and challenges for the next generation of therapeutic antibodies. Nat Rev Immunol 2010; 10:345-52

[2] **Reichert JM.** Marketed therapeutic antibodies compendium. MAbs 2012; 4:413-5

[3] **Ambrogelly A, Liu YH, Li H, Mengisen S, Yao B, Cannon-Carlson S, et al**. Characterization of antibody variants during process development The tale of incomplete processing of N-terminal secretion peptide. MAbs 2012; 4:1-9

[4] **Damen CWN, Chen W, Chakraborty AB, van Oosterhout M, Mazzeo JR, Gebler JC, et al**. Electrospray ionization quadrupole ion-mobility time-of-flight mass spectrometry as a tool to distinguish the lot-to-lot heterogeneity in N-glycosylation profile of the therapeutic monoclonal antibody trastuzumab. J Am Soc Mass Spectrom 2009; 20:2021-33

[5] **Committee for medicinal products for human use (CHMP), European Medicines Agency (EMA)**. Guideline on development, production, characterization and specifications for monoclonal antibodies and related products 2009; EMEA/CHMP/ BWP/157653/2007

[6] **McCamish M, Woollett G**. Worldwide experience with biosimilar development. MAbs 2011; 3:209-17

[7] **Reichert JM, Beck A, Iyer H**. European Medicines Agency workshop on biosimilar monoclonal antibodies: July 2, 2009, London, UK. MAbs 2009; 1:394-416

[8] **Zhang Z, Pan H, Chen X**. Mass spectrometry for structural characterization of therapeutic antibodies. Mass Spectrom Rev 2009; 28:147-76

[9] **Harris RJ, Chin ET, Macchi F, Keck RG, Shyong BJ, Ling VT, et al**. Analytical characterization of monoclonal antibodies: linking structure to function. Current Trends in Monoclonal Antibody Development and Manufacturing, Springer 2010; 193-205

[10] **Wagner-Rousset E, Bednarczyk A, Bussat MC, Colas O, Corvaïa N, Schaeffer C, et al**. The way forward, enhanced characterization of therapeutic antibody glycosylation: comparison of three level mass spectrometry-based strategies. J Chromatogr B Analyt Technol Biomed Life Sci 2008; 872:23-37

[11] **Beck A, Bussat MC, Zorn N, Robillard V, Klinguer-Hamour C, Chenu S, et al**. Characterization by liquid chromatography combined with mass spectrometry of monoclonal anti-IGF-1 receptor antibodies produced in CHO and NS0 cells. J Chromatogr B Analyt Technol Biomed Life Sci 2005; 819:203-18

[12] **Xie H, Chakraborty A, Ahn J, Yu YQ, Dakshinamoorthy DP, Gilar M, et al**. Rapid comparison of a candidate biosimilar to an innovator monoclonal antibody with advanced liquid chromatography and mass spectrometry technologies. MAbs 2010; 2:379-94

[13] **Alvarez M, Tremintin G, Wang J, Eng M, Kao YH, Jeong J, et al**. On-line characterization of monoclonal antibody variants by liquid chromatography-mass spectrometry operating in a two-dimensional format. Anal Biochem 2011; 419:17-25

[14] **Michels DA, Tu AW, McElroy W, Voehringer D, Salas-Solano O**. Charge heterogeneity of monoclonal antibodies by multiplexed imaged capillary isoelectric focusing immunoassay with chemiluminescence detection. Anal Chem 2012; 84:5380-6

[15] **Kaschak T, Boyd D, Yan B**. Characterization of glycation in an IgG1 by capillary electrophoresis sodium dodecyl sulfate and mass spectrometry. Anal Biochem 2011; 417:256-63

[16] **Rustandi RR, Wang Y**. Use of CE-SDS gel for characterization of monoclonal antibody hinge region clipping due to copper and high pH stress. Electrophoresis 2011; 32:3078-84

[17] **Moini M**. Simplifying CE-MS operation. 2. Interfacing low-flow separation techniques to mass spectrometry using a porous tip. Anal Chem 2007; 79:4241-6

[18] **Haselberg R, Ratnayake CK, de Jong GJ, Somsen GW**. Performance of a sheathless porous tip sprayer for capillary electrophoresis-electrospray ionization mass spectrometry of intact proteins. J Chromatogr A 2010; 1217:7605-11

[19] **Faserl K, Sarg B, Kremser L, Lindner H**. Optimization and evaluation of a sheathless capillary electrophoresiselectrospray ionization mass spectrometry platform for peptide analysis: comparison to liquid chromatography-electrospray ionization mass spectrometry. Anal Chem 2011; 83:7297-305

[20] Busnel JM, Schoenmaker B, Ramautar R, Carrasco-Pancorbo A, Ratnayake C, Feitelson JS, et al. High capacity capillary electrophoresis-electrospray ionization mass spectrometry:

coupling a porous sheathless interface with transient-isotachophoresis. Anal Chem 2010; 82:9476-83

[21] **Ramautar R, Busnel JM, Deelder AM, Mayboroda OA**. Enhancing the coverage of the urinary metabolome by sheathless capillary electrophoresis-mass spectrometry. Anal Chem 2012; 84:885-92

[22] **Haselberg R, Harmsen S, Dolman MEM, de Jong GJ, Kok RJ, Somsen GW**. Characterization of druglysozyme conjugates by sheathless capillary electrophoresis- time-of-flight mass spectrometry. Anal Chim Acta 011; 698:77-83

[23] **Whitmore CD, Gennaro LA**. Capillary electrophoresis-mass spectrometry methods for tryptic peptide mapping of therapeutic antibodies. Electrophoresis 2012; 33:1550-6

[24] **Demeule B, Palais C, Machaidze G, Gurny R, Arvinte T**. New methods allowing the detection of protein aggregates: a case study on trastuzumab. MAbs 2009; 1:142-50

[25] **Hudis CA**. Trastuzumab--mechanism of action and use in clinical practice. N Engl J Med 2007; 357:39-51

[26] **Carter P, Presta L, Gorman CM, Ridgway JBB, Henner D, Wong WL, et al**. Humanization of an anti-p185HER2 antibody for human cancer therapy. Proc Natl Acad Sci U S A 1992; 89:4285-9

[27] **Harris RJ, Kabakoff B, Macchi FD, Shen FJ, Kwong M, Andya JD, et al**. Identification of multiple sources of charge heterogeneity in a recombinant antibody. J Chromatogr B Biomed Sci Appl 2001; 752:233-45

[28] **Vlasak J, Bussat MC, Wang S, Wagner-Rousset E, Schaefer M, Klinguer-Hamour C, et al**. Identification and characterization of asparagine deamidation in the light chain CDR1 of a humanized IgG1 antibody. Anal Biochem 2009; 392:145-54

[29] **Gennaro LA, Salas-Solano O, Ma S**. Capillary electrophoresis-mass spectrometry as a characterization tool for therapeutic proteins. Anal Biochem 2006; 355:249-58

[30] **Beck A, Sanglier-Cianférani S, Van Dorsselaer A**. Biosimilar, biobetter, and next generation antibody characterization by mass spectrometry. Anal Chem 2012; 84:4637-46

[31] **Houde D, Peng Y, Berkowitz SA, Engen JR**. Posttranslational modifications differentially affect IgG1 conformation and receptor binding. Mol Cell Proteomics 2010; 9:1716-28

[32] **Liu H, Gaza-Bulseco G, Faldu D, Chumsae C, Sun J**. Heterogeneity of monoclonal antibodies. J Pharm Sci 2008; 97:2426-47

[33] **Jefferis R**. Glycosylation of recombinant antibody therapeutics. Biotechnol Prog 2005; 21:11-6

[34] **Beck A, Reichert JM**. Marketing approval of mogamulizumab: a triumph for glyco-engineering. MAbs 2012; 4:1-7

[35] **Beck A, Wagner-Rousset E, Bussat MC, Lokteff M, Klinguer-Hamour C, Haeuw JF, et al**. Trends in glycosylation, glycoanalysis and glycoengineering of therapeutic antibodies and Fc-fusion proteins. Curr Pharm Biotechnol 2008; 9:482-501

[36] **Beck A, Cochet O, Wurch T**. GlycoFi's technology to control the glycosylation of recombinant therapeutic proteins. Expert Opin Drug Discov 2010; 5:95-111

[37] **Damen CWN, Chen W, Chakraborty AB, van Oosterhout M, Mazzeo JR, Gebler JC, et al**. Electrospray ionization quadrupole ion-mobility time-of-flight mass spectrometry as a tool to distinguish the lot-to-lot heterogeneity in N-glycosylation profile of the therapeutic monoclonal antibody trastuzumab. J Am Soc Mass Spectrom 2009; 20:2021-33

[38] **Wilm MS, Mann M**. Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last? Int J Mass Spectrom Ion Process 1994; 136-167

3. Conclusion

Les travaux présentés dans cette seconde partie détaillent le développement d'une méthode d'analyse par CE-ESI-MS/MS, par l'intermédiaire du système CESI-MS, pour la caractérisation d'anticorps monoclonaux. La méthodologie adoptée dans le cadre de cette caractérisation est basée sur une stratégie de type protéomique *bottom-up*. Les échantillons ont au préalable subi une digestion protéolytique à la trypsine en utilisant un protocole de digestion en solution. Aucune étape de dessalage n'a été effectuée afin d'éviter la perte des peptides hydrophiles, non retenus sur la phase stationnaire lors de l'extraction sur phase solide (SPE). Le digestat a ensuite été séparé et analysé par CESI-MS/MS.

Les résultats obtenus pour la caractérisation de la séquence d'acides aminés ont permis d'en déduire 100% de recouvrement de séquence en une seule injection correspondant à 100 fmol de peptides digérés. Un tel résultat est sans précédent. Ainsi la littérature ne décrit pas un recouvrement complet d'une protéine par l'analyse MS/MS en une unique injection. La mise en œuvre du couplage CESI-MS pour ce type de caractérisation joue un rôle central dans l'obtention de ce résultat par plusieurs aspects. L'identification de peptides hydrophiles et les peptides de faibles masses moléculaires (1-3 acides aminés) peut s'avérer problématique par des approches classiques utilisant des méthodes de RP-LC-MS/MS. En effet, ces peptides ne sont pas ou peu retenus par la phase stationnaire et sont généralement élués dans le volume mort. En CE, tous les peptides migrent sous l'effet du champ électrique quel que soit leur structure ou leur taille et sont donc transférés successivement au spectromètre de masse. Il en résulte la possibilité en CE-MS d'obtenir la séparation et l'analyse MS/MS d'une grande variété de peptides : l'analyse développée a montré la possibilité d'obtenir la séparation et l'identification d'un peptide de faible masse moléculaire composé de seulement 4 acides aminés et également d'un peptide composé de 63 acides aminés dont la masse excède les 6 kDa. L'excellent rendement d'ionisation délivré par le système CESI-MS, décrit dans la partie précédente, a également démontré un intérêt singulier dans ce type de caractérisation. L'analyse de seulement 100 fmol de peptides digérés a permis d'obtenir le recouvrement

complet de la séquence du mAb illustrant la sensibilité que permet d'atteindre le système. De plus le rendement d'ionisation a un impact significatif sur la qualité des spectres générés notamment en MS/MS. Les spectres ont montré l'identification d'un nombre important de fragments, ce qui se traduit en la capacité de caractérisation de 109/120 et 97/108 des acides aminés composant le domaine variable de trastuzumab respectivement sur la HC et la LC. Un tel niveau de caractérisation est particulièrement intéressant et illustre l'intérêt de ce type de couplage.

Parallèlement, la caractérisation des PTMs *hot-spots* a été réalisée en utilisant les mêmes données que la caractérisation de la séquence. Les données CESI-MS/MS ont ainsi permis la caractérisation de 6 hot spots présents sur le mAb. Les résultats ont montré la capacité, par le biais de la méthodologie développée, à obtenir conjointement le spectre de fragmentation du peptide modifié et de son homologue intacte, ce qui offre la possibilité entre autres de définir de manière précise l'acide aminé concerné par la modification.

Enfin les différentes glycosylations portées par trastuzumab ont aussi pu être caractérisées en utilisant le même set de données que précédemment. Les données MS/MS ont ainsi permis l'identification des 5 glycosylations majeures portées par trastuzumab, qui selon la littérature représente plus de 80% des glycosylations portées par le mAb. L'étude des spectres MS/MS des différents glycopeptides a montré la fragmentation du glycanne porté par ce peptide, ce qui a permis d'en déduire les structures des glycosylations. Une fois de plus l'utilisation de la CE en tant que méthode séparative apparait comme un atout. Les différents glycopeptides ont pu être partiellement séparés en fonction du glycoforme, notamment dans le cas d'une différence d'un galactose. Cette sélectivité est particulièrement intéressante afin de limiter les effets de suppression d'ions et participe de manière significative à la sensibilité de la caractérisation sur cet aspect de la protéine.

La méthode développée a permis de caractériser, en une seule injection, l'anticorps monoclonal trastuzumab simultanément sur plusieurs niveaux définissant sa structure. Ces résultats montrent le potentie en termes de sensibilité et de qualité spectrale offerts par ce type de couplage CE-ESI-MS. De plus l'utilisation de la CE pour cette caractérisation a permis de mettre en lumière l'intérêt d'utiliser la sélectivité de cette méthode séparative.

Il convient de démontrer que la méthodologie développée dans cette partie permet d'obtenir une caractérisation aussi avancée que dans le cas de trastuzumab pour d'autres protéines et d'étudier en détail l'apport que la séparation par CE peut apporter à la caractérisation structurale de protéines par MS.

II - Caractérisation complète de la structure primaire et des micro-hétérogénéités d'anticorps en une seule injection par isotachophorèse et CESI-MS/MS

1. Introduction

L'électrophorèse capillaire (CE) est une technique permettant la séparation d'analytes sous l'effet d'un champ électrique. La CE a pour avantage majeur de permettre la réalisation de séparation en quelques minutes tout en maintenant une grande efficacité de séparation. La littérature fait ainsi état de séparations dont l'efficacité est supérieure à 500 000 plateaux théoriques. Une telle efficacité de séparation s'explique notamment par l'absence de phase stationnaire ce qui a pour effet de diminuer de manière importante la diffusion longitudinale et l'opposition au transfert de masse, à l'origine de l'élargissement des pics. Malgré d'importants développements instrumentaux, les systèmes permettant le couplage entre la CE et la spectrométrie de masse à source électrospray (ESI-MS) sont employés de manière marginale par rapport au couplage chromatographique. Cette tendance est directement liée à la difficulté à maintenir le champ électrique, nécessaire à la séparation, à l'intérieur du capillaire de CE tout en positionnant l'extrémité de celui-ci à l'intérieur de la source ESI. De plus, la plupart des systèmes permettant la mise en œuvre du couplage CE-ESI-MS utilisent généralement un système pour maintenir le contact électrique induisant une diminution de la sensibilité. A titre d'exemple l'interface dite *sheath liquid* utilise un liquide additionnel pour maintenir la continuité du champ électrique durant la séparation, ce liquide induit un effet de dilution de l'échantillon et tend à diminuer la sensibilité du signal. Par opposition, le couplage entre la chromatographie liquide et la spectrométrie de masse (LC-ESI-MS), plus simple d'un point de vue instrumental, est souvent préféré du fait de sa facilité de mise en œuvre et de sa robustesse.

Depuis son développement, l'ESI a largement démontré son intérêt pour l'ionisation des protéines et des peptides. Ainsi une séparation électrophorétique en complément de l'analyse MS devrait d'un point de vue théorique être privilégiée pour les analytes biologiques par plusieurs aspects. En particulier la CE est une méthode séparative miniaturisée, ce qui est favorable à la formation d'un nanoESI et garantie une ionisation optimale. De plus, la CE est particulièrement adaptée à la séparation de ce type de molécules, ainsi l'électrophorèse est largement utilisée en biologie pour la séparation et la purification d'analytes variées (ADN, ARN, protéines, peptides). La CE mise en œuvre en couplage avec la MS, peut permettre d'outrepasser certaines des limitations relatives aux couplages utilisant la chromatographie liquide, notamment du fait que les principes physico-chimiques misent en jeu pour induire la

séparation sont différents. Le système CESI-MS permet la mise en œuvre du couplage CE-ESI-MS sans utiliser de liquide additionnel ce qui permet d'obtenir une sensibilité maximale d'autant que les débits générer par le flux électroosmotique ne sont que de quelques dizaines de nL/min. Les travaux présentés dans cette partie ont pour objectif de démontrer l'intérêt d'utiliser une méthode de séparation électrophorétique en amont de l'analyse MS, par l'intermédiaire du système CESI-MS, pour la caractérisation de la structure primaire de protéines.

Ces travaux décrivent la mise en œuvre et le développement d'une méthode analytique permettant la caractérisation fine de la structure primaire de plusieurs anticorps monoclonaux (mAbs) dont notamment trastuzumab (Herceptin®, Roche) et cetuximab (Erbitux®, Merck KgaA). La caractérisation d'échantillons de mAbs apparait d'un fort intérêt car ces glycoprotéines regroupent différentes problématiques communément observées sur une large gamme de protéines : caractérisation de la séquence d'acides aminés, glycosylation, modifications post-traductionnelles (PTMs). Ils représentent donc des molécules de choix pour démontrer le degré de caractérisation pouvant être atteint par la méthodologie développée.

Les échantillons ont tout d'abord subis une digestion trypsique. Par rapport aux résultats présentés dans la partie précédente, le protocole de digestion a été modifié afin d'améliorer la compatibilité entre l'échantillon et la méthode d'analyse par CESI-MS/MS, ce qui a pour effet d'améliorer significativement le niveau de caractérisation. Le digestat de peptides a ensuite été séparée et caractérisée par isotachophorèse transitoire (t-ITP) et CESI-MS/MS. L'isotachophorèse transitoire est une méthode de préconcentration utilisée en électrophorèse capillaire. La t-ITP autorise un volume d'injection correspondant à 25% de volume total du capillaire de CE tout en maintenant une efficacité de séparation optimale. Les données de CESI-MS/MS ont dans un premier temps été utilisées pour caractériser la séquence d'acides aminés. Les résultats obtenus ont montré la capacité de la méthodologie développée à obtenir systématiquement 100% de recouvrement de séquence en une seule injection pour chacun des mAbs étudiés. Pour chacun des mAbs étudiés, la caractérisation de la séquence d'acides aminés a pu être réalisée uniquement par l'identification de peptides sans *miscleavages* et ne portant aucune PTMs. Les résultats obtenus illustrent l'important rendement d'ionisation du système CESI-MS ce qui impacte directement la qualité spectrale : l'identification des peptides a montré la capacité en utilisant cette méthode d'en déduire systématiquement plus de 70% des fragments y/b et dans le cas de trastuzumab cette valeur excède les 90%.

Les données de t-ITP CESI-MS/MS ont également été utilisées pour caractériser les glycosylations portées par ces différents mAbs. Dans le cas de trastuzumab, il a été possible de caractériser au total 15 glycosylations différentes dont les abondances se sont avérées

largement hétérogènes. Les performances de la méthode développée, au regard de la caractérisation des glycosylations, sont directement reliées avec la capacité de la CE à séparer les différents glycopeptides en fonction du glycanne qu'ils portent.

Les PTMs *hot-spots* ont également pu être caractérisés pour chacun des mAbs. Dans chacun des cas, l'ensemble des PTMs *hot-spots* portés par la protéine ont pu être caractérisés. L'utilisation de la CE en tant que méthode séparative en amont de l'analyse MS a démontré sa pertinence en particulier dans le cadre de la caractérisation des PTMs. En effet les résultats obtenus ont montré la faculté de la méthodologie développée à séparer complètement un peptide modifié de son homologue resté intact. Ce type de sélectivité, délivrée par l'utilisation de la séparation électrophorétique, a pu être observée même dans le contexte de modification impactant relativement peu le peptide tel que la déamidation d'une asparagine (+0.98 Da) ou encore l'isomérisation de l'acide aspartique (+ 0 Da). La séparation de cette dernière modification est particulièrement intéressante puisqu'elle ne peut être caractérisée en MS par l'intermédiaire d'instruments utilisant un analyseur TOF. Des expériences complémentaires ont d'ailleurs été réalisées afin de démontrer sans aucune ambiguïté la séparation de peptides portant des isomères différents d'acide aspartique.

Les résultats présentés dans cette partie font l'objet d'une publication dans le journal scientifique *Analytical Chemistry (*ACS publications).

2. Full antibody primary structure and micro-variant characterization in a single injection using transient isotachophoresis and sheathless capillary electrophoresis – tandem mass spectrometry

Rabah Gahoual, Jean-Marc Busnel, Alain Beck, Yannis-Nicolas François, Emmanuelle Leize-Wagner, ***Analytical Chemistry*** 2014, 86(18), 9074-9081.

Abstract

Here we report the complete characterization of the primary structure of a multimeric glycoprotein in a single analysis by capillary electrophoresis (CE) coupled to mass spectrometry (MS). CE was coupled to electrospray ionization tandem MS by the mean of a sheathless interface. Characterization was based on an adapted bottom-up proteomic strategy. Using trypsin as the sole proteolytic enzyme and data from a single injection per considered protein, 100% of the amino acid sequences of four different monoclonal antibodies could be achieved. Furthermore illustrating the effectiveness and overall capabilities of the technique, such results were possible through identification of peptides without tryptic miscleavages nor posttranslational modifications illustrating the potency of the technique. In addition to full sequence coverages, PTMs were simultaneously identified demonstrating further the capacity of this strategy to structurally characterize glycosylations as well as faint modifications such as asparagine deamidation or aspartic acid isomerization. Together with the exquisite detection sensitivity observed, the contribution of both CE separation mechanism and selectivity were essential to the fineness of the characterization in regard to what is achieved with conventional MS strategies. The quality of the results indicates that recent improvements in interfacing CE-MS coupling leading to a considerably improved sensitivity, allows characterizing the primary structure of proteins in a robust and faster manner. Taken together, these results open new research avenues for characterization of proteins through MS.

Introduction

Capillary Electrophoresis (CE) has been commercially introduced as a separation technique during the early 1980's [1-3], though electrokinetically driven separation strategies have been applied in laboratories since the beginning of the 20th century [4,5]. In CE, analytes are separated under an electrical field; this technique has some major advantages including the possibility to obtain separations within minutes while maintaining exceptional separation efficiency. This is partially explained by the absence of a stationary phase reducing tremendously longitudinal dispersion responsible for peak broadening. Despite the considerable effort made regarding instrumental development, platforms combining CE and electrospray ionization (ESI) mass spectrometry (MS) is still marginally used as compared to chromatography based methods. This is mainly related to the difficulty to maintain the CE electrical field while positioning the capillary outlet inside the ESI source, meaning in the open air. Another aspect is related to the fact that CE-ESI-MS platforms rarely provide optimal sensitivity as common interfaces rely on strategies that by nature induces losses of sensitivity [6,7]. On another hand, coupling of high-performance liquid chromatography (HPLC) with MS is more straightforward and tend to be preferentially used for separation ahead of MS due to its ease of coupling and excellent robustness. However as ESI-MS demonstrated its suitability for the study of biological samples such as protein and peptides [8,9], electrophoresis should be theoretically the preferred separating technique for biological samples especially because it is a miniaturized technique which should favors the formation of a nanoESI, thus enhancing the ionization process. Biologists are routinely using electrophoresis to reduce the complexity of samples or to isolate proteins [10] likewise DNA [11,12] or RNA [13,14]. Recently a novel sheathless interface has been introduced for CE-ESI-MS hyphenation [15]. It is originally based on a sheathless design by Moini *et a*l [16] and is here referred as CESI-MS. It allows the hyphenization of CE to ESI-MS without giving away the sensitivity because it does not require any sheath liquid to maintain the electrical contact responsible of analyte dilution. Separation performances and characteristics of CE are in terms of efficiency and selectivity well applicable to the range of analytes typically well separated by reverse phase liquid chromatography (RP-LC). Additionally, CE could also alleviate some of the drawbacks usually encountered when using RP-LC such as separation and elution of very small and hydrophilic peptides that may elute with the dead volume in RP-LC or large ones that could be adsorbed irreversibly on the stationary phase. Here we are reporting the capacity of a CESI-MS methodology to enable the complete amino acid (AA) sequence characterization for a protein in a single injection.

MAbs are tetrameric glycoproteins having molecular mass of approximately 150kDa. They are composed of two heavy chains (HC) and two light chains (LC) linked to each other by several

disulfide bonds. HC bears at least one N-glycosylation site17. The first mAbs studied here is trastuzumab which is approved for the treatment of HER2 positive breast cancer [18] and the second antibody is cetuximab directed against the epidermal growth factor receptor (EGFR) and used to treat colorectal, head and neck cancer [19,20]. From an analytical standpoint, these proteins have an interesting trait due to their structural complexity. They present a large number of microheterogeneities commonly found in proteins such as PTMs including glycosylations and small chemical modifications [21]. Four monoclonal antibodies (mAbs) were studied. Alongside to the characterization of the AA se-quence, other aspects of the primary structure of the studied proteins could be characterized with an unprecedented reliabil-ity. Separation mechanisms provided by CE demonstrated their interest for protein characterization by MS as it has been possible to separate peptides having only minor differences as faint as one AA conformational change.

Experimental section

Materials. Chemicals used were of analytical grade or high purity grade and purchased from Sigma-Aldrich (Saint Louis, MO, USA). Water used to prepare buffers and sample solutions was obtained using an ELGA purelab UHQ PS water purification system (Bucks, UK). Trastuzumab and cetuximab samples are the EMA/FDA approved formulation purchased respectively from Genentech (San Francisco, CA, USA) and Merck (Whitehouse Station, NJ, USA). RapiGest SF surfactant was purchased from Waters (Milford, MA, USA).

Sample preparation. For each mab sample, a volume corresponding to 100 µg of protein was sampled using final formulation for the approved mAbs (trastuzumab and cetuximab) and samples directly coming from the bioreactor for the mAbs in development samples. Samples were diluted using milliQ water to a final concentration of 6.7 µg/µL. Samples were then diluted using 0.1% RapiGest surfactant to a final concentration of 3.35 µg/µL and heated to 40°C during 10 min. Dithiothreitol (DTT) was added to the sample in order to obtain a final concentration of 25 mM. Samples were then heated to 95°C during 5 min. After being cooled down to room temperature (RT), iodoacetamide (IDA) was added to the sample to a final concentration of 10 mM. Afterward samples were place in the dark for 20 min to allow alkylation of cysteine (Cys). A volume of 1 µL of trypsin (0.5µg/µL) was added to the sample which was left at room temperature for 3h, another volume of 1µL was added afterward. Digestion was per-formed overnight at 37°C. After digestion completion, formic acid (FA) was added to the samples at a final concentration of 1% (v/v) in order to cleave the surfactant and samples were left at RT for 2h. Samples were finally diluted to a final concentration in protein of 2.2 µM using ammonium acetate 50 mM (pH 4.0).

Capillary electrophoresis. The CE experiments were performed with a PA 800 *plus* capillary electrophoresis system from Beckman Coulter equipped with a temperature controlled autosampler and a power supply able to deliver up to 30 kV. Hy-phenation was realized using a CESI prototype made available by Sciex separation (Brea, CA, USA). Prototype of bare fused-silica capillaries (total length 100 cm; 30 µm i.d.) with characteristic porous tip on its final end on 3 cm, a second capillary (total length 80 cm; 50 µm i.d.) filled during experiments with BGE allows electric contact. New capillaries were flushed for 10 min at 75 psi (5.17 bar) with methanol, then 10 min with 0.1 M sodium hydroxide, followed 10 min with 0.1 M hydrochloric acid and water for 20 min also at 75 psi. Finally, the capillary was flushed 10 min at 75 psi with BGE which was acetic acid 10%. Hydrodynamic injection (410 mbar for 1 min) corresponding to a total volume of 90 nL of sample injected was used. Separations were performed using a voltage of +20 kV.

Mass spectrometry. For antibody characterization, the CESI system was hyphenized to a 5600 TripleTOF mass spectrometer (ABSciex, Darmstadt, Germany). The 5600 MS is equipped with a hybrid analyzer composed of quadrupoles followed by a time-of-flight (TOF) analyzers. ESI source parameters was set as follow: ESI voltage -1.75kV while Gas supplies (GS1 and GS2) were deactivated, source heating temperature 150°C and curtain gas value 5. Experiments were performed in Top15 information de-pendent acquisition (IDA), accumulation time was 250 msec for MS scans and 100 msec for MS/MS scans leading to a total duty cycle of 1.75 sec. Mass/charge (m/z) range was 100-2000 in MS and 50-2000 in MS/MS. Using those parameters, the mean resolution provided by the instrument is 40000 in MS (m/z 485.251) and 25000 in MS/MS (m/z 345.235).

MS/MS data analysis. Data obtained from the CESI-MS/MS experiments were analyzed using Peakview software (ABSciex, San Francisco, CA). Purely tryptic peptides (without miscleavages nor PTMs except cys carbamidomethylation) were determined theoretically from considered mAbs amino acid sequences available through literature. Additional peptides were identified using Mascot search engine provided by Matrix science; tryptic cleavage rules were applied. Carbamidomethylation of cysteine (+57.02 Da), N-deamidation of aspartic/isoaspartic acid (+0.985 Da) or succinimide intermediate (-17.03 Da) were selected as a variable modifications. Methionine oxidation (+15.99 Da) and N-terminal glutamic acid cyclization (-17.02 Da) were also selected as variable modifications. The mass tolerance allowed for search algorithm identification, for precursor ions was set to ± 5 ppm and ±0.05 Da for fragment ions.

Results and discussion

MAbs were characterized in a bottom-up proteomic adapted strategy, samples were digested by trypsin using an in-solution digestion protocol. We reported previously the development of a method using sheathless CE-ESI-MS/MS for monoclonal antibodies characterization [22]. In the current work, digestion protocol was modified in order to enhance proteolytic digestion efficiency while increasing the compatibility of samples content to capillary zone electrophoresis (CZE) and control matrix effect, likewise the use of transient isotachophoresis (t-ITP). The sample preparation was conducted without any desalting treatment in order to prevent any potential loss of peptides due to either poor or irreversible retention during reverse phase solid phase extraction (SPE). After digestion, the sample was diluted to a final concentration of 2.2 µM in ammonium acetate (50 mM, pH 4.0). Ammonium acetate was chosen as a sample matrix for its compatibility with both ESI-MS and t-ITP. T-ITP is an electrokinetic-based preconcentration process, commonly used in CZE, which enables larger sample injections without any detrimental effect on separation efficiency. Contrarily, the integration of t-ITP often enables an improvement of separation efficiency as compared to conventional CZE [23]. In t-ITP, the sample buffer used has an electrophoretic mobility superior to the background electrolyte (BGE); under the electrical field applied during the separation the sample content is staked in a reduced capillary volume com-pared to the actual injected volume. The use of t-ITP leads to the capacity to inject significantly larger volumes without losing separation efficiency: maximum 25% of the total capillary volume while only 1-2% in conventional CZE [15]. The separation is per-formed under an electrical field of 210 V/cm in a background electrolyte (BGE) composed of 10% acetic acid. This BGE has two advantages, it presents a rather low conductivity as it is not a strong acid and it is fully compatible with ESI ionization process. MAbs digests were analyzed through CESI-MS/MS, the injection volume corresponded to a quantity of 200 fmol of digested peptides. Peptides identification was performed through a peptide fragment fingerprinting (PFF) strategy where peptides are identified based on their complete molecular mass and fragmentation pattern with a mass accuracy systematically better than 5 ppm.

As it is emphasized in Figure 1, the CESI-MS/MS analysis of the mAb tryptic digest allowed in a single injection to obtain 100 % sequence characterization for both HC and LC. Furthermore the full sequence characterization could be performed exclusively through identification of peptides without posttranslational modifications (PTMs) nor miscleavages. To our knowledge, this is the first time that a protein tryptic digest could be entirely characterized in a single injection in such manner. As expected, addi-tional peptides exhibiting miscleavages or various PTMs could also be identified during the experiment though; they can be used to confirm parts of the AA sequence if necessary. Such relevant result is explained by the use of t-ITP CESI-

MS/MS for the method along to the sample preparation which was adapted to allow highly efficient digestion. Additionally sample preparation permits to obtain a complete compatibility to the CE separation conditions which allows to take profit of the electrokinetic separation and ESI ionization to the fullest. The same experiment was performed by nanoLC-MS/MS on trastuzumab digest using the same instrumental settings. Results from a single injection could not permit to obtain the complete sequence coverage (Supporting Information Figure S-1). A simple solution to complete the sequence coverage would be to use a different proteolytic enzyme such as chymotrypsin and concatenate all peptides identified in each digest.

Figure 1. Sequence coverage obtained for trastuzumab by CESI-MS/MS methodology. Experimental conditions: 90 nL injected (200 fmol). CESI-MS/MS spectra recorded on 5600 TripleTOF (AB Sciex, San Francisco, US). Constant domain (blue), variable domain (orange) and complementarity determining region (red) are represented for heavy chain and light chain.

This capacity to characterize without restrictions every peptide composing the digest opens new possibilities for protein primary structure characterization. In particular, the possibility to go beyond the DNA sequence usually used for peptide identification or else identify mutations and/or transcription mismatches. Such capability can be explained by the low pH (2.2) of the BGE, which induces every peptide to be positively charged in solution. With those conditions, all peptides are migrating toward the MS inlet regardless of their chemical nature under the electrophoresis as it can be demonstrated by the separation and identification in the same experimental conditions of peptides having from 2 to 63 AAs and a large range of isoelectric

points (pI). Results point out additional advantages provided by the CESI-MS technique for protein characterization; N-terminal as well as C-terminal parts of the protein could be completely and robustly characterized. Those specific parts of proteins are known to be difficult to characterize in RPLC-MS/MS experiments as they tend to be very poorly retained on reverse stationary phases. Moreover in this case, the N-terminal parts known as the variable domain of the antibody are involved in the antibody-antigen recognition and requires a high level of characterization. Tandem MS (MS/MS), through gas-phase fragmentation of tryptic peptides, allows until a certain extent to precisely identify the AAs order of a peptide depending on the spectra quality [24]. Results demonstrated the capacity of the CESI-MS technique to obtain almost all y/b ions of peptides from mAb variable domains and even for trastuzumab its totality. Over the whole protein, systematically more than 70% of the y/b fragment ions could be obtained during the experiment and more than 90% in the case of trastuzumab (Table 1); depending largely on the size of the tryptic peptides generated. Y and B ions are generated by peptide fragmentation in collision induced dissociation (CID) in MS/MS [25]. Peptide identifications are partially based on those fragment ions. The possibility to detect near the totality of the fragment ions allows on one hand to improve the confidence in the identification. On the other hand, fragment ions give precious information about the precise succession order of AA along the sequence and determine exactly AA experiencing chemical modifications. Such capability could be confirmed for both HC and LC of thefour different mAbs studied (trastuzumab, cetuximab, mab 1 and mab 2). Additionally, three different digestions were characterized for each samples and considered as technical replicates allowed to obtain similar results, proving the robustness of the designed methodology. MS/MS results describe the superior spectra quality obtained while coupling CE to MS by the means of the CESI interface. Spectra quality is a direct consequence of the ionization efficiency which directly impacts the achievable sensitivity and signal/noise ratio. In the case of the CESI interface, the ability to generate a very stable spray at quite low flow rates enablesto robustly operate in the nanoESI regime. Intrinsic characteristics of the CESI interface have akey role in the ionization yield of the interface [26]. Briefly, in nanoESI smaller droplets are initially formed, favoring Rayleigh division but also desolvation process and finally resulting in readily improved ionization and signal/noise ratio compared to standard ESI [27].

MAbs are glycosylated proteins, those glycans are naturally incorporated in the protein during secretion into the extracellular environment [28]. Glycosylation has been implicated in mAbs safety, pharmacokinetic/pharmacodynamics (PK/PD) and is one of the main sources of heterogeneity among this type of protein. Therefore, extensive characterization in terms of structure and relative abundance are mandatory. Concomitantly to primary sequence

characterization, using the same CESI-MS/MS data, in-depth characterization of glycosylation was possible.

	Trastuzumab	Cetuximab	mab in-dev #1	mab in-dev #2
sequence coverage	100%	100%	100%	100%
% MS2 y/b ions	> 90%	> 70 %	> 90%	> 70%
identified glycosylations	15	15	10	16
other PTMs hotspots				
glutamic acid cyclization	1 / 1	1 / 1	1 / 1	1 / 1
methionine oxidation	2 / 2	0 / 0	2 / 2	0 / 0
asparagine deamidation	4 / 4	4 / 4	2 / 2	4 / 4
aspartic acid isomerization	6 / 6	2 / 2	3 / 3	2 / 2

Table 1. Summarized results obtained for a single analysis of each antibody studied using the CESI-MS/MS analysis showing the robustness of the methodology developed and the extended of the primary structure characterization.

For example, in the case of trastuzumab, 15 different glycoforms, were identified demonstrating the outstanding sensitivity of the CESI-MS method (Figure 2). Glycopeptides were identified based on accurate mass measurement in MS1 provided by the high resolution MS (sub 2 ppm) and additionally fragmentation spectra. Indeed, MS/MS spectra exhibited the fragmentation of glycans present on the glycopeptide giving structural information of the glycans along with reinforcing the confidence of the identification. Furthermore, the electropherogram obtained showed partial separation of the different glycopeptides demonstrating the benefit of using CE as the separating technique for this type of characterization. As displayed in Figure 2, particular glycopeptides having a difference of one galactose (meaning a mass difference of 162 Da) could be baseline separated. The capacity to separate peptides having such small differences is clearly interesting since they tend to compete against each other during the ionization process potentially interfering with the relative quantification. Therefore their separation participates to ease their ionization, defining the rare sensitivity regarding glycosylations characterization with the CESI-MS/MS methodology developed. Such sensitivity could be achieved while the entire peptide digest mixture was characterized without glycan release followed by extraction which is commonly performed in glycan analysis by MS [29]. Cetuximab contains a second N-glycosylation sites on the HC: one is located in the Fc/2 domain (Asn299) similarly to trastuzumab while the second one is located in the Fd domain on Asn88 [30]. The CESI-MS/MS experiments on cetuximab indicated the two different sites, each glycosylation site could be precisely located based on the CESI-MS/MS data. Additionally, structural glycan characterization as well as relative quantification could be both established independently for each site in the same experiment.

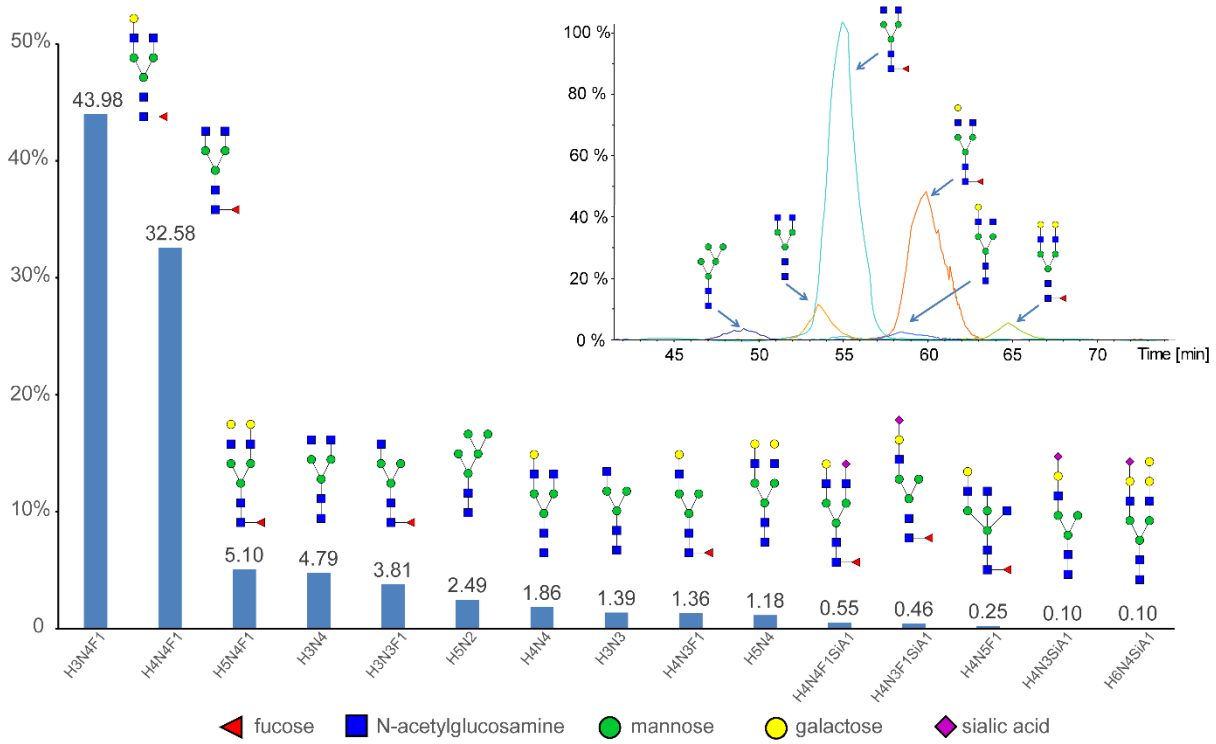

Figure 2. Glycoforms determination obtained for trastuzumab using the CESI-MS/MS method in a single analysis (left-hand side). Extracted Ion Electropherogram (EIE) corresponding to the m/z of the most abundant glycoforms, illustrating the separa-tion selectivity obtained with CE regarding mAbs glycopeptides (right-hand side).

Additional PTMs were also analyzed in the same run. For example, trastuzumab HC N-terminal extremity contains a glutamic acid which can undergo partial cyclization leading to pyroglutamic acid [17]. The extraction of the m/z ratios corresponding to the native N-terminal peptide and the pyroglutamic acid variant migrates as two different peaks separated by several minutes. This results can be explained by the fact that glutamic acid cyclization entails for the AA a mass loss of 17.02 Da. As CE separates compound regarding to their size and charge state in solution, such PTM involves a significant modification of the electrophoretic mobility.

Another common micro-variant is methionine oxidation. This modification implies for the peptide a mass increase of 15.99 Da while the charge density remains the same. In a similar way as previously, the same CESI-MS/MS data highlight the capacity of CE to separate the modified peptide undergoing methionine oxidation from the native peptide (Supporting Information Figure S-2). Those results open perspectives for improved relative quantification regarding the level of occurrence of those modifications similarly to glycosylation characterization.

Deamidation, is associated with the removal of the amide group present on the side chain of asparagine (Asn) and, to a lesser extent, of glutamine (Gln) residues [31]. These modifications are observed by separation methods such as isoelectric focusing (IEF) and cationic exchange chromatography (CEX) in combination with offline MS methods. In contrast, the CESI-MS/MS

method allowed obtaining complete separation between the parent and the degraded peptide (Figure 3). That characteristic in separation could be confirmed for every deamidation hot-spots identified both on trastuzumab and cetuximab (4 different sites each). The ability to separate those modified peptides is particularly important. Deamidation involves a loss of only 0.98 Da; such little difference would lead during the ionization process to competition lowering the sensitivity of the MS signal in the case if both peptides could not be separated. That appears to be quite relevant as the deamidation sites on the studied mAbs exhibited low level of modification usually below 5%, urging the necessity to prevent ionization competition. The excellent separation provided by the CE allows in the case of this characterization to obtain the best sensitivity for both peptides. One direct consequence is that the fragmentation spectra quality could be significantly improved allowing to precisely locate the modified AA even if other Asn or Gln were present in the peptide.

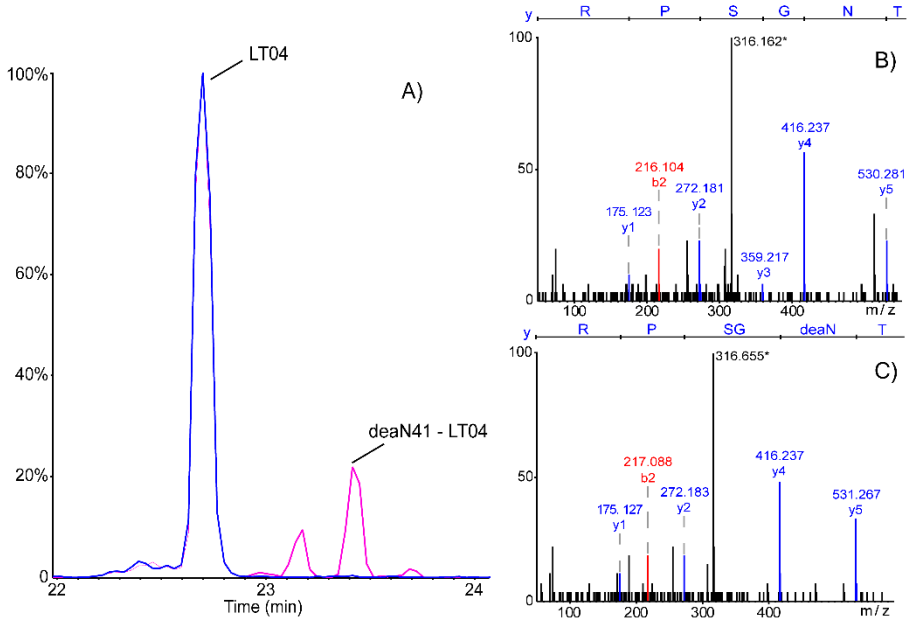

Figure 3. A) EIE corresponding to the m/z of peptide LT04 (light chain, position 40-45) and LT06 with deamidatedAsp41. De-convoluted MS/MS spectra corresponding to B) peptide LT04 and C) LT04deamidatedAsp41 (deamidation represented by deaN).

The last considered PTM is aspartic acid isomerization; this modification is particularly difficult to characterize. Indeed the change of conformation of aspartic acid (Asp) could not induce a significant variation of affinity toward the reverse stationary and requires particular analytical methodologies giving only access to a specified aspect of the protein [31-33]. Furthermore the conformation change does not induce a change in the mass of the peptide thus ESI-MS using hybrid analyzers such as a quadrupole-time of flight (Q-TOF) does not allow to distinguish a potential Asp isomerization. From the CESI-MS/MS data, extraction of the m/z ratio corresponding to a peptide potentially presenting Asp isomerization systematically exhibited two consecutive peaks as it is appearing in Figure 4. The important acquisition rate capacity

provided by the MS used therefore enables, from the CESI-MS/MS data, to obtain the fragmentation spectra for both peak. From fragmentation pattern, MS/MS spectra presented in Figure 4 unambiguously proved that the two peaks are corresponding to the same peptides. Fragmentation is obtained inside the MSCID; in this fragmentation mode the energy induced to neutral particles (usually N2 or Ar) is limited to a few tenths eV. Such energy levels allow to activate the fragmentation of the peptide backbone enabling the detection of specific b and y fragments ions [34]. In the context of this study, two consecutive peaks leading to the same fragmentation spectrum suggests that those CE conditions are enabling the separation of the same peptide having different Asp isomers. From a theoretical aspect, electrophoretic mobility is significantly influenced by the hydrodynamic radius of the molecule. Two similar peptides with the same AA sequence containing different aspartic acid isomers should be differently oriented. This would most likely induce a difference in their respective hydrodynamic radius and therefore implying a difference of electrophoretic mobilities among them. Also the different potentially Asp isomerization sites studied, on both samples, exhibited the same behavior while peptide having no Asp did not presented such characteristic, reinforcing the assertion on separation based on Asp isomerization.

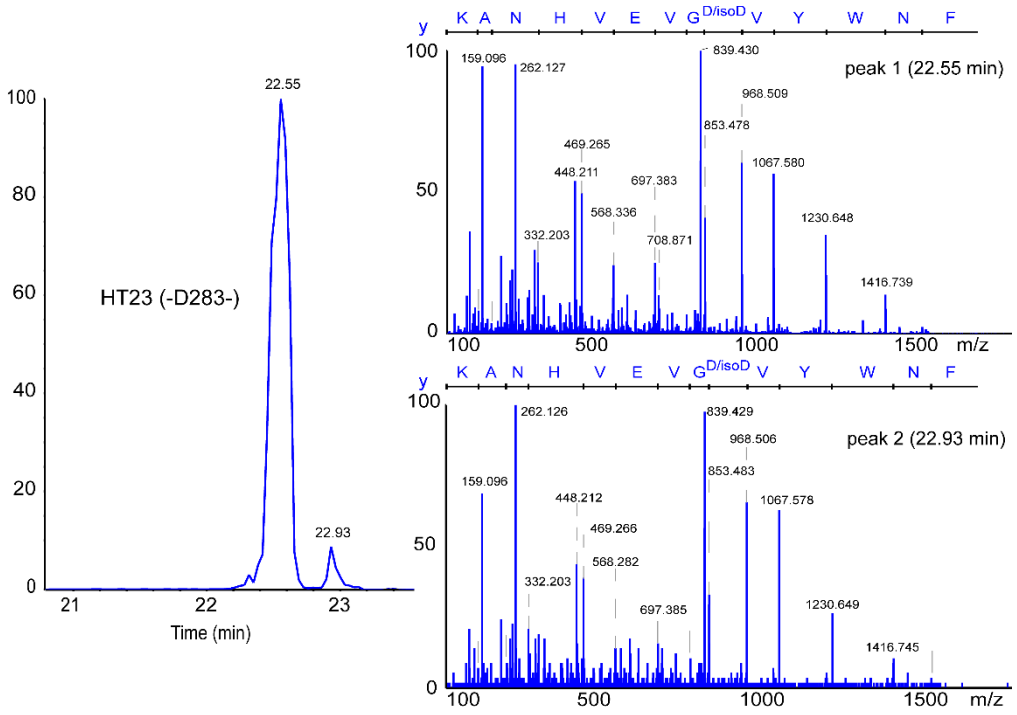

Figure 4. EIE corresponding to m/z for digested peptide HT23 (heavy chain; position 278-291) experiencing aspartic acid isomer-ization. Raw MS/MS spectra for both peaks (right-hand side) demonstrated the same fragmentation pattern additionally to precursor m/z and charge state values.

In order to validate with certainty the capacity of the developed CE method to separate peptides regarding to Asp isomeriza-tion, two peptides were specially synthetized. Those peptides, composed of 20 AAs, have exactly the same AA sequence and contains one Asp;

each synthetic peptide bearing a different Asp isomer. As it is emphasize in the Supporting InformationFigure S-3, several samples composed of a mixture of both synthetic peptides in different ratio were analyzed using the same t-ITP-CESI-MS/MS conditions as the mAbs characterization. Results obtained for the different mixture ratios exhibitstwo consecutive peaks for the m/z ratio corresponding to the synthetic peptide. On the contrary when only a single peptide is injected, the ex-tracted ion electropherogram (EIE) showed only one peak. To reinforce the result, peaks heights illustrate relatively the evolution in proportion of one peptide compared to the other. Such results demonstrates without ambiguity the selectivity of the separation in the case of a peptide experiencing Asp isomerization. This results furtheremphasize the relevance of using CE separation for protein primary structure by MS. It indeed allows to discriminate peptides with Asp isomerization in a robust manner, thereby enriching further the information that can be obtained by MS on a given protein molecule.

Conclusion

To summarize, we report here the use of t-ITP-CESI-MS/MS for the characterization of four different therapeutic mAbs. The instrumental settings used are composed in particular of a CE-ESI-MS interface which has been recently developed. Using a single injection, we were able to characterize the primary structure of those antibodies in a robust manner at an unprecedented level. We managed to obtain the complete AA sequence characterization while only relying on tryptic peptide without miscleavages nor exogenous modifications. It is the first time that such level of characterization could be achieved in a single injection/run, suggesting new approaches for bottom-up proteomics in particular. Simultaneously to the AA sequence, PTMs including glycosylation were also characterized. Results highlighted the benefit of using electrophoretic separation in complement to chromatographic separation which is conventionally applied in this type of study. CE separation selectivity showed the ability to separate peptides having only minor differences while the sensitivity provided by the CESI-MS led to the improvement of the MS/MS characterization. Indeed the opportunity to separate peptides having only an isomerization of one AA or a difference of 0.98 Da enables to cancel ionization competition between the different peptides and explains the capacity of the CESI-MS/MS methodology to characterize in the same experiment the intact and the modified peptide. Similarly CE proved through MS to ease primary structure characterization as it was possible to detect aspartic acid isomerization on several peptides from the same analysis along with its other attributes. Glycosylations were also characterized from the same experiment, thus 15 different glycans could be characterized for trastuzumab, showing the advantages of using CESI-MS improved sensitivity; note that no glycans release was necessary reducing the

sample treatment and the necessity to use different experimental conditions to characterize glycosylation along the other characteristics of the primary structure of the protein. Finally such experiment could be achieved by injecting a quantity of sample corresponding to 200 fmol of digested peptide illustrating the suitability of the CESI-MS/MS method for small available amount of sample. The CESI-MS/MS data reported here indicate that electrophoretic separa tion, combined to the highly efficient CESI interface become a viable alternative to LC-ESI-MS/MS for innovative approaches in MS proteomics such as identifying AA mutations or transcription mismatches.

Ackowledgments

Authors would like to thank Sciex separations Inc. for lending a CESI prototype and Dr. M. Anselme from Sciex separations Inc. for his support. LSMIS would like to thank Philippe Hammann, Lauriane Kuhn and Johana Chicher (Institut de Biologie Moléculaire et Cellulaire, Strasbourg, France) for giving access to their mass spectrometer. Authors would like to acknowledge Dr. Joseph Vlasak (Merck, West Point, PA) for critical reading of the manuscript. The authors would like also to express their gratitude to M. Biacchi (LSMIS, Strasbourg, France), Dr. E. Wagner-Rousset, Dr. D. Ayoub, MC. Janin-Bussat and O. Colas (Centre d'immunologie Pierre Fabre, St Julien en Genevois, France) for helpful discussions around antibody structural characterization.

References

[1] **Jorgenson, J. W.; Lukacs, K. D.** Analytical Chemistry 1981, *53*, 1298-1302.

[2] **Jorgenson, J.; Lukacs, K.** Science1983, 222, 266-272.

[3] **Hjertén, S.** Journal of Chromatography A 1983, 270, 1-6.

[4] **Tiselius, A. U. u.** *The moving boundary method of studying the electrophoresis of proteins*. Almqvist & Wiksell, Uppsala, 1930.

[5] **Longsworth, L. G.; MacInnes, D. A.** Chemical Reviews 1939, *24*, 271-287.

[6] **Mokaddem, M.; Gareil, P.; Belgaied, J. E.; Varenne, A.** Electrophoresis 2008, *29*, 1957-1964.

[7] **Mokaddem, M.; Gareil, P.; Belgaied, J. E.; Varenne, A.** Electrophoresis 2009, *30*, 1692-1697.

[8] **Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M.** Science 1989, *246*, 64-71.

[9] **Tito, M. A.; Tars, K.; Valegard, K.; Hajdu, J.; Robinson, C. V.** Journal of the American Chemical Society 2000, *122*, 3550-3551.

[10] **Schagger, H.; Vonjagow, G.** Analytical Biochemistry 1991, *199*, 223-231.

[11] **Klepárník, K.; Boček, P.** Chemical Reviews 2007, *107*, 5279-5317.

[12] **Salas-Solano, O.; Carrilho, E.; Kotler, L.; Miller, A. W.; Goetzinger, W.; Sosic, Z.; Karger, B. L.** Analytical Chemistry 1998, *70*, 3996-4003.

[13] **Berk, A. J.; Sharp, P. A.** Cell 1977, *12*, 721-732.

[14] **Boodram, S. N.; McCann, L. C.; Organ, M. G.; Johnson, P. E.** Analytical Biochemistry **2013**, *442*, 231-236.

[15] **Busnel, J.-M.; Schoenmaker, B.; Ramautar, R.; Carrasco-Pancorbo, A.; Ratnayake, C.; Feitelson, J. S.; Chapman, J. D.; Deelder, A. M.; Mayboroda, O. A.** Analytical Chemistry 2010, *82*, 9476-9483.

[16] **Moini, M.** Analytical Chemistry 2007, *79*, 4241-4246.

[17] **Beck, A.; Wagner-Rousset, E.; Ayoub, D.; Van Dorsselaer, A.; Sanglier-Cianferani, S.** Analytical Chemistry 2013, *85*, 715-736.

[18] **Hudis, C. A.** New England Journal of Medicine 2007, *357*, 39-51.

[19] **Ayoub, D.; Jabs, W.; Resemann, A.; Evers, W.; Evans, C.; Main, L.; Baessmann, C.; Wagner-Rousset, E.; Suckau, D.; Beck, A.** mAbs 2013, *5*, 699-710.

[20] **Li, S.; Schmitz, K. R.; Jeffrey, P. D.; Wiltzius, J. J. W.; Kussie, P.; Ferguson, K. M.** Cancer Cell 2005, *7*, 301-311.

[21] **Walsh, G.; Jefferis, R.** Nat Biotech 2006, *24*, 1241-1252.

[22] **Gahoual, R.; Burr, A.; Busnel, J.-M.; Kuhn, L.; Hammann, P.; Beck, A.; François, Y.-N.; Leize-Wagner, E.** mAbs 2013, *5*, 479-490.

[23] **Malá, Z.; Gebauer, P.; Boček, P.** Electrophoresis 2013, *34*, 19-28.

[24] **Domon, B.; Aebersold, R.** Science 2006, *312*, 212-217.

[25] **Paizs, B.; Suhai, S.** Mass Spectrometry Reviews 2005, *24*, 508-548.

[26] **Gahoual, R.; Busnel, J.-M.; Wolff, P.; François, Y.; Leize-Wagner, E.** Anal Bioanal Chem 2014, *406*, 1029-1038.

[27] **Wilm, M. S.; Mann, M.** International Journal of Mass Spectrometry and Ion Processes 1994, *136*, 167-180.

[28] **Beck, A.; Wagner-Rousset, E.; Bussat, M. C.; Lokteff, M.; Klinguer-Hamour, C.; Haeuw, J. F.; Goetsch, L.; Wurch, T.; Van Dorsselaer, A.; Corvaia, N.** Current Pharmaceutical Biotechnology 2008, *9*, 482-501.

[29] **Wagner-Rousset, E.; Bednarczyk, A.; Bussat, M.-C.; Colas, O.; Corvaïa, N.; Schaeffer, C.; Van Dorsselaer, A.; Beck, A.** Journal of Chromatography B 2008, *872*, 23-37.

[30] **Beck, A.; Sanglier-Cianférani, S.; Van Dorsselaer, A.** Analytical Chemistry 2012, *84*, 4637-4646.

[31**] Vlasak, J.; Bussat, M. C.; Wang, S.; Wagner-Rousset, E.; Schaefer, M.; Klinguer-Hamour, C.; Kirchmeier, M.; Corvaïa, N.; Ionescu, R.; Beck, A.** Analytical Biochemistry 2009, *392*, 145-154.

[32] **Kern, W.; Mende, R.; Denefeld, B.; Sackewitz, M.; Chelius, D.** Journal of Chromatography B 2014, *955–956*, 26-33.

[33] **Zhang, J.; Yip, H.; Katta, V.** Analytical Biochemistry 2011, *410*, 234-243.

[34] **Mitchell Wells, J.; McLuckey, S. A.** In Methods in Enzymology, Burlingame, A. L., Ed.; Academic Press, 2005, pp 148-185.

3. Additional discussion regarding characterization of mAbs by t-ITP CESI-MS/MS

3.1. Sample preparation

As it has been described previously, the sample preparation was modified for different purposes: improve digestion yield and perform transient isotachophoresis. MAbs structures are ambivalent, the Y shaped quaternary structure is “souple” however their primary structure is quite organized in order to interact selectively with the antigen epitope. That organization tend to limit the access of trypsin to several lysine and arginine which is the cause of proteolytic digestion miscleavages occurrence.

In order to limit the occurrence of peptide with miscleavages, it is required to disrupt the primary structure of the mAbs, without chemical modifying the protein, prior to the tryptic digestion. That precaution enables trypsin to have access efficently to the protein. In the characterization presented in the previous chapter a chaotropic agent composed of guanidine hydrochloride 6M. Using this chaotropic agent, identification results showed a significant proportion of peptides with miscleavages and the complete sequence coverage could be successfully by the identification of digested peptides exhibiting miscleavages. Guanidine hydrochloride was replaced by another chaotropic agent which is Rapigest SF (Waters). That surfactant is originaly used to perform highly efficient tryptic digestion in a reduced time but demonstrated its ability to improved the digestion yield especially for mAbs digestion. After the digestion, lowering the pH of the sample enables to cleave and precipitate the surfactant allowing its elimination

Results showed the use of that surfactant improves significantly the digestion yield. The digestion still is not complete for the entire sample however all lysine and arginine can be accessible to trypsin as it has been demonstrated by the complete sequence coverage performed using only tryptic peptides without miscleavages (figure 1).

3.2. Y / B fragment ions recovery

The characterization performed by t-ITP CESI-MS/MS enabled to systematically obtain more than 70% of the y/b ions induced by the fragmentation of the digested peptides. This results imply the possibility to retrace the exact amino acids succession on an extended part of the protein and also emphasize the important amount of information fragmentation spectra obtained using the CESI-MS/MS system are providing.

To determine the proportion of y/b ions retrieved, all the MS/MS spectra leading to the identification of a peptide were considered, including peptides containing miscleavages or

PTMs as long as they could be identified. Fragment ions identification from a single analysis were considered to estimate the proportion of y/b ions retrieved. The figure 5 shows the fragment ions obtained in the case of trastuzumab.

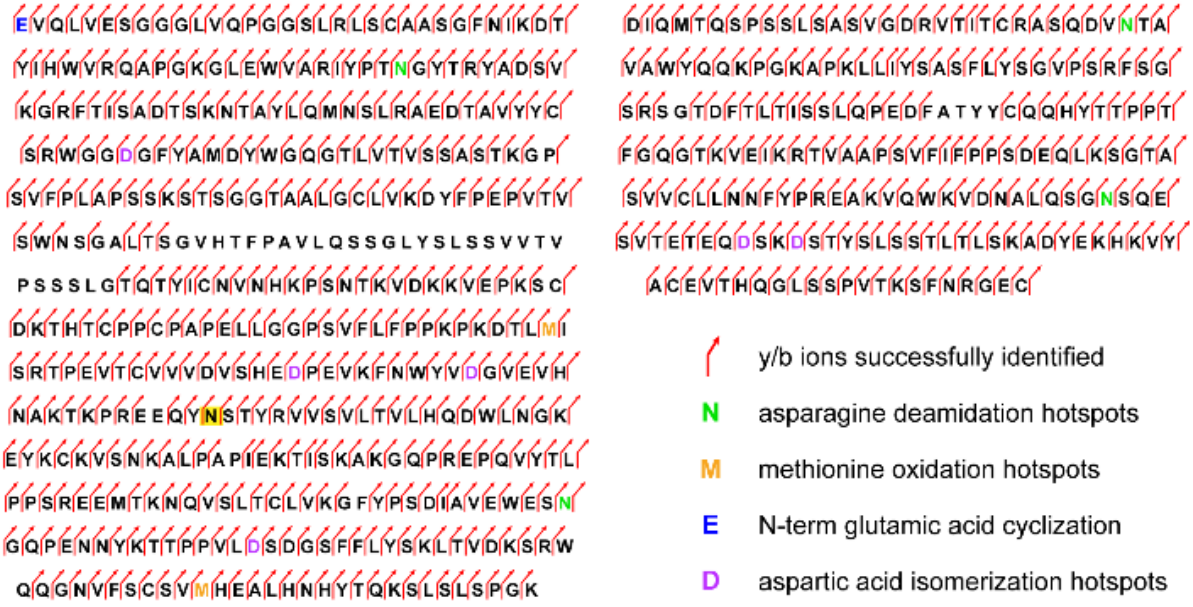

Figure 5. Schematic representation of the fragment ions identified by the characterization of trastuzumab by t-ITP CESI-MS/MS.

The nature of the fragmented peptide is clearly influencing the proportion of its fragment which can be identified. Digested peptides composed of a few amino acids enables the identification of all the fragment whereas the fragment ion identification is less efficient in the case of bigger peptides. That trend is linked first to the collision energy range used for the fragmentation in CID. Also, MS/MS spectra are showing the different fragmentation possible for the selected precursor ion. As the number of amino acids is increased, the number of possibilities is mechanically increased as well. As a consequence, it is necessary to ionize and select an important number of the selected precursor ions to generate as much fragment ions which will give a decent signal for all the possible y/b fragment leading to their identification.

The improvement of the digestion yield is showing its interest. A complete digestion is generating a mixture composed of the smallest peptides, those peptides are allowing the identification of a maximum number of fragment ions. Indeed the part of the protein which could not see their fragment ions retrieved are located on high molecular mass peptides. Those peptides have not been cleaved further because they do not contain a lysine or an arginine. In that aspect it is possible to affirm that the proportion of y/b ions is partially conditioned by the peptides generated from its digestion which explains in this work differences from one mAb to another. The ionization efficiency of the CESI-MS system is also participating to the ability to retrieve almost the entire y/b ions.

The beneficial ionization efficiency provided by the CESI-MS is a direct consequence of the separation conditions used for the CE separation and the ionization process. Indeed the internal diameter of the capillairy (30 µm) is enabling the formation of droplets having a reduced volume. In addition the flow rate of BGE generated by the electroosmotic flow is only of a few tenth nL/min which is favorable to ESI ionization process by enabling to evaporate the BGE easily.

The use of an electrophoretic separation before the MS/MS analysis demonstrated previously the ability to separate and transfer successfully to the MS an extended variety of peptides regarding their chemical structures and properties. That characteristic cumulated to the ability to retrieve almost in totality the y/b ions from peptides fragmentation is particularly interesting. Those characteristics suggest that the implementation of the t-ITP CESI-MS/MS method can give the opportunity to develop alternative MS characterization strategies especially because the consistency of the performance of the system.

3.3. Electrophoretic separation of the modified peptides

The following sub-chapter is describing the separation of digested peptides affected by PTMs which results could only be briefly mentioned in the article presented previously. That sub-chapter also contains additional experiments performed to validate unambiguously the separation of peptides because of aspartic acid isomerization.

a) Methionine oxidation

Similarly to the separation of peptides because of a deamidation of an asparagine, the separation selectivity provided by the implementation of CE enables to generate the separation of a peptide which has an oxidated methionine from the very same peptide without modification. As a reminder, methionine oxidation involves a mass shift of 15.99 Da however it does not imply a change of the charge state of the amino acid at the BGE pH.

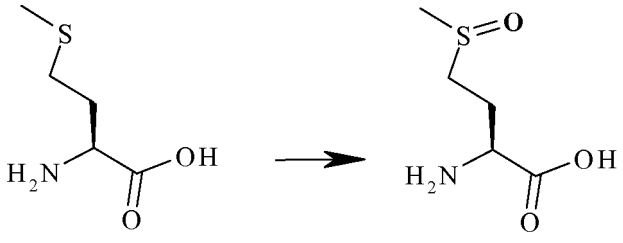

Figure 6. Reaction mechanism describing the oxidation of methionine lateral chain to give methionine sulfoxide.

Characterization of the primary structure performed using the t-ITP CESI-MS/MS methodology developed in this work demonstrated the ability of the conditions used to perform the electrophoretic separation lead to a change of the mobilities of the modified peptide compared

to the unmodified one. Study of the corresponding MS/MS spectra proves the separation without any ambiguity as it is emphasize in figure 7.

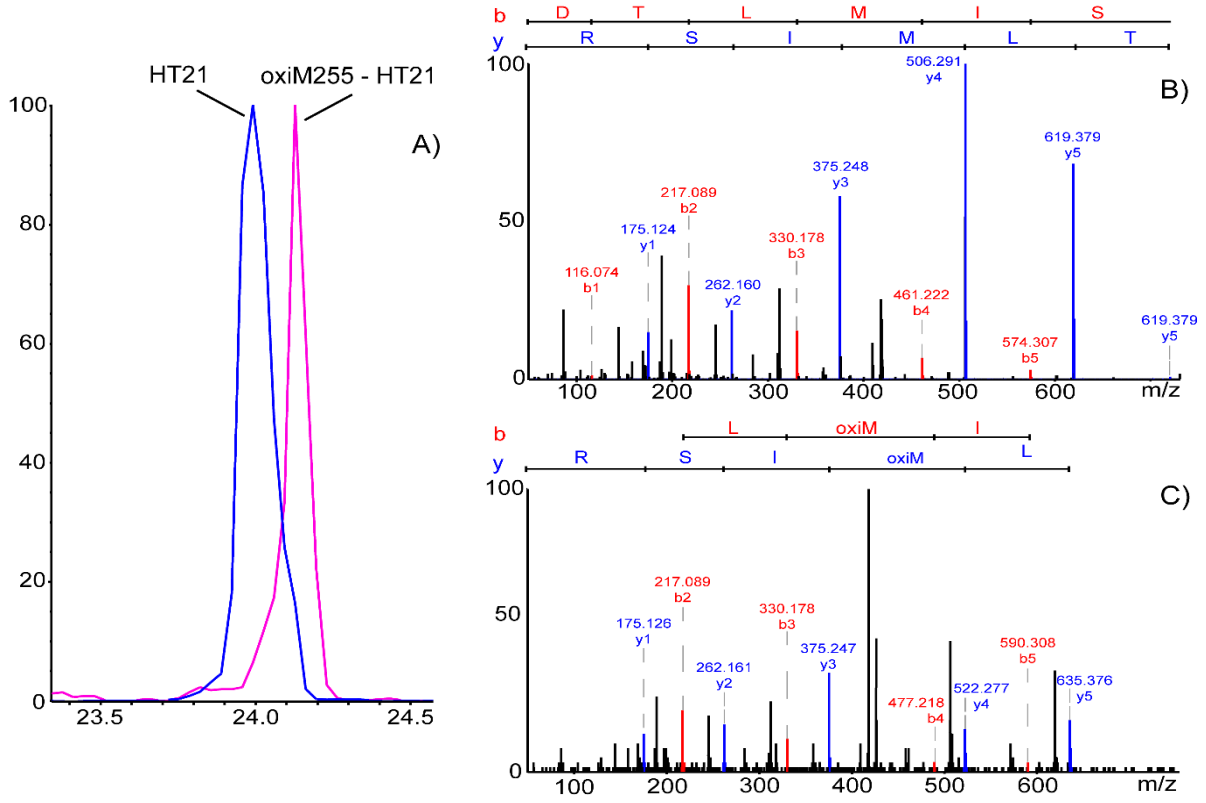

Figure 7. A) EIE corresponding to the m/z of trastuzumab peptide HT21 (heavy chain, position 252-258) and HT21 with oxidated Met255. MS/MS spectra corresponding to B) peptide HT21 and C) HT21 with oxidated met255 (oxidation represented by oxiM).

b) Aspartic acid isomerization

The separation of similar peptides because they are composed of aspartic acid in two different conformation has been described previously. The isomerization of aspartic acid does not imply a mass change of the entire peptide or a distinguishable change in the fragmentation which mean that it can not be identified from the MS or MS/MS data. As a consequence, in case of a co-migration, this modification could not be identified. Therefore the separation in this case is interesting because it allows to give additional information in the context of the primary structure characterization.

As separation in based also on the difference of hydrodynamic radius of the analytes, a conformation change of one amino acid composing the peptide is modifying the hydrodynamic radius of the affected peptide. A change of the hydrodynamic radius of the modified peptide leads to a difference of electrophoretic mobilities between the two peptides which is observed during the electrophoretic separation.

The article previously presented shows that the MS/MS spectra corresponding to both peaks in CE leaded to the same identification suggesting that the separation is induced by different

aspartic acid isomers. In order to confirm this assumption, additional experiments have been performed using synthetic peptides.

Those synthetic peptides were designed to be similar to a digested peptides, their amino acid sequences are as follow:

- $_2$HN-GLEWIGYISY**D**GTNNYKPSLK-OH
- $_2$HN-GLEWIGYISY**isoD**GTNNYKPSLK-OH

Different mixtures composed of those synthetic peptides were analyzed using the same t-ITP CESI-MS/MS method as the mAb characterization. Peptide ratios were different from one mixture to another in order to distinguish and identify the separation of the peptides. Figure 8 shows the separation obtained fro the different mixtures.

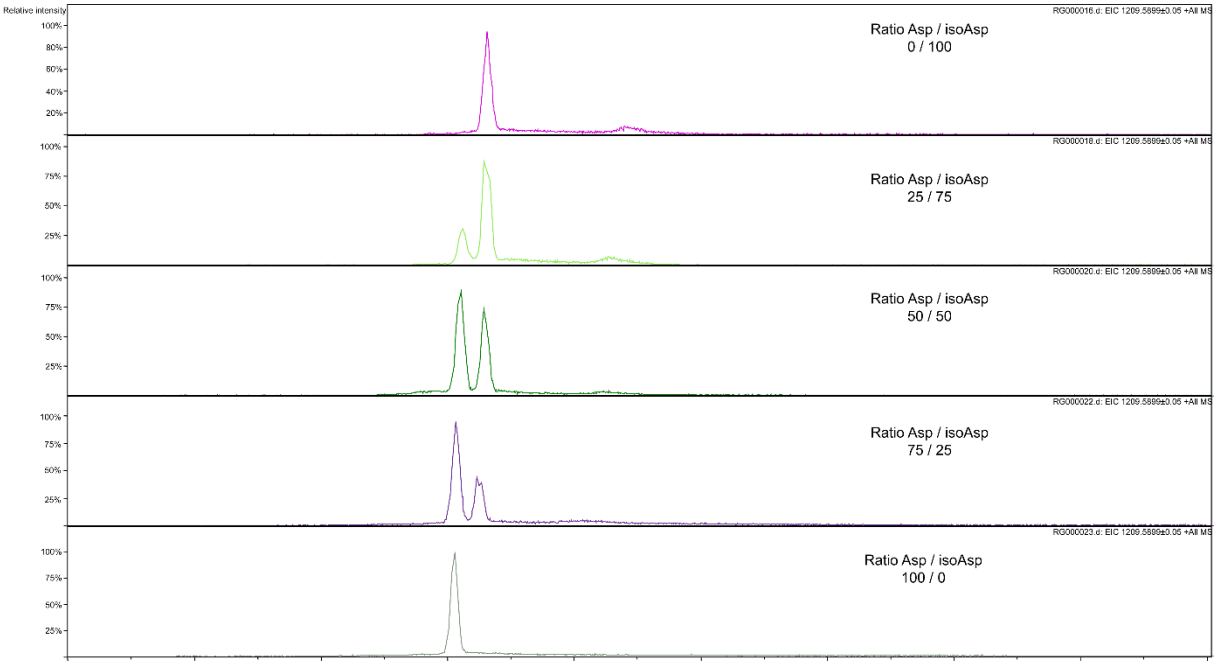

Figure 8. EIE corresponding to the m/z of the synthetic peptides considered for Asp isomerization separation study. Theoretical mass: 1209.5899± 0.05 (charge state 2+). Ratios expressed between two peptide conformations are based on absolute injected quantities.

The extraction of the m/z ratio corresponding to the peptides shows systematically two distinctive peak in the case a mixture of both peptides and one peak when only one of the peptide conformation was injected. In addition the abundances are in complete agreement with the ratio prepared. Those results validate without any ambiguity that the CE separation conditions used for the characterization allows to separate peptides based uniquely on aspartic acid isomerization.

These results was published in *Journal of Mass Spectrometry* (Wiley) in 2015.

4. Conclusion

Les travaux présentés dans cette partie détaillent l'utilisation du système CESI-MS afin de réaliser en une seule injection, la caractérisation complète de la structure primaire d'un anticorps et de démontrer l'intérêt d'utiliser une séparation électrophorétique en amont de l'analyse MS/MS pour fournir une caractérisation fine de la structre primaire de ces protéines. Dans le cadre de ces travaux une stratégie de protéomique *bottom-up* a été adoptée. La méthodologie a été significativement adaptée du point de vue de la préparation d'échantillon afin d'améliorer l'efficacité de la digestion et la compatibilité entre le contenu de l'échantillon et la séparation électrophorétique. L'isotachophorèse transitoire (t-ITP) a également été introduite dans la méthode d'analyse. Le bénéfice de cette technique est double : la t-ITP est une méthode de pré-concentration utilisée en CE, elle permet notamment de travailler avec des concentrations proches de celles utilisées en protéomique classique. De plus la t-ITP permet une standardisation du contenu de l'échantillon et tend à améliorer dans une certaine mesure l'efficacité de la séparation. Le traitement des données MS/MS à également dû être adapté en partie afin d'améliorer l'identification des peptides de faibles masses moléculaires. Différents mAbs ont été caractérisés. Ces protéines regroupent diverses problématiques de caractérisation généralement rencontrées dans une grande variété de protéines. Des réplicas techniques ont été également analysés afin de démontrer la robustesse de la méthodologie développée.

Les résultats de ces travaux ont montré pour une injection correspondant à 200 fmol de peptides, la possibilité d'obtenir la caractérisation complète de la séquence d'acides aminés de manière systématique sur les différents mAbs analysés en utilisant la méthodologie développée. La caractérisation de la séquence d'acides aminés a pu être effectuée uniquement par l'identification de peptides ne portant ni *miscleavages* ni PTMs. Cette caractéristique est particulièrement intéressante et démontre que l'utilisation de la CE couplée à la MS permet la séparation et l'identification de l'ensemble des peptides composant le mélange indépendamment de leur structure. La caractérisation de la totalité de la séquence uniquement par l'intermédiaire de peptides intacts offre la possibilité de générer des cartes peptidiques qui pourront être utilisées pour le contrôle de la production de mAbs. L'annotation des spectres MS/MS a également montré que systématiquement plus de 70% des ions y/b ont pu être identifiés, cette valeur dépassant même les 90% dans le cas de trastuzumab. Ce résultat en particulier, a permis d'en déduire l'enchainement complet des acides aminés composant le domaine variable de trastuzumab. La fonction principale du domaine variable des mAbs est de reconnaitre et d'intéragir avec l'antigène. La possibilité de caractériser de manière aussi précise une partie aussi importante de la protéine apparait donc comme un atout majeur. Ces résultats de *peptide mapping* corrélés sur les différents mAbs montrent que

la mise en œuvre du couplage CESI-MS permet d'améliorer le degré de caractérisation par MS par rapport à une approche protéomique utilisant une méthodologie analytique conventionnelle (nanoLC-MS/MS). La possibilité d'identifier une large majorité de fragments et l'ensemble des peptides composant le digestat ouvrent la voie à de nouvelles applications pour la caractérisation de protéines notamment en se détachant progressivement des banques de données de protéines dont l'analyse protéomique est souvent tributaire ou encore d'identifiés des erreurs de transcription.

En concomitance avec la caractérisation de la séquence d'acides aminés, les différentes glycosylations portées par ces mAbs ont également été analysées. L'incorporation de la t-ITP a permis d'augmenter de manière significative le nombre de glycoformes pouvant être caractérisés. Ainsi dans le cas de trastuzumab, 15 glycosylations différentes ont pu être identifiées. L'intensité du signal MS correspondant aux glycopeptides a aussi été utilisée pour estimer l'abondance relative de chacune des glycosylations. Les résultats montrent la compatibilité de la gamme dynamique délivrée par le couplage CESI-MS pour cet aspect de la caractérisation : la glycosylation la plus abondante représentant une abondance de 44% tandis que la glycosylation la plus faiblement représentée possède une abondance de 0,1%. Notons que parmi les glycosylations identifiées, certaines contiennent un acide sialique. La sélectivité de la séparation joue un rôle dans la détection de glycoformes si faiblement abondants. L'étude de la séparation a montré que la sélectivité de la CE entraine la séparation complète des glycopeptides possédant un galactose de différence. De la même manière, les glycopeptides portant une différence d'un fucose se sont vu partiellement séparés. La séparation des différents glycopeptides tend à minimiser les effets de compétition au moment de l'ionisation (suppression d'ions), ce qui se traduit au niveau du résultat par une sensibilité importante et une quantification relative dont la justesse est améliorée. La méthodologie décrite dans cette partie permet la caractérisation des glycannes encore liés de manière covalente sur le peptide. Les spectres MS/MS illustrent la fragmentation du glycanne ce qui laisse en déduire la structure des glycannes.

Dans le cas de cetuximab, dont l'une des particularités est d'abriter deux sites de glycosylations, l'analyse par CESI-MS/MS a permis de considérer chaque site de glycosylations indépendamment. Les résultats de cette étude ont permis d'identifier les différentes glycosylations portées par chacun des sites de N-glycosylation et d'en déduire la structure pour la plupart d'entre elles.

Dans le cadre de la caractérisation de la structure primaire de la protéine, les PTMs ont également été étudiées. La caractérisation de PTMs illustre clairement la pertinence de l'utilisation d'une séparation électrophorétique en amont de l'analyse MS/MS. En effet ces

modifications induisent une variation de la mobilité électrophorétique du peptide ce qui entraine, sous l'effet du champ électrique, une séparation du peptide modifié et de son homologue n'ayant pas subi la modification. L'application de la CE a permis la séparation de peptide dont les modifications ont une influence significative sur la masse du peptide (cyclisation de l'acide glutamique N-terminal, oxidation de méthionine) mais également de modifications dont l'impact sur la masse du peptide est faible (déamidation d'asparagine).

En particuliers, ces travaux démontrent que l'utilisation de la CE comme méthode séparative permet d'obtenir une séparation complètement résolue pour des peptides de séquence identique et dont la seule différence est la présence d'un isomère différent d'acide aspartique. Ce résultat est tout fait significatif car l'isomérisation de l'acide aspartique n'entraine pas de variation de la masse du peptide, les différents isomères ne peuvent être identifiés en MS par l'intermédiaire d'un analyseur TOF et nécessitent des stratégies analytiques spécifiques pour être caractérisés. Les travaux montrent que l'identification de ces isomères a pu être intégrée dans la méthodologie globale de la caractérisation de la structure primaire des mAbs étudiés. L'application d'une séparation électrophorétique permet par conséquent dans ce cas d'étoffer le degré de caractérisation pouvant être atteint.

Les résultats présentés illustrent donc la possibilité de réaliser en une seule analyse une caractérisation avancée de la structure primaire de protéines thérapeutiques. Différents mAbs ont été étudiés afin de démontrer la robustesse de la caractérisation obtenue par l'intermédiaire de la méthodologie développée. De la même manière, des réplicas techniques ont également été analysés afin de prouver la reproductibilité de la caractérisation avec succés.

Les mAbs étudiés étant relativement bien détaillé dans la littérature, il convient de démontrer l'applicabilité de la stratégie de caractérisation par t-ITP CESI-MS/MS dans le cas de protéines inconnues ou ayant subies des modifications.

III - Evaluation de la biosimilarité d'anticorps monoclonaux par isotachophorèse transitoire et CESI-MS/MS

1. Introduction

Les anticorps monoclonaux (mAbs) ont été introduit en tant que traitement thérapeutique seulement à la fin des années 1980, ils représentent à l'heure actuelle la classe d'agent thérapeutique qui enregistre la croissance la plus importante. Les mAbs du fait de leur spécificité pour l'antigène associé ouvrent de nouvelles perspectives en termes de traitement thérapeutique, de plus certaines de leurs propriétés sont favorables à l'application pour le traitement de pathologies (réductions des effets secondaires, pharmacocinétique et pharmacodynamique favorables). A l'heure actuelle, une quarantaine de mAbs ont été approuvés comme traitement thérapeutique dans le monde et 30 autres mAbs sont en Phase 3 d'essai clinique. Les champs d'application de ce type de protéines sont principalement l'oncologie, ils sont également utilisés dans le traitement de maladie inflammatoire, de maladies auto-immune ou plus récemment pour le traitement de la maladie d'Alzheimer.

Les brevets protégeant la première génération de mAbs, dont plusieurs *blockbusters* (traitements générant plus d'un milliards de dollar de chiffre d'affaires par an), tomberont dans les prochains mois dans le domaine public. La fin de cette protection offre l'opportunité à des entreprises tierces de produire des versions « bio-générique » de ces mAbs. Ces copies portent le nom de biosimilaires.

Les anticorps monoclonaux sont des protéines recombinantes produites à partir de lignées cellulaires sélectionnées et de ce fait protégées. La complexité structurale des mAbs et le fait qu'ils soient soumis à d'importantes micro-hétérogénéités sont autant de difficultés s'opposant à la production d'une protéine strictement similaire par des entreprises tierces comme cela est communément le cas dans le domaine des petites molécules. Prenant en compte ce type de limitations les autorités du médicament, EMA et FDA en pointe, travaillent depuis plusieurs années sur l'établissement de lignes directrices pavant la voie à l'approbation de biosimilaires sans recourir à l'ensemble des essais cliniques comme pour l'approbation d'un nouveau traitement. Ces directives prennent en compte les limitations induites par la production de protéines recombinantes et définissent des critères de qualité (*critical quality attributes*, CQA) qui doivent impérativement être communs entre un mAb original et un candidat pour être considéré comme biosimilaire, ces critères concernent notamment la structure de la protéine. Ces critères suggèrent ainsi que l'immunogénicité aussi bien que les PK/PD entre un mAb et

son biosimilaire ne sont pas significativement différentes. Ces directives récemment publiées ont par conséquent pour but d'aider à la démonstration de la biosimilarité entre un mAb déjà commercialisé et un candidat biosimilaire.

Preuve de l'intérêt industriel pour le développement de mAb biosimilaires, l'EMA a approuvé la commercialisation du premier mAb biosimilaire (infliximab) fin de l'année 2013. L'évaluation de la biosimilarité entre deux protéines implique une caractérisation complète de sa structure et de propriétés physico-chimiques tout comme l'étude de ses PK et PD. Certaines différences mineures peuvent être tolérer, toutefois il convient de souligner que l'émergence de différences significatives pourra faire l'objet d'essais cliniques afin de démontrer que celles-ci n'induisent sur les propriétés toxicologiques et cliniques du biosimilaire.

Dans le but de fournir une caractérisation structurale la plus aboutie possible, l'application de méthodologie analytique innovante pourra jouer un rôle important dans le développement et la validation de biosimilaires mais également de mAbs de nouvelle génération. Il est impératif de développer des outils analytiques permettant d'obtenir une large gamme d'information concernant la structure de la protéine de manière fiable et robuste. De tels outils pourraient se révéler des atouts majeurs d'un point de vue sécurité et pourraient aider aux développements de nouveaux traitements. La possibilité par le biais de données analytiques d'éliminer durant la phase de R&D un candidat biosimilaire non satisfaisant permettrait d'améliorer les chances de réussite dans le contexte du développement d'un biosimilaire. La spectrométrie de masse (MS) joue un rôle de plus en plus important en ce qui concerne la caractérisation de mAbs notamment du fait qu'elle permet d'accèder à une grande variété d'informations concernant la structure de la protéine. Pourtant la MS nécessite l'utilisation de différentes techniques orthogonales (RP-LC, SEC, CEX) afin de réaliser l'étude de la protéine sur les différentes facettes qui composent sa structure.

Les travaux de thèse détaillés dans cette partie décrivent l'application de la méthodologie t-ITP CESI-MS/MS, développée dans la partie précédente, pour l'étude de biosimilarité de plusieurs mAbs. Ces travaux ont pour objectif de démontrer ou d'infirmer la similarité entre un mAb déjà commercialisé et un candidat biosimilaire du point de vue de la structure primaire de ces protéines. Trasuzumab et cetuximab ont ainsi été confronté chacun à un candidat biosimilaire respectif. Les comparaisons ont été effectuées en utilisant les données provenant d'une seule injection pour chacun des échantillons considérés.

Les résultats obtenus pour la caractérisation de la séquence d'acides aminés ont montré pour chacun des mAbs analysés la possibilité de caractériser la totalité de la séquence en une seule analyse. Dans chacun des cas, la caractérisation est basée sur l'identification de peptides sans *miscleavages* ni modifications post-traductionnelles (PTMs). Les données CESI-MS/MS

ont permis dans le cadre de l'étude trastuzumab/biosimilaire de mettre en évidence une différence d'un unique acide aminé entre les deux protéines. Les données ont ainsi permis de caractériser rapidement la présence d'une différence majeure au niveau de la séquence d'acide aminés du candidat biosimilaire, mais également de déterminer l'acide aminé en cette position dans le cas du candidat. Dans le cadre de l'étude de la biosimilarité cetuximab/biosimilaire, les données CESI-MS/MS ont permis de démontrer sans ambiguïté une complète similitude entre cetuximab et le candidat du point de vue de la séquence d'acides aminés.

Concernant les différentes glycosylations portées par les mAbs étudiés. Les glycannes ont pu être caractérisés du point de vue de leur structure et les abondances relatives ont été estimées, afin de comparer les profils d'expression des glycosylations entre les mAbs originaux et leur candidat biosimilaire. Les résultats de l'analyse CESI-MS/MS ont montré la capacité de cette méthodologie à permettre pour chacun des échantillons étudiés l'identification d'un nombre significatif de glycosylations (entre 10 et 16 selon le mAb étudié). Les comparaisons des profils de glycosylation entre les mAbs originaux et leur candidat biosimilaire démontrent la capacité de la méthode pour caractériser des différences mineures entre les profils de glycosylation. La comparaison entre trastuzumab et le candidat biosimilaire a mis en évidence uniquement des différences d'abondance entre les deux échantillons.

Pour cetuximab, les deux sites de glycosylations ont pu être caractérisés de manière totalement indépendante. Les données de CESI-MS/MS montrent des différences d'abondances entre cetuximab et son candidat biosimilaire. De plus les données recueillies ont servi pour prouver qu'une partie des glycosylations exprimées par le candidat biosimilaire n'étaient pas présente chez cetuximab, ce qui a abouti à l'impossibilité de considérer ces deux protéines comme étant biosimilaires.

Dans le contexte de l'évaluation de la biosimilarité entre les échantillons de mAbs originaux et les candidats biosimilaires, les PTMs d'intérêt ont également été caractérisées. Les données issues des analyses CESI-MS/MS ont permis l'identification de ces PTMs, le degré de modification a aussi été estimé à partir des données. L'objectif de cet aspect est de déceler des modifications surexprimées pouvant mettre en doute l'activité du candidat biosimilaire par rapport au mAb original. Les résultats présentés dans cette partie montrent qu'il a été possible d'identifier chacune des PTMs d'intérêts présentes sur les échantillons caractérisés, quelle que soit la nature de la modification. Les comparaisons effectuées, sur cet aspect de la protéine, entre les mAbs originaux et leur candidat biosimilaire ont démontré la capacité de la méthode à mettre en évidence des différences de degré de modification relativement faibles.

Dans la partie précédente, nous avons démontré la séparation de peptides portant une isomérisation d'un acide aspartique. Cette caractéristique a pu être mise à profit dans les travaux présentés ci-après. Ainsi les niveaux de modifications ont pu également être obtenus pour cette modification, en l'intégrant directement au processus de caractérisation.

Les résultats de la caractérisation réalisée par t-ITP CESI-MS/MS ont permis de conclure sur la biosimilarité de chacun des candidats aussi bien dans le cas de trastuzumab ou cetuximab. D'autant plus, la qualité des données générées et la diversité de la caractérisation pouvant être accomplie en une seule analyse a permis dans chacune des études de mettre en lumière l'aspect du candidat qui n'a pu être jugé comme suffisamment similaire au mAb original.

2. Monoclonal antibodies biosimilarity assessment using transient isotachophoresis capillary zone electrophoresis-tandem mass spectrometry

Rabah Gahoual, Michaël Biacchi, Johana Chicher, Lauriane Kuhn, Philippe Hammann, Alain Beck, Emmanuelle Leize-Wagner, Yannis-Nicolas François, ***mAbs*** 2014, 6(6), 1464-1473.

Abstract

Out of all categories, monoclonal antibody (mAb) therapeutics attract the most interest due to their strong therapeutic potency and specificity. Six of the ten top-selling drugs are antibody-based therapeutics that will lose patent protection soon. The European Medicines Agency has pioneered the regulatory framework for approval of biosimilar products and approved the first biosimilar antibodies by the end of 2013. As highly complex glycoproteins with a wide range of micro-variants, mAbs require extensive characterization through multiple analytical methods for structure assessment rendering manufacturing control and biosimilarity studies particularly product and time-consuming. Here, capillary zone electrophoresis coupled to mass spectrometry by a sheathless interface (CESI-MS) was used to characterize marketed reference mAbs and their respective biosimilar candidate simultaneously over different facets of their primary structure. CESI-MS/MS data were confronted between approved mAbs and their biosimilar candidates to prove/disconfirm biosimilarity regarding recent regulation directives. Using only a single sample injection of 200 fmol, CESI-MS/MS data enabled 100% amino acids (AA) sequence characterization, which allows a difference of even one AA between two samples to be distinguished precisely. Simultaneously glycoforms were characterized regarding their structures and position through fragmentation spectra and glycoforms semiquantitative analysis was established, showing the capacity of the developed methodology to detect up to 16 different glycans. Other posttranslational modifications hotspots were characterized while their relative occurrence levels were estimated and compared to biosimilars. These results proved the value of using CESI-MS because the separation selectivity and ionization efficiency provided by the system allowed substantial improvement in the characterization workflow robustness and accuracy. Biosimilarity assessment could be performed routinely with a single injection of each candidate enabling improvements in the biosimilar development pipeline.

Introduction

Although monoclonal antibodies (mAbs) were introduced as treatments for disease in the late 1980s (muromonab-CD3 was approved in 1986), they currently represent the most rapidly growing category of therapeutic molecule [1]. Some of their properties naturally explain such success. For example, their therapeutic efficiency, reduction of side-effects, favorable pharmacokinetic (PK) and pharmacodynamics (PD) lead to an intensive appeal in research and development (R&D) activities regarding this type of protein for the last decade [2]. Currently more than 40 mabs have been approved by regulation agencies, such as the US Food and Drug Administration (FDA) or the European Medicine Agency (EMA), and 30 additional candidates are currently in Phase 3 evaluation [3, 4]. Their applications are mainly in the field of oncology, inflammation, immune mediated disorders and neurological diseases, including Alzheimer's treatments [5, 6]. Patents protecting the first generation blockbuster mAbs will expire in the next five years, giving the opportunity to many companies to produce "biogeneric versions". These copies are referred as biosimilars or in some cases follow-on biologics. A biosimilar is defined by the EMA as a biological medicinal product that contains a version of the active substance of an already authorized original biological medicinal product [7]. Recombinant mAbs are produced from proprietary cell lines; their structural complexity and important micro-variability induced by a given production process explain the greater difficulty that generic companies have in producing exactly reproduction of biologic molecules compared to small molecules. In the past few years, regulation agencies have worked toward establishment of guidelines that, while taking into account such limitations, determine critical quality attributes (CQAs) that must be in common between a reference mAb and its biosimilar suggesting that their immunogenicity, as well as their PK/PD, should not be significantly different. Those guidelines should help demonstrate similarity between a biosimilar candidate compared to the reference product. The work by EMA and FDA led to the publication of guidelines [7] and to the approval for marketing in the EU of a first biosimilar antibody (infliximab), thus paving the way for the whole product class [8]. Biosimilarity assessment includes extensive physicochemical characterization likewise PK and PD study, performed in a comprehensive manner. As different structural heterogeneities emerged from comparison of a biosimilar candidate with the reference molecule, more complementary studies should be performed in order to demonstrate the absence of toxicological and clinical effect [9]. To perform a complete structural characterization in such a context, innovative analytical methodologies have a major role to play for the development and approbation of biosimilars, as well as for next generation mAbs. The possibility of obtaining a large variety of structural information over different aspects of the protein (e.g. AA sequence, posttranslational modifications, disulfide bonds, glycosylation), in an accurate and robust manner, is obviously

necessary and could be an important asset for safety and product development. One illustration of that trend is the recent intensification of research work aiming to address new issues regarding mAbs characterization [10-12]. A crucial point that could significantly improve the biosimilar development pipeline is to obtain suitable structural information in order to quickly eliminate unsatisfying candidates. Mass spectrometry (MS) has progressively taken an important role in the characterization of therapeutic mabs due to its outstanding selectivity, sensitivity and the possibility to obtain structural information. For example, characterization of therapeutic mAbs by MS has been applied to product and process development [13, 14] and batch consistency assessment. [15] However, MS still needs to be used in concomitance with orthogonal analytical techniques to be able to focus on the different facets of the protein including mainly liquid chromatography (LC) [16-22]. Recently, we described a method using capillary zone electrophoresis (CZE) coupled to tandem MS with a sheathless interface to perform the simultaneous characterization of several aspects of a protein in one injection, including AA sequence and posttranslational modifications (PTMs) [23].

This sheathless CE-ESI-MS interface, referred to as CESI-MS, has been developed based on an original design proposed by Moini et al [24]. Detailed description of this interface has been given by Haselberg et al [25] The CESI system has been used to perform successful CE-ESI-MS experiments in proteomics [26, 27] metabolomics [28] intact proteins [25] therapeutic small molecules [29] and as an infusion platform [30], and it demonstrated a drastically increased sensitivity compared to sheath-liquid CE-MS interfaces because of its low operating flow rates below 100 nL/min.

In the present work, we developed a transient isotachophoresis CE-ESI-MS/MS methodology, with the CESI-MS system, to first obtain the characterization of several mabs. This extensive characterization over different levels of the protein, was performed in a single analysis of each sample. In a second time, each mab was compared to a candidate biosimilar to establish if the methodology developed could be used to assess the biosimilarity between two samples. The comparison was made using the data from a single injection of each sample. Transient isotachophoresis (t-ITP) was introduced in this work: t-ITP is a preconcentration method widely used with CZE [31]. The use of t-ITP enables, in the case of bottom-up proteomics, use of sample concentrations similar to those required for nanoLC-MS/MS analysis [26].

The mAbs included are trastuzumab, cetuximab and their respective candidate biosimilars. Trastuzumab is a humanized immunoglobulin gamma 1 (IgG-1) directed against the HER2/neu receptor overexpressed in around 20-30% of invasive breast cancer (HER2 positive) and more recently to treat HER2+ gastric cancer [32, 33]. It was first approved as a therapeutic in the US and EU in 1998 and 2000, respectively. Trastuzumab is composed of two heavy chains

(HC) and two light chains (LC) representing a total of 1328 AA. Trastuzumab HC bears an N-glycosylation consensus site on the asparagine (Asn) in position Asn-300 [2] (Figure 1).

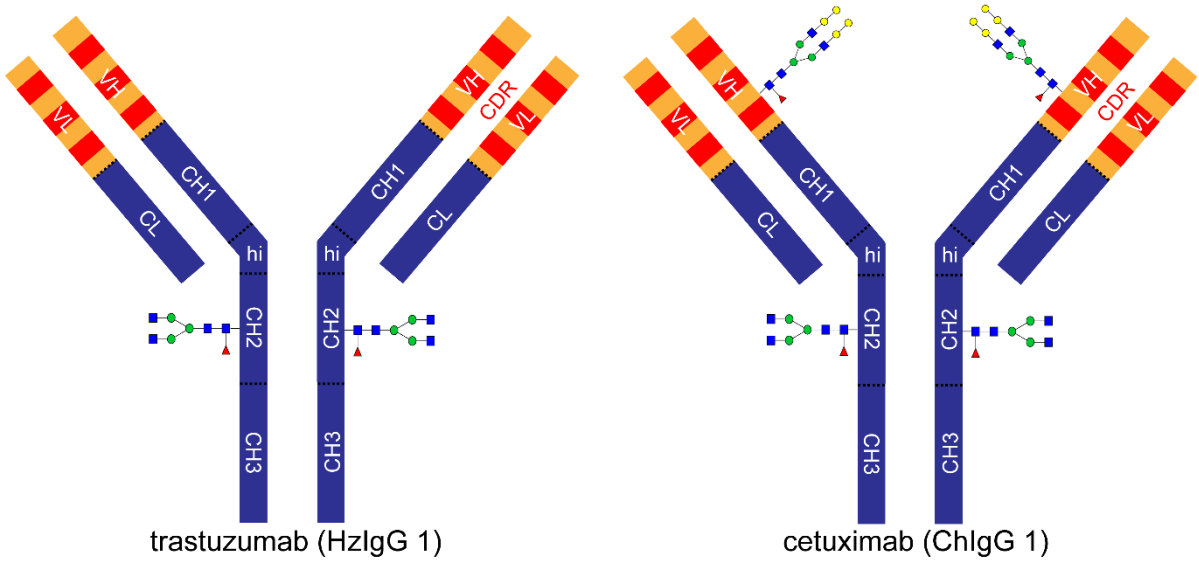

Figure 1. Schematic representation of trastuzumab and cetuximab. Blue parts of the protein represent the constant domain of the mAbs. Yellow parts of the scheme represent the variable domain of the mAbs and red parts represent the complementary determining region (CDR).

Cetuximab is human/murine chimeric IgG-1 directed against the epidermal growth factor receptor (EGFR) overexpressed in advanced-stage EGFR positive colorectal cancer. Cetuximab was approved in the US and EU in 2004 and 2005, respectively. It is composed of two HC and two LC as well for a total of 1326 AA and bears two N-glycosylation sites, both present on the HC. One is located on the Asn-299 and the other is present on the Asn-88 [34] (Figure 1). Note that mAbs' glycosylations are subject to extensive characterization and profiling because it significantly influences antibody - dependent cell - mediated cytotoxicity (ADCC) and complement dependent cytotoxicity (CDC) [35-37]

A tryptic digest of each sample was prepared using an adapted in-solution digestion protocol. The digests were then separated and characterized by t-ITP CESI-MS/MS. The peptide mapping was first performed using a peptide fragment fingerprinting (PFF) strategy similar to conventional bottom-up proteomics. Glycosylations were structurally characterized using MS/MS spectra and a semiquantitative profiling was established. Other posttranslational modifications (PTMs) of interest were investigated to complete the characterization of mab samples. Finally trastuzumab and cetuximab CESI-MS/MS characterization data were confronted to the data obtained for their respective biosimilar candidates in an equivalent perspective as required by regulation agencies to assess biosimilarity. This comparison was performed in order to determine if the developed t-ITP CESI-MS/MS methodology could be applied to demonstrate the biosimilarity between two mAbs in a robust and hastened manner, thereby improving biosimilars development pipeline.

Results and discussion

Amino acid sequences characterization. AA sequence is a key criteria for demonstrating the biosimilarity of a candidate. Indeed, along with other attributes, regulations require that the AA sequence of a candidate is proven to be the same as the reference product in order to be assimilated as biosimilar. Each mAb sample was digested by trypsin using an adapted digestion protocol in order to obtain a peptide mixture that could be characterized by CESI-MS/MS enabling demonstration/invalidation of the biosimilarity of two mabs. Another aspect of the sample preparation is the source of the sample, reference mAbs (trastuzumab and cetuximab) in their commercial formulation were compared to two potential biosimilars (trastuzumab-B and cetuximab-B) in development that were sampled directly from the test bioreactor and purified, meaning that matrixes were significantly different. Sample preparation requires then cancelation of matrix effects, which then renders CESI-MS/MS data completely comparable.

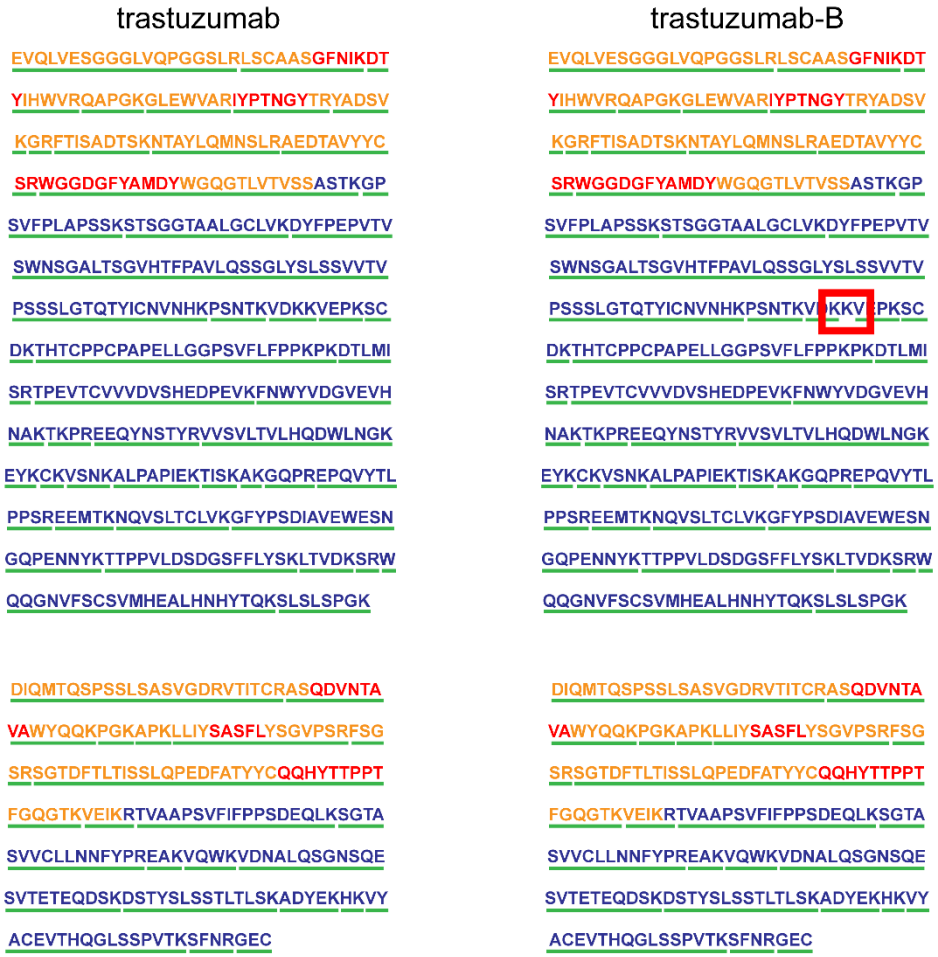

Figure 2. Sequence coverage obtained by CESI-MS/MS for trastuzumab (left-hand side) and trastuzumab-B (right hand side) showing only one amino acid (Lys-217) unidentified in the case of the candidate biosimilar. Experimental conditions: bare fused silica capillary (95cm; 30µm i.d.); sample mAbs tryptic digest 2.22 µM; injection volume 90nL.

For trastuzumab, the analysis of the tryptic digest using the CESI-MS/MS method allowed acquisition, in a single analysis, of 100% sequence coverage on both HC and LC while injecting only 32 ng of digested protein. In a previous work, we demonstrated the capacity to obtain the complete sequence coverage in a single analysis using CESI-MS/MS [23]. It is important to note, however, that significant improvement of the methodology enabled us to obtain the complete AA sequence characterization exclusively through the identification of peptides, without miscleavages nor posttranslational modifications, as emphasized in **Figure 2**. Trastuzumab-B was characterized using the same methodology. As shown in Figure 2, the AA sequence characterization of this biosimilar allowed demonstratation of a complete AA sequence similarity to the reference protein except for HC lysine217 (lys-217). Even by considering tryptic peptides with miscleavages, the AA lys-217 could not be confirmed, suggesting that this particular AA may be substituted in trastuzumab-B compared to the reference mAb. The interpretation of unassigned MS/MS spectra allowed us to demonstrate that the AA composing trastuzumab-B in position 217 was an arginine instead of a lysine, as illustrated by the spectra in **Figure 3**. Several ions pointed out this particular AA substitution reinforcing the confidence of this interpretation (**Figure 3**). From a regulatory standpoint, the substitution of one AA does not with the highly similar paradigm, and thus trastuzumab-B would not be considered biosimilar by either the EMA or FDA.

The same experiment was performed for cetuximab and cetuximab-B head-to-head. Once again, one single injection of cetuximab digest allowed the complete characterization of the AA sequence in this case as well through peptides without modifications nor miscleavages, demonstrating the robustness of the methodology. Cetuximab-B was produced in Chinese hamster ovary cells (CHO) in contrast to the reference product, which was produced in mouse myeloma cells (Sp2/0), but based on the same amino-acid sequence. The comparison of the CESI-MS/MS data, from unique injections of each sample, proved the complete equivalence of the AA sequence between cetuximab and its biosimilar candidate.

The experiment was repeated in three technical replicates and the same results were obtained in a single analysis of each mAb sample, demonstrating the reproducibility in term of characterization of the developed methodology. The CESI-MS/MS interface demonstrates its capacity to provide an impressive sensitivity, as it allows advanced characterization of the AA sequence using only 32 ng of digested mAbs. Even more, the MS/MS data exhibited the capacity to retrieve systematically more than 70% of the y/b fragment ions. In the case of trastuzumab, more than 90% of the y/b ions were obtained, which resulted into the possibility to unravel the totality of the AA organization over the variable domain.

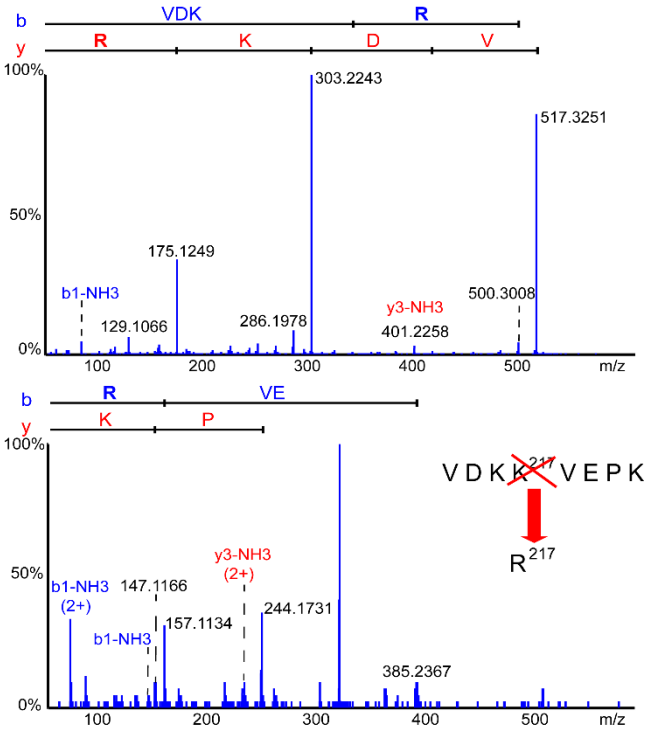

Figure 3. Raw MS/MS, extracted from trastuzumab-B CESI-MS/MS analysis, presenting fragmentation of ions m/z 517.3095 (1+) and m/z 314.6937 (2+) characterizing the amino acid substitution between trastuzumab and its candidate biosimilar. VDKK$_{217}$VEPK for trastuzumab (**G1m17 allotype**) and VDKR$_{217}$VEPK in the case of trastuzumab-B (**G1m3 allotype**).

That trend in MS/MS spectra quality, even on a MS using a TOF analyzer enabling ions accumulation only on a reduced time, is explained by the excellent ionization efficiency when coupling CE to MS through sheathless interfacing. Indeed the low flow rate induced by the electroosmotic flow (< 40 nL/min) and the intrinsic characteristics of the CESI interface contribute to the formation of a nano electrospray (nanoESI). The nanoESI improves analyte ionization and reduces ion suppression effects, therefore impacting positively on MS/MS spectra quality [30]. The capacity to obtain the totality of the y/b ions, meaning by deduction the AA succession order, for the variable domain could be a huge asset in mAbs development as because this part of the protein is directly responsible for antigen epitope recognition, and is therefore crucial for mAbs potency.

Glycosylation structural characterization and semiquantitative profiling. Glycosylation is a PTM that occurs naturally during excretion of antibodies from the expression system to the extracellular medium. It only represents around 3% of the total mass of the protein but is subject to extensive studies due to its significant influence on mAbs ADCC and CDC.

No glycan release was performed during the sample preparation and glycans could be characterized directly on the corresponding digested peptide. In the case of the trastuzumab/trastuzumab-B biosimilarity study, the injection of the reference mAb allowed identification of 16 different glycoforms borne by trastuzumab. Alongside of glycans structural

characterization, the semiquantitative glycoforms profiling was performed for trastuzumab and its biosimilar candidate as shown in Figure 4. To estimate glycoforms distribution, maximum intensities of each charge states of the glycopeptides were considered. The glycoprofiling established for the protein demonstrated the possibility to access lowly abundant glycans which is of interest in comparability and biosimilarity studies. These results suggest the compatibility between the MS dynamic range and the CE loading capacity. Note also that the glycosylation profiling established is consistent to data available in the literature [38].

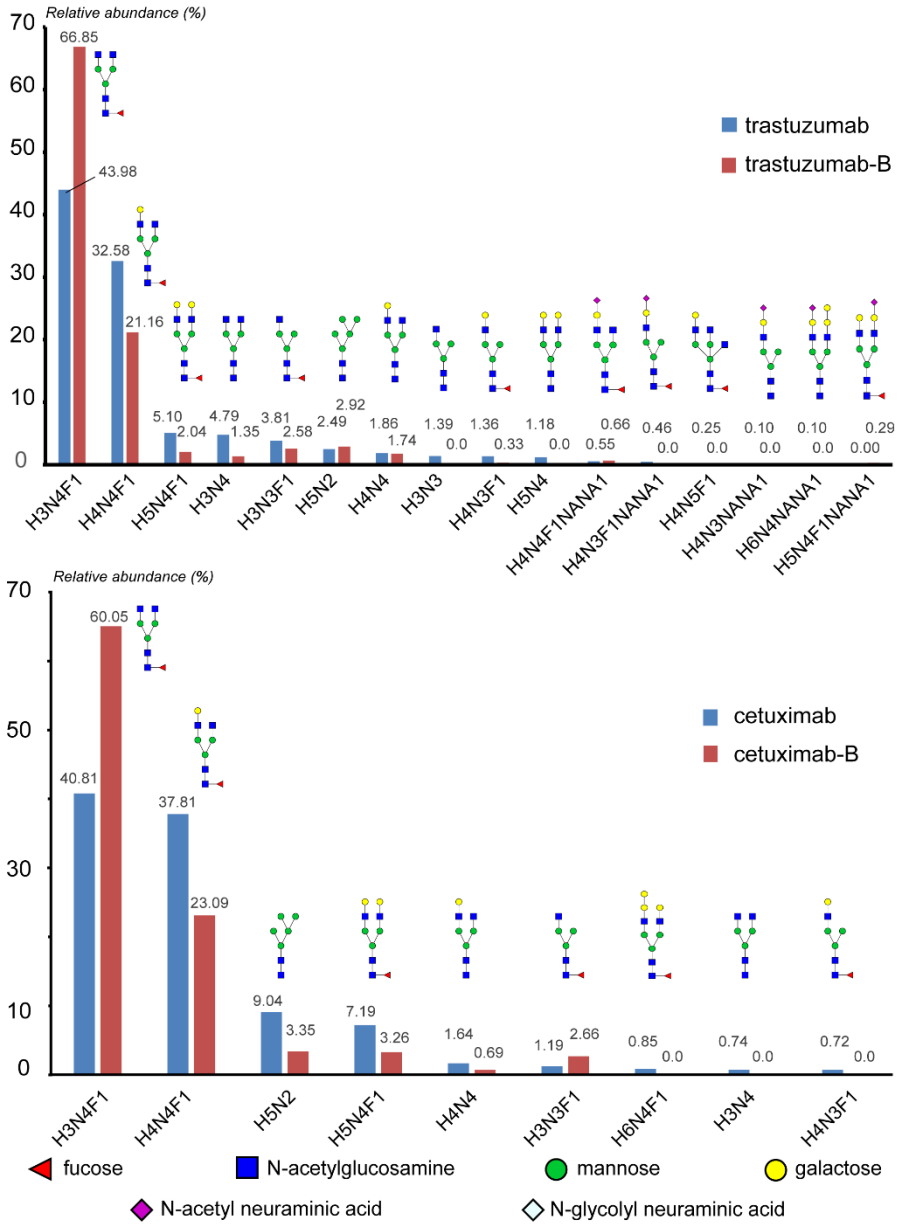

Figure 4. Glycoforms semiquantitative analysis results obtained through the CESI-MS/MS data for the Fc peptide of trastuzumab/trastuzumab-B and cetuximab/cetuximab-B. Glycoform profiling could be compared between original mAbs and their candidate biosimilar.

Even subtle glycoforms distribution differences between trastuzumab and the candidate biosimilar could be distinguished from the results. Trastuzumab-B was also shown to exhibit a rather different glycoforms distribution while carrying no glycans containing sialic acid (Figure 4).

The cetuximab/cetuximab-B biosimilarity study exhibited different particularities regarding glycosylations. Cetuximab bears two N-glycosylation sites on each HC; the first one is located in the Fc/2 domain and common to all IgGs, while the second one in the Fd domain [39] Focusing on the glycopeptide allowed us to distinguish glycans with regard to their glycosylation sites, which cannot be done when using methodologies involving glycans release. In this study as well, glycosylation profiles established demonstrated the identification of a significant number of glycosylation on each sites as shown in Figure 4-B and Figure 5. Significant differences in the glycosylation profile between cetuximab and its biosimilar candidate could be characterized showing the applicability of the method to biosimilarity assessment (figure 4-B). Several cases of hypersensitivity to cetuximab were reported in the literature; the cause of that side-effect was related to the galactose-α-1,3-galactose present on the Fd glycosylation site of the protein [40]. For the marketed version of cetuximab produced in SP2/0 cells, 18 different glycoforms were identified, with 88% capped by at least one alpha-1,3-galactose residue, 23% capped by a N-glycolylneuraminic acid (NGNA) residue and traces of oligomannose. Development of a cetuximab biobetter should take into account that side-effect in order to prevent the occurrence of such glycosylation and provide an improved product, e.g. by production in CHO cells. Fd glycosylation profiles established from CESI-MS/MS data demonstrated that cetuximab-B would not be a suitable biosimilar candidate but a possible biobetter. Thus the glycans identified on the Fd glycosylation site of cetuximab-B were significantly different from those present on the innovator mAb. The profile established for cetuximab-B therefore demonstrated that the abundance of glycoforms containing galactose-α-1,3-galactose could be significantly reduced (Figure 5).

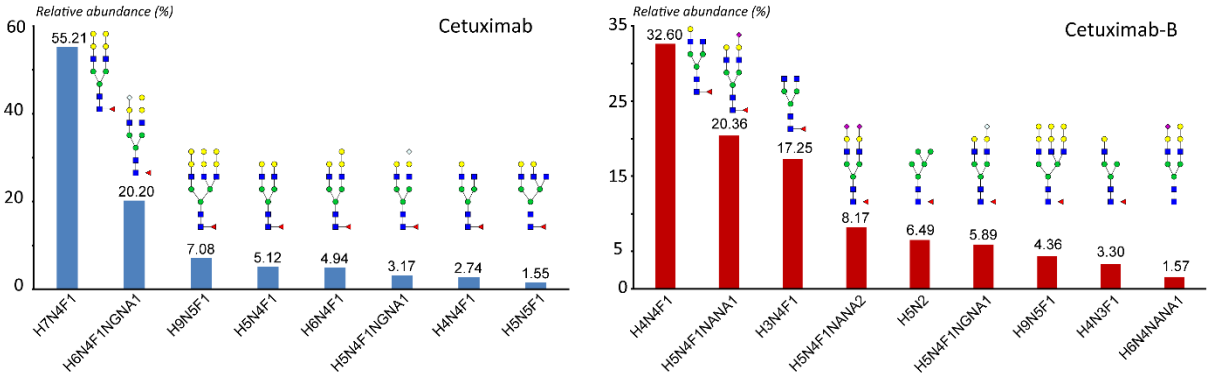

Figure 5. Glycoform semiquantitative analysis results regarding Fd domain glycosylation of cetuximab and cetuximab-B.

Highly sensitive glycosylation semiquantitation is often performed by removing the glycans using specific enzymes [41]. Results proved that the methodology used in this study additionally to reduce sample treatment, lowering the potentiality of artefacts, allows researchers to specifically locate the glycosylation sites and study them independently, with a sensitivity challenging already applied methods. The adopted strategy has allowed, concomitant with AA sequence characterization, to demonstrate/disconfirm the biosimilarity regarding the glycosylation profile between reference mAbs and their candidate biosimilar.

PTMs hot-spots comparison study. Other PTMs besides glycosylation must also be considered in biosimilarity assessment. PTMs referred as hot-spots are described in the literature as being induced by changes of the protein structure. Some of them classified as CQAs may influence the immunogenicity, the PK and PD of the protein. It is therefore necessary to prove that none of those PTMs are over represented in a biosimilar candidate, including N-terminal glutamine/glutamic acid cyclization leading to the formation of N-terminal pyroglutamic acid (pE), asparagine deamidation (deaN), methionine oxidation (oxiM) and aspartic acid isomerization (isoD) on various position depending on the considered mAb [2]. CESI-MS/MS data were also used to characterize the PTM hot-spots of the different mAbs samples prepared. PTMs occurrence levels were estimated from the data and finally compared between a biosimilar candidate and the corresponding originator antibody.

Data showed that every single PTM hot-spot monitored could be successfully characterized (Table 1). The use of electrophoretic separation is particularly pertinent to characterize PTMs. Indeed PTMs induced significant change of electrophoretic mobilities between the modified digested peptides and its unmodified homologs. Therefore, those peptides could be separated by the electrokinetically driven separation of CE from 0.5 min in case of deaN to several minutes for pE. The use of CE as a separating technique is clearly of great interest. As emphasized in Figure 6, it was possible to separate the same peptide differing solely by the conformation of its aspartic acid. The separation in this particular modification allows to obtain additional information regarding the characterization. Indeed the isomerization of aspartic acid does not involve a mass change of the peptide and cannot be identified by MS using a time-of-flight (TOF) analyzer. Moreover the ability to separate, upfront to MS analysis, a peptide having experienced chemical modifications from the same peptide without modification allows sensitivity to be maximized. In the case of co-migration, those peptides would compete against each other during the ionization process.

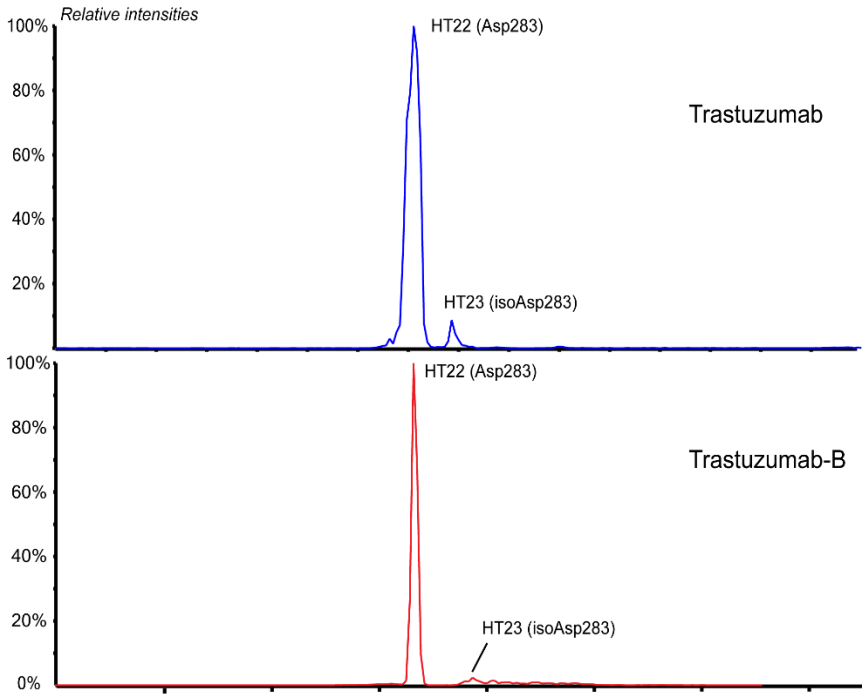

Figure 6. Extracted Ion Electropherogram (EIE) corresponding to m/z (839.408, 2+) for digested peptide HT22 (position 278-291) experiencing aspartic acid isomerization.

Signal is improved by their separation and occurrence levels accuracy theoretically optimal. Specificity of CE separation enables the separation of digested peptides experiencing aspartic acid isomerization; that characteristic is particularly valuable as the characterization of such modification requires particular methods giving access solely to this information,21 in contrast, here the study of aspartic acid isomerization could be incorporated within the framework of the overall primary structure characterization. The excellent MS/MS spectra quality obtained was particularly useful for PTM hot-spots characterization as it was possible to point out precisely the AA affected by the modification. PTM hot-spots semiquantitative measure was performed for each sample and occurrence levels were compared between the reference mAb and its respective candidate biosimilar as emphasized in Table 1 and 2. In this part, relative occurrence levels were estimated from the maximum intensities of the ions corresponding to the modified peptides compared to the abundance of the corresponding unmodified peptide.

In the context of the trastuzumab/trastuzumab-B biosimilarity study, results showed that endogenous deamidation due to proteolytic digestion conditions could be significantly controlled as several deamidation sites exhibited modifications levels below 7% and even 4% for dean-55 (Table 1).

Results demonstrated the capacity of this methodology to distinguish significant variation of modification levels. Therefore results showed that isoD-170 (LC) abundance was superior to 13.4% in trastuzumab while its candidate biosimilar exhibited only 7.1% on this modification (Table 1), a similar trend could also be unraveled with the presented results for isoD-167 (LC).

position	sequence	PTM	Trastuzumab distribution		Trastuzumab-B distribution	
			unmodif	modif	unmodif	modif
1 -19	**E**VQLVESGGGLVQPGGSLR	E^{1} / pE^{1}	98.2	1.8	97.3	2.7
51 - 59	IYPT**N**GYTR	N^{55} / $deaN^{55}$	89.4	10.6	92.3	7.7
99 - 124	WGG**D**GFYAMDYWGQGTLVTVSSASTK	D^{102} / $isoD^{102}$	84.9	15.1	86.0	14.0
252 - 258	DTL**M**ISR	M^{255} / $oxiM^{255}$	95.3	4.7	94.5	5.5
259 - 277	TPEVTCVVVDVSHE**D**PEVK	D^{272} / $isoD^{272}$	94.4	5.6	94.5	5.5
278 - 291	FNWYV**D**GVEVHNAK	D^{283} / $isoD^{283}$	92.0	8.0	96.5	3.5
374 - 395	GFYPSDIAVEWES**N**GQPENNYK	N^{387} / $deaN^{387}$	85.8	14.3	100.0	0.0
420 - 442	WQQGNVFSCSV**M**HEALHNHYTQK	M^{431} / $oxiM^{431}$	97.8	2.2	95.2	4.8
396 - 412	TTPPVLDS**D**GSFFLYSK	D^{404} / $isoD^{404}$	93.0	7.0	93.9	6.1
25 - 42	ASQDV**N**TAVAWYQQKPGK	N^{30} / $deaN^{30}$	96.1	3.9	96.0	4.0
150 - 169	VDNALQSGNSQESVTEQ**D**SK	D^{167} / $isoD^{167}$	64.0	36.0	91.6	8.4
170 - 183	**D**STYSLSSTLTLSK	D^{170} / $isoD^{170}$	86.6	13.4	92,9	7.1

Table 1. Table summarizing the study regarding PTM hotspots occurrence levels for trastuzumab/trastuzumab-B obtained by CESI-MS/MS for each tryptic digest, modification levels (distribution % intact / % modification) were determined through signal intensity for each targeted PTM.

These high levels of modifications were correlated to the prolonged time the sample of reference trastuzumab had been stored. Regarding deaN, results of trastuzumab/trastuzumab-B showed dissimilarities for deaN-387 (HC) only, indeed trastuzumab-B exhibited less than 1% of this modification. OxiM results obtained from this study proved similar level of oxidation on the designated sites (Table1).

Relative occurrence levels of PTM hot-spots were also studied and compared in the case of cetuximab/cetuximab-B. Results obtained were compiled in Table 2. As in the first study, results obtained allowed successful characterization of each PTMs hot-spots present on cetuximab. Results of the semiquantitative analysis regarding PTMs hot-spots occurrence were compared. Notably, both mAbs demonstrated specific characteristics supporting their similarity: N-terminal glutamine (Gln) cyclization was shown to be complete because the intact peptide could not be detected in samples of either mAbs. This trend could also be confirmed on the HC for $isoD^{403}$ and $deaN^{158}$ (Table 2); indeed those modifications exhibited absolutely no occurrence in cetuximab or its candidate biosimilar, illustrating the similarity of those proteins on such specific aspect. Cetuximab, as opposed to cetuximab-B, demonstrated slightly superior deamidation levels in the case of $deaN^{172}$ and $deaN^{386}$ for the HC and $deaN^{41}$ on the LC (Table 2), which were correlated with a longer conservation period for the reference mAb sample. However, these results illustrate the sensitivity of the method, which allows only

minor variations in PTMs occurrence levels to be distinguished. Regarding the biosimilarity assessment of PTM hot-spots, results showed those antibodies could comply to be considered as biosimilars on this particular aspect.

position	sequence	PTM	Cetuximab distribution		Cetuximab-B distribution	
			unmodif	modif	unmodif	modif
1 -5	**Q**VQLK	Q^{1} / pQ^{1}	100.0	0.0	100.0	0.0
150 - 212	DYFPEPVTVSW**N**SGALTSGVHTF PAVLQSSGLYSLSSVVTVPSSSL GTQTYICNVNHKPSNTK	N^{161} / deaN^{161}	25.6	74.4	59.5	40.5
277 - 290	FNWYV**D**GVEVHNAK	D^{282} / isoD^{282}	95.2	4.8	96.3	3.7
373 - 394	GFYPSDIAVEWES**N**GQPENNYK	N^{386} / deaN^{386}	45.6	54.4	95.7	4.3
395 - 411	TTPPVLDS**D**GSFFLYSK	D^{403} / isoD^{403}	100.0	0.0	100.0	0.0
40 - 45	T**N**GSPR	N^{41} / deaN^{41}	87.8	12.2	94.7	5.3
150 - 169	VDNALQSG**N**SQESVTEQDSK	N^{158} / deaN^{158}	96.6	3.4	100.0	0.0

Table 2. Table summarizing the study regarding PTM hotspots occurrence levels for cetuximab/ cetuximab-B. Using the CESI-MS/MS data for each tryptic digest, modification levels (distribution % intact / % modification) were determined through signal intensity for each targeted PTM.

The characterization strategy developed in this study involved the use of CESI-MS/MS analysis. Data obtained from only one injection of each sample tryptic digest allowed characterization of each mAbs simultaneously over several levels defining their primary structure: AA sequence, glycosylation structure/relative abundance and specific PTMs hot-spots. The purpose was to study the biosimilarity between two approved mAbs (trastuzumab and cetuximab) and two respective biosimilar candidates. Biosimilarity study results unraveled that in both cases, biosimilar candidates would hardly be considered as such according to current regulations. In the case of trastuzumab-B, one AA difference compared to the original mAb led to this conclusion, while in the case of cetuximab-B, major glyco-profile variations were highlighted. Among the biosimilarity assessments performed successfully for each mAbs, data collected during the CESI-MS/MS analysis pointed out specifically which attribute of the protein did not comply with the reference mAb giving precious information from an R&D perspective. For example, the difference of one AA for trastuzumab-B strongly supports the need of re-engineering of the vector to produce the right allotype.42 However, in the case of cetuximab-B, each aspects of the primary structure of the mAb demonstrated the biosimilarity except for the glycosylation. Yet recent advances in mAbs glycoengineering permits optimization of the mAbs glycoform expression, thereby suppressing undesired glycosylations [43, 44].

In summary, the CESI-MS/MS strategy developed in the context of this study appeared to be fully compatible with use as an orthogonal technique concomitantly to the now commonly used analytical techniques for characterization of mAbs, biosimilar, biobetters and, more generally, biotherapeutic proteins.

Conclusion

We report here an innovative strategy involving CE coupled to high resolution tandem MS to characterize mAbs simultaneously over different angles of their primary structure, in order to perform biosimilarity assessments between approved mAbs (trastuzumab and cetuximab) and candidate biosimilars. Prior to analysis, samples were digested by trypsin using an adapted in-solution digestion protocol allowing improvement in digestion yield and a standardized final sample content compatible with CE separation. The analysis of each sample allowed establishment of the similarity/dissimilarities concerning AA sequence between several approved mAbs and their respective biosimilar candidates.

The sensitivity and MS/MS spectra quality obtained, mainly supported by the instrumental setting used in this work, enabled AA sequence characterization on an unprecedented level as 100% sequence coverage could be obtained in a single analysis through identification of only tryptic peptides without miscleavages or PTMs. Furthermore more than 70% of the y/b ions could be systematically retrieved, providing detailed structural data over this aspect of the protein. Along with this characterization, structural and semiquantitative analysis of the glycoforms could be performed, which demonstrated the capacity of the method to locate precisely different glycosylations sites. The glycoforms profiling obtained demonstrated the capacity of the established methodology to detect quite low abundance glycans, enabling the detection of potentially undesired glycoforms. Likewise the profiling accuracy was also highlighted as relative abundances obtained could be correlated to other methods. CESI-MS/MS data enabled subtle differences in glycoforms distribution between reference mAbs and their candidate biosimilars to be distinguished, proving that the dynamic range achievable with CE-ESI-MS coupling is fully compatible with this type of analytical problem. Finally, specific PTMs hot-spots were characterized using the very same CESI-MS/MS data used to characterize mAbs samples previously; results demonstrated the capacity of the method to characterize precisely the targeted sites.

The sensitivity provided by the CESI-MS system allowed, in most cases, the precise identification of the AA affected by the PTMs even in the circumstances of weakly abundant modifications. CE separation specificity proved valuable because peptides with chemical modification, as weak as aspartic acid isomerization, could be separated, enabling their

concomitant characterization. The results illustrated the capacity of the method to observe differences in modification levels between the reference mAb samples and their respective candidate biosimilars studied. Data obtained from CESI-MS/MS analysis established dissimilarities that existed amongst the candidate biosimilars compared to the approved mAbs, demonstrating the ability of the developed methodology to address such analytical issues in a structured and comprehensive way while easing sample treatment and reducing the number of experiment required. Furthermore, results permitted identification of precisely which aspects of the candidate antibody did not comply with their assignment as biosimilars providing important structural information that could be used to strategically plan subsequent R&D work. This study illustrates without ambiguity the possibility of using CE coupled to MS as an orthogonal technique, along with techniques routinely applied, to address complex analytical issues such as mAbs biosimilarity assessment, and points out more generally the capacity of this type of coupling to characterize biotherapeutic proteins.

Materials and methods

Materials. Chemicals used were of analytical grade or high purity grade and purchased from Sigma-Aldrich (Saint Louis, MO, USA). Water used to prepare buffers and sample solutions was obtained using an ELGA purelab UHQ PS water purification system (Bucks, UK). Trastuzumab and cetuximab samples are the EMA approved products and formulation purchased from Roche (Penzberg, GE) and Merck KGaA (Darmstadt, GE) respectively. The biosimilar versions were produced at the Centre d'Immunologie Pierre Fabre (Saint Julien en Genevois, FR) for analytical methods development. Marketed mAbs were stored at 4°C for one year before characterization while biosimilar versions was analyzed soon after their reception (approximately 2 months), otherwise all samples were stored at 4°C. RapiGest SF surfactant was purchased from Waters (Milford, MA, USA).

Sample preparation. A volume corresponding to 100 µg of protein were used. Samples were first diluted using milliQ water to a final concentration of 6.7 µg/µL. Samples were then diluted using 0.1% RapiGest surfactant to reach a final concentration of 3.35 µg/µL and heated to 40°C during 10 min. Dithiothreitol (DTT) was added to the sample to a final concentration of 25 mM. Samples were then heated for a 5 min incubation at 95°C. After being cooled to room temperature (RT), iodoacetamide (IDA) was added to the sample to a final concentration of 10 mM and samples were kept in the dark for 20 min to allow alkylation of cysteines (Cys). A volume of 1 µL of trypsin (0.5µg/µL) was added to the sample which was left at room temperature for 3h and another volume of 1µL was added after this time. Digestion was performed overnight at 37°C. After digestion completion, 1% (v/v) formic acid (FA) was added

to the samples, which were left at RT in order to cleave the surfactant. Samples were finally diluted to a final concentration of 2.2 µM of protein using ammonium acetate 50 mM (pH 4.0).

Capillary electrophoresis. The CE separations were performed with a PA 800 plus capillary electrophoresis system from Beckman Coulter equipped with a temperature controlled autosampler and a power supply able to deliver up to 30 kV. Hyphenation was carried out using a CESI prototype made available by Sciex separation (Brea, CA, USA). Prototype of bare fused-silica capillaries (total length 100 cm; 30 µm i.d.) with a characteristic porous tip on its final 3 cm supplied by Sciex separation, a second capillary (total length 80 cm; 50 µm i.d.) filled during experiments with BGE allows electric contact. New capillaries were flushed at 75 psi (5.17 bar) for 10 min with methanol, 10 min with 0.1 M sodium hydroxide, then 10 min with 0.1 M hydrochloric acid and water for 20 min. Finally, the capillary was flushed 10 min at 75 psi with 10% acetic acid, which is the BGE used for the separation.

Mass spectrometry. For antibody characterization, the CESI system was hyphenized to a 5600 TripleTOF mass spectrometer (ABSciex, Darmstadt, Germany). The 5600 MS is equipped with a hybrid analyzer composed of quadrupoles followed by a time-of-flight (TOF) analyzer. ESI source parameters were set as follows: ESI voltage -1.75kV, gas supplies (GS1 and GS2) were deactivated, source heating temperature 150°C and curtain gas value 5. Experiments were performed in Top15 information dependent acquisition (IDA), accumulation time was 250 msec for MS scans and 100 msec for MS/MS scans leading to a total duty cycle of 1.75 sec. Mass/charge (m/z) range was set to 100-2000 in MS and 50-2000 in MS/MS. Using those parameters, the mean resolution provided by the instrument is 40000 in MS (m/z 485.251) and 25000 in MS/MS (m/z 345.235).

MS/MS data analysis. Data obtained from the CESI-MS/MS experiments were analyzed using Peakview software (ABSciex, Darmstadt, Germany). Tryptic peptides (without miscleavages or PTMs except cys carbamidomethylation) were determined theoretically from mabs' amino acid sequences available through literature.23, 45 Additional peptides were identified using Mascot search engine provided by Matrix science; tryptic cleavage rules were applied. Carbamidomethylation of cysteine (+57.02 Da), N-deamidation of aspartic/isoaspartic acid (+0.985 Da) or succinimide intermediate (-17.03 Da), methionine oxidation (+15.99 Da) and N-terminal glutamic acid cyclization (-17.02 Da) were selected as variable modifications. The mass tolerance for precursor ions was set to ± 5 ppm and ±0.05 Da for fragmentation ions.

Acknowledgments

Authors would like to thank Sciex separations Inc. for lending a CESI prototype, Dr. M. Anselme and Dr. Stephen Lock from AB Sciex Inc. for their support. The authors would like also to express their gratitude to Dr. E. Wagner-Rousset, MC. Janin-Bussat and O. Colas

(Centre d'immunologie Pierre Fabre, St Julien en Genevois, France) for helpful discussions around antibody structural characterization.

References

[1] **Beck A, Carter PJ, Gerber H-P, Lugovskoy AA, Wurch T, Junutula JR, Kontermann R, Mabry R, Ghayur T**. 8th Annual European Antibody Congress 2012: November 27–28, 2012, Geneva, Switzerland. mAbs 2013; 5:339-57.

[2] **Beck A, Wagner-Rousset E, Ayoub D, Van Dorsselaer A, Sanglier-Cianferani S**. Characterization of Therapeutic Antibodies and Related Products. Analytical Chemistry 2013; 85:715-36.

[3] **Beck A, Wurch T, Bailly C, Corvaia N**. Strategies and challenges for the next generation of therapeutic antibodies. Nat Rev Immunol 2010; 10:345-52.

[4] **Reichert JM**. Antibodies to watch in 2013: Mid-year update. mAbs 2013; 5:513-7.

[5] **Kotsovilis S, Andreakos E**. Therapeutic Human Monoclonal Antibodies in Inflammatory Diseases. In: Steinitz M, ed. Human Monoclonal Antibodies: Humana Press, 2014:37-59.

[6] **Waldmann TA**. Immunotherapy: past, present and future. Nat Med 2003; 9:269-77.

[7] **Agency EM**. Guideline on Similar Biological Medicinal Products. 22 May 2013; CHMP/437/04 Rev 1.

[8] **Beck A, Reichert JM**. Approval of the first biosimilar antibodies in Europe: A major landmark for the biopharmaceutical industry. mAbs 2013; 5:621-3.

[9] **Beck A, Diemer H, Ayoub D, Debaene F, Wagner-Rousset E, Carapito C, Van Dorsselaer A, Sanglier-Cianferani S**. Analytical characterization of biosimilar antibodies and Fc-fusion proteins. Trac-Trends in Analytical Chemistry 2013; 48:81-95.

[10] **Fornelli L, Ayoub D, Aizikov K, Beck A, Tsybin YO**. Middle-Down Analysis of Monoclonal Antibodies with Electron Transfer Dissociation Orbitrap Fourier Transform Mass Spectrometry. Analytical Chemistry 2014; 86:3005-12.

[11] **Rosati S, Thompson NJ, Barendregt A, Hendriks LJA, Bakker ABH, de Kruif J, Throsby M, van Duijn E, Heck AJR**. Qualitative and Semiquantitative Analysis of Composite Mixtures of Antibodies by Native Mass Spectrometry. Analytical Chemistry 2012; 84:7227-32.

[12] **Beck A, Sanglier-Cianférani S, Van Dorsselaer A**. Biosimilar, Biobetter, and Next Generation Antibody Characterization by Mass Spectrometry. Analytical Chemistry 2012; 84:4637-46.

[13] **Ambrogelly A, Liu Y-H, Li H, Mengisen S, Yao B, Xu W, Cannon-Carlson S**. Characterization of antibody variants during process development: The tale of incomplete processing of N-terminal secretion peptide. mAbs 2012; 4:701-9.

[14] **European Medicines Agency (EMA)**. Guideline on development, production, characterization and specifications for monoclonal antibodies and related products. 2009:EMEA/CHMP/BWP/157653/2007.

[15] **Damen CN, Chen W, Chakraborty A, Oosterhout M, Mazzeo J, Gebler J, Schellens JM, Rosing H, Beijnen J**. Electrospray ionization quadrupole ion-mobility time-of-flight mass spectrometry as a tool to distinguish the lot-to-lot heterogeneity in N-glycosylation profile of the therapeutic monoclonal antibody trastuzumab. J Am Soc Mass Spectrom 2009; 20:2021-33.

[16] **Rustandi RR, Wang Y**. Use of CE-SDS gel for characterization of monoclonal antibody hinge region clipping due to copper and high pH stress. Electrophoresis 2011; 32:3078-84.

[17] **Michels DA, Tu AW, McElroy W, Voehringer D, Salas-Solano O**. Charge Heterogeneity of Monoclonal Antibodies by Multiplexed Imaged Capillary Isoelectric Focusing Immunoassay with Chemiluminescence Detection. Analytical Chemistry 2012; 84:5380-6.

[18] **Wollacott RB, Casaz PL, Morin TJ, Zhu HL, Anderson RS, Babcock GJ, Que J, Thomas WD, Ozturk SS**. Analytical characterization of a monoclonal antibody therapeutic reveals a three-light chain species that is efficiently removed using hydrophobic interaction chromatography. Mabs 2013; 5:925-35.

[19] **Heudi O, Barteau S, Zimmer D, Schmidt J, Bill K, Lehmann N, Bauer C, Kretz O**. Towards absolute quantification of therapeutic monoclonal antibody in serum by LC-MS/MS using isotope-labeled antibody standard and protein cleavage isotope dilution mass spectrometry. Analytical Chemistry 2008; 80:4200-7.

[20] **Beck A, Bussat MC, Zorn N, Robillard V, Klinguer-Hamour C, Chenu S, Goetsch L, Corvaia N, Van Dorsselaer A, Haeuw JF**. Characterization by liquid chromatography combined with mass spectrometry of monoclonal anti-IGF-1 receptor antibodies produced in CHO and NSO cells. Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences 2005; 819:203-18.

[21] **Vlasak J, Bussat MC, Wang S, Wagner-Rousset E, Schaefer M, Klinguer-Hamour C, Kirchmeier M, Corvaïa N, Ionescu R, Beck A**. Identification and characterization of asparagine deamidation in the light chain CDR1 of a humanized IgG1 antibody. Analytical Biochemistry 2009; 392:145-54.

[22] **Fekete S, Beck A, Veuthey J-L, Guillarme D**. Theory and practice of size exclusion chromatography for the analysis of protein aggregates. Journal of Pharmaceutical and Biomedical Analysis.

[23] **Gahoual R, Burr A, Busnel J-M, Kuhn L, Hammann P, Beck A, François Y-N, Leize-Wagner E**. Rapid and multi-level characterization of trastuzumab using sheathless capillary electrophoresis-tandem mass spectrometry. mAbs 2013; 5:479-90.

[24] **Moini M**. Simplifying CE−MS Operation. 2. Interfacing Low-Flow Separation Techniques to Mass Spectrometry Using a Porous Tip. Analytical Chemistry 2007; 79:4241-6.

[25] **Haselberg R, Ratnayake CK, de Jong GJ, Somsen GW**. Performance of a sheathless porous tip sprayer for capillary electrophoresis-electrospray ionization-mass spectrometry of intact proteins. Journal of Chromatography A 2010; 1217:7605-11.

[26] **Busnel J-M, Schoenmaker B, Ramautar R, Carrasco-Pancorbo A, Ratnayake C, Feitelson JS, Chapman JD, Deelder AM, Mayboroda OA**. High Capacity Capillary Electrophoresis-Electrospray Ionization Mass Spectrometry: Coupling a Porous Sheathless Interface with Transient-Isotachophoresis. Analytical Chemistry 2010; 82:9476-83.

[27] **Sarg B, Faserl K, Kremser L, Halfinger B, Sebastiano R, Lindner HH**. Comparing and Combining Capillary Electrophoresis Electrospray Ionization Mass Spectrometry and Nano–Liquid Chromatography Electrospray Ionization Mass Spectrometry for the Characterization of Post-translationally Modified Histones. Molecular & Cellular Proteomics 2013; 12:2640-56.

[28] **Ramautar R, Busnel JM, Deelder AM, Mayboroda OA**. Enhancing the Coverage of the Urinary Metabolome by Sheathless Capillary Electrophoresis-Mass Spectrometry. Analytical Chemistry 2012; 84:885-92.

[29] **Bonvin G, Veuthey JL, Rudaz S, Schappler J**. Evaluation of a sheathless nanospray interface based on a porous tip sprayer for CE-ESI-MS coupling. Electrophoresis 2012; 33:552-62.

[30] **Gahoual R, Busnel J-M, Wolff P, François Y, Leize-Wagner E**. Novel sheathless CE-MS interface as an original and powerful infusion platform for nanoESI study: from intact proteins to high molecular mass noncovalent complexes. Anal Bioanal Chem 2014; 406:1029-38.

[31] **Larsson M, Lutz ESM**. Transient isotachophoresis for sensitivity enhancement in capillary electrophoresis-mass spectrometry for peptide analysis. Electrophoresis 2000; 21:2859-65.

[32] **Li S, Schmitz KR, Jeffrey PD, Wiltzius JJW, Kussie P, Ferguson KM**. Structural basis for inhibition of the epidermal growth factor receptor by cetuximab. Cancer Cell 2005; 7:301-11.

[33] **Gunturu KSW, Yanghee; Beaubier, Nike; Remotti, Helen E.; Saif, Wasif M**. Gastric cancer and trastuzumab: first biologic therapy in gastric cancer. Therapeutic Advances in Medical Oncology 2013; 5:143-51.

[34] **Dubois M, Fenaille F, Clement G, Lechmann M, Tabet J-C, Ezan E, Becher F**. Immunopurification and Mass Spectrometric Quantification of the Active Form of a Chimeric Therapeutic Antibody in Human Serum. Analytical Chemistry 2008; 80:1737-45.

[35] **Dalziel M, Crispin M, Scanlan CN, Zitzmann N, Dwek RA**. Emerging principles for the therapeutic exploitation of glycosylation. Science 2014; 343.

[36] **Krapp S, Mimura Y, Jefferis R, Huber R, Sondermann P**. Structural Analysis of Human IgG-Fc Glycoforms Reveals a Correlation Between Glycosylation and Structural Integrity. Journal of Molecular Biology 2003; 325:979-89.

[37] **Jefferis R**. Recombinant antibody therapeutics: the impact of glycosylation on mechanisms of action. Trends in Pharmacological Sciences 2009; 30:356-62.

[38] **Stadlmann J, Pabst M, Kolarich D, Kunert R, Altmann F**. Analysis of immunoglobulin glycosylation by LC-ESI-MS of glycopeptides and oligosaccharides. PROTEOMICS 2008; 8:2858-71.

[39] **Qian J, Liu T, Yang L, Daus A, Crowley R, Zhou Q**. Structural characterization of N-linked oligosaccharides on monoclonal antibody cetuximab by the combination of orthogonal matrix-assisted laser desorption/ionization hybrid quadrupole–quadrupole time-of-flight tandem mass spectrometry and sequential enzymatic digestion. Analytical Biochemistry 2007; 364:8-18.

[40] **Chung CH, Mirakhur B, Chan E, Le Q-T, Berlin J, Morse M, Murphy BA, Satinover SM, Hosen J, Mauro D, et al**. Cetuximab-Induced Anaphylaxis and IgE Specific for Galactose-α-1,3-Galactose. New England Journal of Medicine 2008; 358:1109-17.

[41] **Wagner-Rousset E, Bednarczyk A, Bussat M-C, Colas O, Corvaïa N, Schaeffer C, Van Dorsselaer A, Beck A**. The way forward, enhanced characterization of therapeutic antibody glycosylation: Comparison of three level mass spectrometry-based strategies. Journal of Chromatography B 2008; 872:23-37.

[42] **Jefferis R, Lefranc M-P**. Human immunoglobulin allotypes: Possible implications for immunogenicity. mAbs 2009; 1:332-8.

[43] **Beck A, Cochet O, Wurch T**. GlycoFi's technology to control the glycosylation of recombinant therapeutic proteins. Expert Opinion on Drug Discovery 2010; 5:95-111.

[44] **Beck A, Reichert JM**. Marketing approval of mogamulizumab A triumph for glyco-engineering. Mabs 2012; 4:419-25.

[45] **Ayoub D, Jabs W, Resemann A, Evers W, Evans C, Main L, Baessmann C, Wagner-Rousset E, Suckau D, Beck A**. Correct primary structure assessment and extensive glyco-profiling of cetuximab by a combination of intact, middle-up, middle-down and bottom-up ESI and MALDI mass spectrometry techniques. mAbs 2013; 5:699-710.

3. Conclusion

Les travaux présentés dans cette partie détaillent l'application de la méthode de caractérisation par t-ITP CESI-MS/MS pour l'évaluation la biosimilarité entre deux échantillons de mAbs commercialisés et leur candidat biosimilaire respectif. L'objectif de ces travaux était de démontrer la possibilité d'utiliser le couplage entre la CE et la MS pour réaliser une caractérisation avancée des différents échantillons de mAbs et de déceler les éventuelles différences de structures entre un mAb de référence et son candidat biosimilaire. A partir des données générées en une seule injection, les mAbs ont été caractérisés simultanément sur plusieurs aspects composant leur structure primaire : séquence d'acides aminés, glycosylations, PTMs.

Les résultats obtenus concernant la caractérisation de la séquence montrent que, pour chacun des mAbs étudiés, la méthodologie analytique employée permet un recouvrement de séquence de 100% uniquement par l'intermédiaire d'identifications de peptides sans *miscleavages* ni PTMs. La comparaison des données entre trastuzumab et son candidat biosimilaire a mis en évidence une différence d'un unique acide aminé entre les séquences de ces deux protéines. De plus l'excellente qualité spectrale délivrée par le système CESI-MS a permis d'identifier sans ambiguïté que le candidat était composé d'une arginine en position 217 (HC) tandis que trastuzumab possède une lysine pour cette position. Cette différence est rédhibitoire pour le candidat ne pouvant être considéré comme un biosimilaire de trastuzumab conformément aux directives récemment publiées par l'EMA. Dans le cas de l'étude entre cetuximab et son candidat biosimilaire, les données issues de l'analyse CESI-MS/MS ont permis de conclure à une complète similitude entre les deux mAbs concernant cet aspect.

Les glycosylations ont également été caractérisées en utilisant le même set de données. Les données de CESI-MS/MS permettant d'identifier la structure de la glycosylation par l'intermédiaire des spectres de fragmentations, mais aussi d'estimer l'abondance relative de chaque glycanne. Les profils de glycosylation des mAbs originaux ont pu être confrontés à leur candidat biosimilaire respectif.

Les comparaisons de ces profils ont mis en lumière dans le cas de trastuzumab des différences mineures d'abondances relatives même si les glycosylations exprimées se sont avérées communes aux deux mAbs. Dans le cadre de l'étude de la biosimilarité entre cetuximab et son candidat biosimilaire, la méthodologie développée a rendu possible le traitement de chaque site de N-glycosylation indépendamment. Cette caractéristique est particulièrement intéressante pour les mAbs qui comme cetuximab possèdent plusieurs sites de glycosylations. Dans le cas du site présent dans le domaine Fc/2, les données ont montré des différences dans la distribution des glycannes entre les deux protéines. En ce qui concerne

la glycosylation présente dans le domaine Fd du mAb, les données MS/MS ont permis de conclure qu'une partie des glycosylations exprimées n'étaient pas communes aux deux mAbs. Ainsi le profil de glycosylation établi pour le candidat biosimilaire montre que les glycosylations terminée par le motif galactose-α-1,3-galactose sont bien moins abondantes par rapport au profil de cetuximab. Ce type de glycosylation est responsable d'effets secondaires d'hypersensibilité chez certains patients.

Compte tenu de ces résultats, le candidat biosimilaire de cetuximab ne peut être considéré comme étant un biosimilaire de celui-ci. Pour autant les glycosylations exprimées par le candidat indiquent un risque de réaction dû à une hypersensibilité réduit. Ce candidat pourrait donc potentiellement être considéré comme un *biobetter* de cetuximab. On parle de *biobetter* lorsque par exemple le processus de production du mAbs a été amélioré afin de minimiser l'expression d'une des formes du mAbs non satisfaisante ou alors favoriser l'expression de formes dont l'activité est importante. De nombreux travaux de recherches illustrent cette volonté notamment afin d'obtenir un profil de glycosylation optimale (*glyco-engineering*). L'autorisation de mise sur le marché de *biobetter* est également prévue par les directives de l'EMA.

La dernière partie de l'étude de la biosimilarité entre ces différents échantillons concernent les PTMs *hot-spots*. En effet ces modifications influent sur l'immunogénicité du mAb, il convient par conséquent de vérifier que les candidats biosimilaires ne voient pas ces PTMs surexprimées. De la même manière que pour la caractérisation des glycosylations, les PTMs ont pu être identifiées à partir des spectres MS/MS et le degré de modification de l'échantillon au regard de chaque PTM a pu être évalué. Les résultats obtenus ont avant tout permis de caractériser l'ensemble des PTMs d'intérêt pour chacun des échantillons caractérisés par CESI-MS/MS. Notamment la méthodologie utilisant le couplage entre la CE et la MS par l'intermédiaire du système CESI-MS a offert la possibilité de mettre en pratique l'identification d'isomérisation d'acide aspartique ; ainsi chacun des sites potentiels d'isomérisation a montré une isomérisation partielle. Dans le cadre de l'étude de la biosimilarité entre trastuzumab et son biosimilaire, cela démontre que la méthode CESI-MS/MS peut distinguer de faibles variations du niveau de modifications. La possibilité de distinguer ces différences mineures est possible car la CE permet d'obtenir la séparation de ces peptides avant leur transfert dans la source du spectromètre ce qui offre une sensibilité optimale. Les niveaux de modification entre ces deux mAbs se sont avérés relativement proche et n'ont pas suggérer une quelconque sur-modification du candidat biosimilaire.

La même comparaison inscrite dans l'étude entre cetuximab et son candidat biosimilaire ont permis également de mettre en évidence des variations faibles du niveau de modifications. La

comparaison entre les mAbs n'a pas montré de taux de modifications anormalement élevés en ce qui concerne le candidat biosimilaire ce qui a permis de conclure que les deux mAbs pouvaient être considéré comme similaire du point de vue des PTMs.

Les travaux présentés dans cette partie illustrent le niveau de caractérisation pouvant être atteint grâce à la méthodologie t-ITP CESI-MS/MS mise en place. En effet une seule analyse de chaque échantillon, correspondant à une injection de 200 fmol de peptides digérés, a permis d'effectuer la caractérisation complète de la structure primaire des différents anticorps mais également de conclure sur leur biosimilarité au niveau de leur structure primaire. Les données CESI-MS/MS ont permis dans chaque cas, d'identifier les différences entre un mAbs référence et son candidat biosimilaire avec une grande finesse délivrée par le couplage entre la CE et la MS. Les données recueillies ont mis en avant l'aspect du candidat qui n'a pu donner satisfaction dans l'étude de la biosimilarité avec le mAb référence. Ce type d'information est particulièrement intéressant puisqu'il permet d'orienter les activités de R&D devant être mené pour obtenir l'autorisation du candidat ou dans le pire des cas abandonner rapidement un candidat jugé non prometteur, afin de maîtriser les coûts de développement.

La mise en œuvre du couplage entre la CE et la MS a montré au travers de ces travaux son intérêt dans ce type de problématique. La sensibilité délivrée par le couplage CESI-MS et son important rendement d'ionisation génère une excellente qualité spectrale notamment en MS/MS, ce qui se traduit concrètement dans la possibilité d'obtenir pratiquement la totalité des ions y/b pour une molécule aussi complexe qu'un mAb. De plus comme cela est évoqué dans les parties précédentes, la séparation de peptides par CE entraine la migration de tous les peptides indépendamment de leur structure et évitent la « perte » de certains peptides. Ces deux caractéristiques contribuent à la finesse de la caractérisation et surtout ont permis d'identifier et de caractériser la substitution d'un seule acide aminé. Outre la validation par une étude pratique de cet aspect évoqué à la fin de la partie III, les performances de la méthode ouvrent des perspectives extrêmement intéressantes notamment pour des approches d'analyse protéomique n'utilisant pas de banque de données de protéines pour l'identification, la caractérisation d'erreur de transcription lors de la production de protéines ou encore l'utilisation en tant que méthode orthogonale aux techniques déjà utilisées pour la caractérisation fine des structures de protéines.

L'utilisation du couplage CE-ESI-MS/MS a enfin montré son intérêt pour la caractérisation des glycosylations et autres PTMs. Ces problématiques bénéficient également de la capacité d'ionisation importante du système CESI-MS. Les mécanismes de séparation mises en jeu lors de l'électrophorèse se sont avérés particulièrement intéressant pour séparer des peptides modifiés en raison de l'influence de ces modifications sur la mobilité électrophorétique du

peptide. Il a été ainsi possible de séparer des glycopeptides ayant pour unique différence un galactose. Il en va de même pour des modifications ne donnant pas lieu en RP-LC à la séparation des peptides modifiés de leurs homologues restés intactes tel que la déamidation. L'exemple de l'isomérisation de l'acide aspartique est d'autant plus flagrant que la capacité à séparer ces peptides en CE donne des informations supplémentaires venant augmenter la profondeur de la caractérisation réalisée. La capacité à séparer ces peptides permet de minimiser les effets de suppression d'ions durant l'ionisation, la conséquence directe est une quantification relative dont la justesse et la sensibilité sont intéressantes comme cela a pu être montré dans ces travaux.

La profondeur de la caractérisation pouvant être obtenue par l'intermédiaire de la méthodologie t-ITP CESI-MS/MS développée dans ces travaux de thèse ne s'applique pas uniquement aux mAbs et d'un point de vue théorique les mêmes caractéristiques peuvent permettre d'améliorer le degré de caractérisation atteint dans d'autre problématique tel que la protéomique quantitative ou encore la métabolomique.

PARTIE III : Caractérisation d'anticorps monoclonaux par approche middle-up à l'aide du couplage indirect CE-MS

I - Séparation et caractérisation de glycoformes de Cetuximab par Off-line CZE-UV/ESI-MS

1. Introduction

Depuis 1986 et l'approbation du murumonab-CD3 par l'administration de régulation des aliments et des médicaments américaine (FDA), les anticorps monoclonaux (mAbs) représentent la classe de molécule dont le développement est en constante croissance dans l'industrie pharmaceutique. Les mAbs sont des glycoprotéines complexes potentiellement porteuses de nombreuses micro-hétérogénéités moléculaires dont les sources sont variées (procédés de production inadaptée, conservation prolongée). La complexité de ces protéines et leur utilisation en tant que traitement thérapeutique justifient la nécessité d'améliorer les méthodologies analytiques de caractérisation. Ceci de manière à obtenir une caractérisation rapide et précise de ce type de molécules. La spectrométrie de masse (MS) a su rapidement s'imposer comme une méthode de choix pour la caractérisation des mAbs, notamment en raison de son excellente spécificité, sensibilité mais aussi car elle permet d'obtenir des informations structurales dans certain cas. La MS peut ainsi être utilisée dans le développement de nouveaux mAbs ou encore pour suivre différents lots d'un mAb afin de distinguer des différences mineures. L'enjeu dans un tel contexte est de fournir rapidement des informations structurales permettant en amont d'orienter les activités de R&D (optimisation du procédé de production, élimination d'un candidat non satisfaisant) et de production (libération de lots). Les méthodes analytiques utilisées doivent ainsi être capables de couvrir un large champ de la structure des mAbs : séquence d'acides aminés, glycosylation, modifications post-traductionnelles (PTMs), ponts disulfures. Concernant plus particulièrement les glycosylations qui sont connue pour être une classe de modification importante pouvant impacter les propriétés immunogéniques des mAbs, leur caractérisation est réalisée suivant différentes stratégies tel que la mesure de masse intacte, l'analyse middle-up des différentes domaines des mAbs ou l'analyse protéomique. Cependant, aujourd'hui, aucune séparation de glycoformes de mAbs n'a encore été décrite à l'aide du couplage de référence LC-MS. Comme il a été détaillée dans la partie précédente (Partie II : Caractérisation de la structure primaire d'anticorps monoclonaux et évaluation de biosimilarité à l'aide du couplage CESI-MS), l'électrophorèse de zone (CZE) a montré un réel intérêt pour la caractérisation de mAbs. De plus, cette méthode apparait comme une méthode puissante pour la séparation de protéines entières. Cette technique mise en place avec une détection optique (UV, Fluorescence) est méthode de référence dans les laboratoires de R&D et de qualité-contrôle des industries biopharmaceutiques. Cependant, due à la présence de sel en forte concentration et à la non-

volatilité des électrolytes support (BGE) nécessaire pour séparer des variants de charges de mAbs, le développement du couplage de la CZE avec la MS est très limité. En effet, les conditions décrites dans la littérature permettant la séparation de mAbs intacts en CZE-UV mettent en jeu une forte concentration d'acide ε-aminocaproïque ainsi que certains additifs (Triethylenetetramine, Tween 20). Malgré la qualité des séparations obtenues, ces BGE ne sont en aucun cas compatible avec une détection directe en ESI-MS.

Faisant suite à l'étude précédemment décrite sur le développement du couplage CZE-UV/MALDI-MS (III - Développement et évaluation du couplage CE-UV/MALDI-MS pour l'analyse de protéines intactes et d'applications protéomiques) et dans le but de pallier au manque de résolution du MALDI-MS, nous avons développé une stratégie pour réaliser un couplage CZE-UV/ESI-MS basée sur l'utilisation de notre robot CZE-UV/collection de fraction suivi d'une infusion des différentes fractions par CESI-MS utilisé en tant que plateforme nanoESI. Dans le but de minimiser les phénomènes d'adsorption de protéines à la surface interne du capillaire et de réduire le flux électroosmotique (EOF), un greffage hydroxypropylcellulose (HPC) a été réalisé. L'un des objectifs de cette étude a été de rendre le BGE compatible avec la détection ESI-MS. Pour cela, des conditions asymétriques mettant en jeu différents BGE aux extrémités du capillaire ont été développés. Le BGE à l'entrée du capillaire composé de 200 mM d'acide ε-aminocaproïque et de 25 mM d'acétate d'ammonium pH 5,70, permet de conserver les performances de la séparation électrophorétique alors que le BGE à la sortie du capillaire composé de 25 mM d'acétate d'ammonium pH 5,70, permet une bonne compatibilité avec la détection ESI-MS. Le mAb étudié est le cetuximab qui est un IgG1 chimérique de type humain/souris principalement utilisé dans le traitement du cancer colorectal. Le cetuximab a pour particularité de contenir deux sites de glycosylation réparties sur la chaine lourde (HC) : un localisé sur le domaine Fc/2 (Asn^{299}) et le second localisé sur le domaine F(ab')2 (Asn^{88}). Cela représente un échantillon de choix pour notre étude qui permettra de caractériser ces deux domaines séparément à l'aide d'une stratégie middle-up par digestion IdeS sans réduction. Les performances de notre stratégie seront évaluées sur la caractérisation et la séparation des variants de masses dus à la perte d'une lysine en C-terminal de la HC et sur les variants de charges de glycoformes du cetuximab.

Les résultats présentés dans cette partie ont fait l'objet d'une publication dans le journal scientifique *Analytical Chemistry (*ACS Publications 2015).

2. Glycoform Separation and Characterization of Cetuximab Variants by Middle-up Off-line CZE-UV/ESI-MS

Michael Biacchi, Rabah Gahoual, Nassur Said, Alain Beck, Emmanuelle Leize-Wagner, Yannis Nicolas François. ***Analytical Chemistry*** 2015, 87(12), 6240-6250.

Abstract

Monoclonal antibodies (mAbs) are highly complex glycoproteins that present a wide range of microheterogeneities that requires multiple analytical methods for full structure assessment and quality control. Capillary zone electrophoresis-Mass spectrometry couplings (CZE-MS), especially by Electrospray (ESI), appear really attractive methods for the characterization of biological samples. However, due to the presence of non- or medium volatile salts in the background electrolyte (BGE), online CZE-ESI-MS coupling is difficult to implement for mAbs isoforms separation. Here we report an original strategy to perform off-line CZE-ESI-MS using CZE-UV/fraction collection technology to performed CZE separation, followed by ESI-MS infusion of the different fractions using CESI interface as nanoESI infusion platform. As the aim is to conserve electrophoretic resolution and complete compatibility with ESI-MS without sample treatment, Hydroxypropylcellulose (HPC) coated capillary was used to prevent analyte adsorption and assymetric CZE conditions involving different BGE at both ends of capillary have been developed. The efficiency of our strategy was validated with the separation of cetuximab charge variant by middle-up approach. Molecular weights were measured for six charge variants detected in the CZE separation of cetuximab subunits. The three first peaks correspond to Fc/2 variants with electrophoretic resolution up to 2.10 and the last three peaks correspond to $F(ab')_2$ variants with average electrophoretic resolution of 1.05. Two Fc/2 C-terminal lysine variants were identified and separated. Moreover separation of Fc/2 fragments allowed the glycoprofiling of the variants with the characterization of 7 different glycoforms. Regarding $F(ab')_2$ domain, 8 glycoforms were detected and separated in three different peaks following the presence of N-glycolyl neuraminic acid residues in some glycan structure. This work highlights the potential of CZE technology to perform separation of mAbs especially when they carry sialic acid carbohydrates.

Introduction

Since 1986 and the approbation of muromonab-CD3 by the US Food and Drug Administration (FDA), monoclo-nal antibodies (mAbs) have taken a major market share in the pharmaceutical industry and their development is constantly increasing [1,2]. MAbs are highly complex glycoproteins potentially displaying many naturally-occurring molecular micro-heterogeneities [3,4]. Patents protecting the first generation blockbuster mAbs will expire in the next 5 years, giving the opportunity to many companies to produce “biogeneric versions”. These copies are referred as biosimilars. Biosimilarity assessment includes extensive physicochemical characterization likewise pharmacokinetic (PK) and pharmacodynamics (PD) study, performed in a comprehensive manner. Analytical high similarity is the most robust scientific basis for comparing independently sourced biologics [5]. As different structural heterogeneities emerged from comparison of a biosimilar candidate with the reference molecule, more complementary studies should be performed in order to demonstrate the absence of toxicological and negative clinical outcome [6].

As a consequence, there is a continuous need for improvement of analytical methods to enable fast and accurate characterization. Mass spectrometry (MS) generally coupled with separation methods such as high-performance liquid chromatography (HPLC), plays a key role in the characterization of therapeutic mAbs [7]. Many levels of characterization are performed following different strategies as intact molecular weight (MW) measurement, top-down, middle-up and bottom-up approaches [7-9]. Concerning glycosylation variants which is known as an important class of modification that can significantly impact the immunogenic properties of mAbs [10], intact MW, middle-up on domain level, as well as bottom-up techniques were also performed [11-14]. However, at this point, no separation of mAbs glycoforms was obtained by HPLC-MS with intact MW or a middle-up approach [15-17]. As an alternative to HPLC, capillary zone electrophoresis (CZE) has been demonstrated to be an useful and powerful separation method for the characterization of intact proteins [18,19]. CZE and related methods using optical detections were fully established at all stages of mAbs discovery. However, due to the presence of high concentration of non-volatile salts in the background electrolyte (BGE) necessary for mAbs variants separation, development of CZE coupled with MS detection have been limited. Recently we published several research papers on bottom-up characterization of mAbs and biosimilars highlighting the potential of CZE-ESI-MS to obtain full primary structure and microvariant characterization as well as biosimilarity assessment [12-14]. However, in bottom-up approach, BGE is totally compatible with ESI-MS, whereas for an intact MW or a middle-up approach, BGE involves high concentration in non-volatile salt or detergent which preclude the use of ESI-MS [20-22]. To our knowledge, some groups defined CZE-UV conditions for the separation of intact mAbs involving always high levels of ε-amino-caproic

acid (EACA) and the addition of triethylenetetramine (TETA) or Tween 20 [23-27]. Despite very interesting results in the separation of charge variants, BGEs are totally incompatible with ESI-MS [28]. In 2014, we developed the first analysis of intact mAb charge variant by CZE using a matrix-assisted laser desorption/ionization-MS (MALDI) detection [29-31]. Unfortunately, limitation of MALDI-MS resolu-tion does not allow to measure the exact mass of the charge variants. More recently, Redman et al published an important paper on the first characterization of intact mAb variants using microfluidic CZE-ESI devices [28].

In this work, we developed a strategy to perform off-line CZE-UV/ESI-MS with the use of CZE-UV/fraction collection technology to performed CZE separation, followed by ESI-MS infusion of the different fractions using CESI interface as nanoESI platform. Hydroxypropylcellulose (HPC) coating was used to minimize analyte adsorption on the capillary wall and reduce electroosmotic mobility (EOF). As one of the aim is to be compatible with ESI-MS, assymetric CZE conditions bringing into play different BGE at the ends of capillary have been developed. Inlet BGE composed of EACA 200 mM and ammonium acetate 25 mM, pH 5.70, allowed to conserve performance of CZE separation while outlet BGE composed only of ammonium acetate 25 mM allowed to be compatible with ESI-MS. Based on the work of Gahoual et al we used CESI interface as nanoESI infusion platform allowed us to generated stable spray at 100 nL/min with sample consumption of 2 µL per fraction avoiding dilution effect [32]. The mAb selected was cetuximab which is human/murine chimeric IgG-1 directed against the epidermal growth factor receptor (EGFR) overexpressed in advanced-stage EGFR positive colorectal cancer [33]. Cetuximab was approved in the US and EU in 2004 and 2005, respectively and will be off-patent soon. Cetuximab contains two sites of glycosylation on the HC: one is located in the Fc/2 domain (Asn299) and the second is located in the F(ab')2 domain on Asn88 [7] (Figure 1). It is an ideal sample for the evaluation of our strategy by middle-up approach. Indeed, cetuximab subunits characterization allowed to locate and to identify glycoforms. The performance of our strategy is demonstrated by the characterization of C-terminal lysine variants and the first separation of glycoforms of cetuximab subunits using CZE-UV off-line hyphenated to ESI-MS after fraction collection.

Experimentals section

Materials. Methanol (HPLC gradient grade) and acetic acid (100%) were obtained from VWR (Radnor, PA, USA). Ammonium acetate (>98%), sodium hydroxide, ε-Amino-caproic acid (>98%), hydroyxypropylcellulose Mw 100000 (HPC) and formic acid (>98%) were purchased from Sigma-Aldrich (Saint Louis, MO, USA). Water used to prepare buffers and sample solutions was obtained using an ELGA purelab UHQ PS water purification system (Bucks, UK).

IdeS (Immunoglobulin-degrading enzyme of Streptococcus pyogenes) also named FabRICATOR was purchased from Genovis (Lund, Sweden). Cetuximab (Erbitux®, Merck KGaA, Darmstadt, Germany) is a sterile, preservative-free solution for intravenous infusion containing 5 mg/mL of cetuximab. The other ingredients are sodium chloride, glycine, polysorbate 80, citric acid monohydrate, sodium hydroxide and water for injections.

Middle-up sample preparation. Cetuximab was cleaved in the hinge region using limited proteolysis by IdeS (FabriCATOR, Genovis) to obtain two Fc/2 fragments (calculated pI 7.74) and one F(ab')2 fragment (calculated pI 7.78) (Figure 1). Sample was diluted using 147.25µL of 50 mM sodium phosphate, 150 mM NaCl, pH 6.60, to a final concentration of 1 µg/µL. A volume of 2.25 µL of IdeS (67 units/µL) was added to the sample which was left at 37°C for 30 min. After digestion completion, sample was desalted using Amicon centrifugal filters (cut off = 10,000 Da) in pure water at 10°C and 14,000 g for 20 min. After desalted step, sample volume was reduced to around 10 µL. Sample was finally diluted to an assumed final concentration of 5 µg/µL in a total volume of 30 µL of pure water.

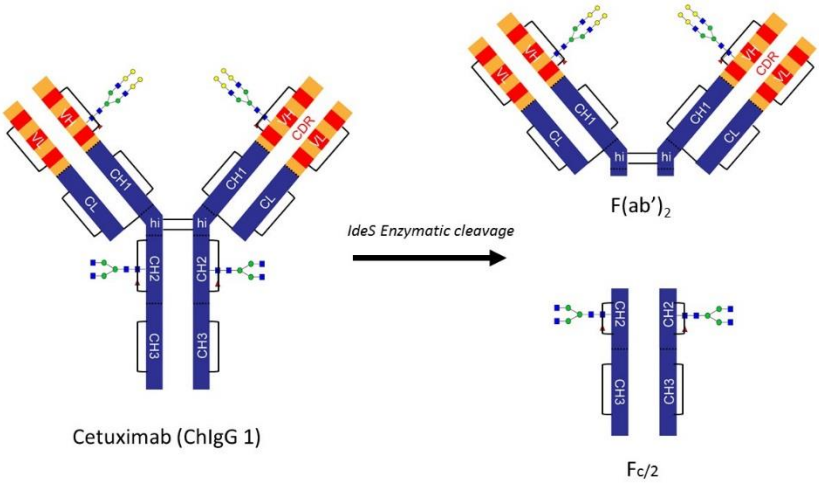

Figure 1. Schematic representation IdeS enzymatic cleavage of cetuximab. Cetuximab is a chimeric IgG: human amino acids are high-lighted in blue (constant domains), murine amino acids in orange (frameworks) and in red (the complementary determining region)

Capillary electrophoresis. The CZE experiments were carried out on a P/ACE MDQTM CE system from Sciex Separation (Brea, CA) equipped with a UV detection, a temperature controlled autosampler and a power supply able to deliver up to 30 kV. A 32 KaratTM 8.0 (Sciex Separation, Brea, CA) was used for instrument control, data acquisition and data handling. Polymicro bare fused-silica capillaries of 75 µm i.d., 375 o.d. (75.5 cm effective length, 82 cm total length) were obtained from Photonlines (St-Germain-en-Laye, France). New capillaries were conditioned by successive flushes with 1.00 M and 0.10 M NaOH and then with water under a pressure of 30 psi for 10 min each. The temperature in the capillary

cartridge and autosampler were set at 25 °C. The acquisition rate was 10 points / s. Capillaries were rinsed with water and dried by air when not in use. UV absorbance fixed at 200 nm. Voltage applied at 20 kV with a ramp of 0.17 and injection sample condition was 0.5 psi for 50s. Concerning modified capillaries, capillaries were coated in laboratory with hydroxypropylcellulose Mw 100000 (HPC) following the protocol described by Shen et al [34]. 5% HPC in pure water (w/v) was prepared to perform capillary coating. Durability of coating is around 20 runs without recoating step. For Cetuximab separation, inlet BGE 200 mM ε-Amino-caproic acid (EACA)-ammonium acetate 25 mM pH 5.70 and outlet BGE ammonium acetate 25 mM pH 5.70 have been used as separation condition. Injection volumes have been calculated using CEToolbox application (Pansanel, GooglePlay).

CZE/Collection Fraction Interface. This interface is describe in a previous study31. Briefly, automated off-line coupling of CZE to MS was performed by using a homemade modified automatic spotting device Proteineer FC (Bruker Daltonics, Bremen, Germany) for the sheath flow-assisted spotting from the CZE capillary end onto a fraction collection target. The original set up of the UV cell in the P/ACE MDQTM (Sciex Separation, Brea, CA) was modified in order to allow the simultaneous UV detection and fraction collection. Hystar 3.2 (Bruker Daltonics, Bremen, Germany) was used for Proteineer FC control.

CESI as nanoESI infusion platform. The infusion experiments were carried out with a PA 800 plus capillary electrophoresis (CE) system from Sciex Separation (Brea, CA) equipped with a temperature controlled autosampler and a power supply able to deliver up to 30 kV. Prototype fused-silica capillaries (total length 95 cm; 30 µm i.d.) whose outlet end (about 3 cm) was etched with hydrofluoric acid were used for all related CESI experiments and initially provided by Sciex Separation (Brea, CA, USA). It is certainly noteworthy to precise here that the inner lumen of these capillaries is not tapered and presents an i.d. of 30 □m throughout its entire length. New capillaries were initially conditioned by flushing them for 10 min with MeOH, 10 min with 0.10 M sodium hydroxide, 10 min with 0.10 M hydrochloric acid and finally with water for 20 min each flushing step being conducted at 75 psi (5.17 bar).

MALDI-TOF-MS. The matrix was prepared by dissolving 2,5-dihydroxybenzoic acid (DHB) (2g/L) in 0.1% trifluoroacetic acid/acetonitrile (TFA/ACN) (30/70)(v/v). Fraction collection was realized using Ground Steel MALDI target (Bruker Daltonics, Bremen, Germany). Mass spectra of the CZE fractions were recorded using an Autoflex II MALDI-TOF (Bruker Daltonics, Bremen, Germany), operating in reflector mode and with FlexControl software. Positively charged ions were detected and sums of 1500 single-shot spectra were acquired automatically from each sample by using the AutoXecute software. Data processing was performed with FlexAnalysis 3.0 provided by the mass spectrometer manufacturer. All spectra were calibrated

according an external calibration using Protein calibration standard I (Bruker Daltonics, Bremen, Germany) for intact protein separation.

ESI-TOF-MS. For sheathless CZE-ESI-MS experiments the CE system was coupled to a maxis 4G (Bruker Daltonics, Bremen, Germany). MS transfer parameters were optimized using the actual sample directly infused via the CE system using a pressure of 5 psi (340 mbar). MS parameters were optimized so that high m/z ions could be properly transferred to the TOF analyzer while avoiding fragmentation. In the case of the maxis 4G, ion funnels were set at values of 300 Vpp and 400 Vpp. The electrospray voltage (capillary voltage) was typically ranging from -1.2 to -1.8 kV. Dry gas was set at 1.5 L/min and source temperature at 180°C. Data processing was performed with DataAnalysis 4.0. Deconvolution of the mass spectra was performed based on maximum entropy analysis using ESI Compass 1.3 Maximum Entropy Deconvolution Option in DataAnalysis 4.0. All spectra were calibrated by external calibration using Pepmix (Bruker Daltonics, Bremen, Germany) and CsI from Sigma-Aldrich (Saint-Louis, MO, USA).

Results and discussion

Optimization of CZE separation conditions.

As the aim is to develop a procedure for the middle-up characterization of cetuximab using the potential of CZE separation method and suitable for ESI-MS direct infusion, we have optimized BGEs as mixtures of components that facilitate isoforms separation and allow the use of ESI ionization process. First we based our research on the study developed by the group of Pr. Somsen on the glycoform profiling of interferon-β-1a and erythorprotietin [19]. However, using condition based on HPC coated capillary and 50 mM acid BGE, no separation of cetuximab isoforms had been obtained. To our knowledge, some groups defined CZE-UV conditions for the separation of intact mAbs involving always high levels of EACA and the addition of TETA or Tween 20 [23-27]. EACA is a zwitterion which can be used to create high ionic strength and low conductivity buffers that minimize the electroosmotic flow (EOF) and improve electrophoretic mobility differences. TETA and Tween 20 behave as a modifier in BGE for dynamic coating to reduce prevent analyte adsorption in the capillary wall [35-37]. Despite very interesting results in the separation of charge variants due to the presence of these modifiers, the major issue of these BGE conditions is the incompatibility with ESI-MS detection. The first step of our optimization consisted to eliminate Tween 20 and TETA of the BGE. Indeed, following the results described by Gassner *et al*, we minimize analyte adsorption only by the use of a HPC coating [25]. This allows to remove the use of additional detergent or oligoamine, and it additionally presents the advantage of being a static coating that avoid

potential polymer release for ESI-MS infusion. In a previous work, we demonstrated the high repeatability and the good robustness obtained with HPC-coated capillary for the separation of intact mAbs (RSD<0.5%, migration time) [31]. The second step of our BGE screening concerned the use of high level of EACA. He *et al* demonstrated the influence of pH and EACA concentration on the mAb separation [23]. Concerning EACA concentration, they demonstrated that increasing EACA concentration may improve separation efficiency and electrophoretic resolution. However, due to zwitterionic properties, EACA based BGE involve significant interference with the sample in the ionization process of ESI-MS. Despite this, the presence of EACA in the BGE remains essential to maintain the separation efficiency particularly due to the low conductivity. Based on the results of Ruesch's team [23] and some difficulties of matrix/sample crystallization for MALDI-MS experiment observed in previous study, we decided to reduce the EACA concentration from 400 mM to 200 mM and to add ammonium acetate at a pH of 5.70. Ammonium acetate is usually employed with ESI-MS mainly for volatility properties of the ammonium ion and good compatibility with the process of ionization. To perform separation we used our homemade CZE-UV/fraction collection described in a previous work [31]. BGEs are placed in the inlet and outlet vials of the CE apparatus. Due to the modification of CE cartridge, outlet BGE has been used as sheath liquid with a flow rate of 0.5µL/min. Fraction collection was directly deposit on a target plate. Each peak is collected on the basis of its apparent mobility described by the equation:

$$\mu_{app} = \frac{\mathrm{L\,l}}{\mathrm{t}_m\,\mathrm{V}}$$

with μ_{app} the apparent mobility which is the sum of effective mobility and residual EOF, L and l the total capillary length and length to detection window, respectively, V the applied voltage and t_m the migration time. Deposition time t_d is then calculated by the equation:

$$t_d = \frac{L^2}{\mu_{app} V}$$

The results presented Figure 2 show the CZE-UV separation of cetuximab IdeS fragments (Figure 1) with different BGEs at pH 5.70 composed by the mixture of EACA 200 mM and various concentration of ammonium acetate from 5 mM to 50 mM. Electropherograms presented Figure 2 show a drastic decrease of electrophoretic resolution with lowest concentration of ammonium acetate until complete loss of separation with 5 mM ammonium acetate (electropherogram A). These results confirm the trend in loss of electrophoretic resolution with the decrease of the ionic strength. Electropherogram D corresponding to BGE composed by EACA 200mM ammonium acetate 50mM pH 5.70, show the best separation with

electrophoretic resolution up to 1.63 for peak 1-3 corresponding to Fc/2 variants and an average electrophoretic resolution of 1.02 for peak 4-6 corresponding to F(ab')2 variants.

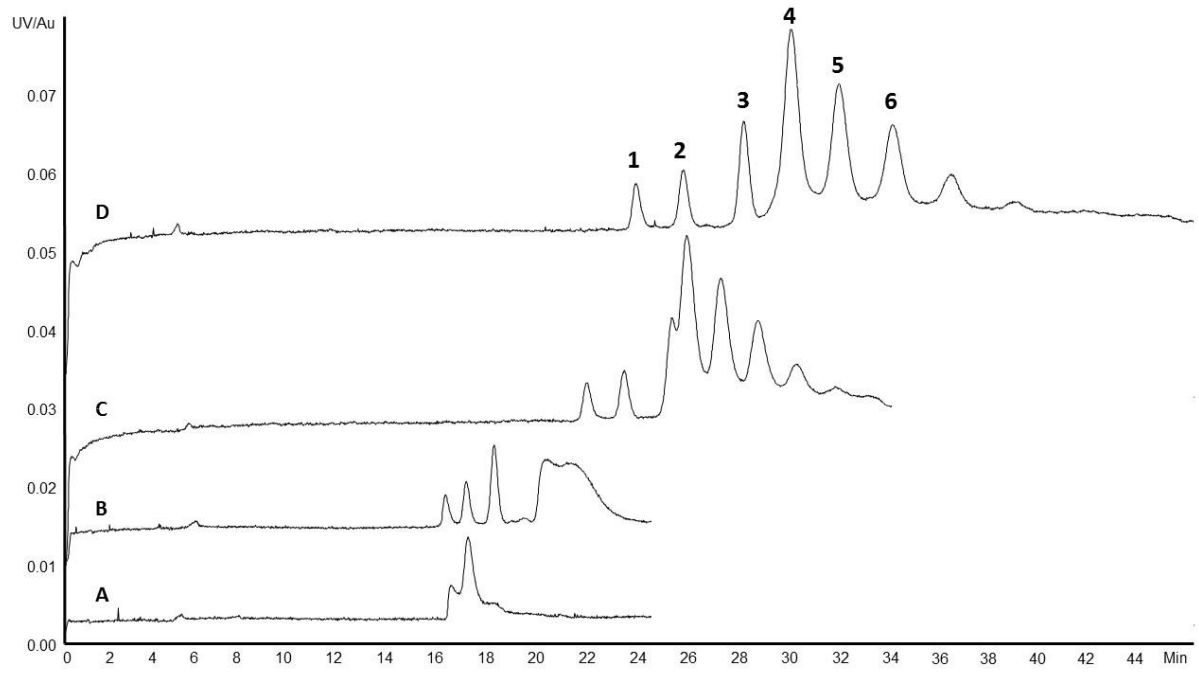

Figure 2. Impact of ammonium acetate concentration in BGE on electrophoretic resolution. BGE composed by a mixture of EACA 200 mM and acetate ammonium (A) 5 mM, (B) 10 mM, (C) 25 mM and (D) 50 mM, at pH 5.70. Peak 1-3 correspond to Fc/2 variants and peak 4-6 correspond to F(ab')2 variants. Experimental conditions: HPC-coated capillary, total/effective length 82/75.5 cm x 75 µm i.d.; voltage, 20 kV; UV absorbance at 200 nm; sample, IdeS digest of cetuximab (5 µg/µL); sample injection 0.5 psi 50s.

However, outlet BGE which plays the role of sheath liquid is composed of 200 mM EACA which can compromise the ESI-MS ionization process. Indeed, to verify the impact of the presence of EACA in the BGE, we perform MALDI-MS mass spectra of the different fractions corresponding to the six major peaks. No signals were observed by MALDI-MS because of the lack of crystallization of the sample with the DHB matrix (Figure S-1 part A). Presence of EACA in the BGE gives white homogenous deposit totally different of the spangled classical one. MALDI-MS detection failure and the experimental ESI-MS infusion of fraction 3 without any MS signal (data not shown) allows us to conclude that direct ESI-MS infusion is completely impossible in these conditions.

Based on the CZE-ESI-MS sheath liquid interface described by the group of Smith [38], and the fact that additional liquid is usually different of BGE composition, we decided to develop asymmetric conditions between the inlet BGE and the outlet BGE. As the aim of the study is to conserve the electrophoretic resolution of the separation, to allow ESI-MS detection and to avoid the Joule heating effect, inlet BGE has been chosen as EACA 200 mM and ammonium

acetate 25 mM pH 5.70. A different BGE, placed in the outlet vial, is composed only with ammonium acetate at pH 5.70.

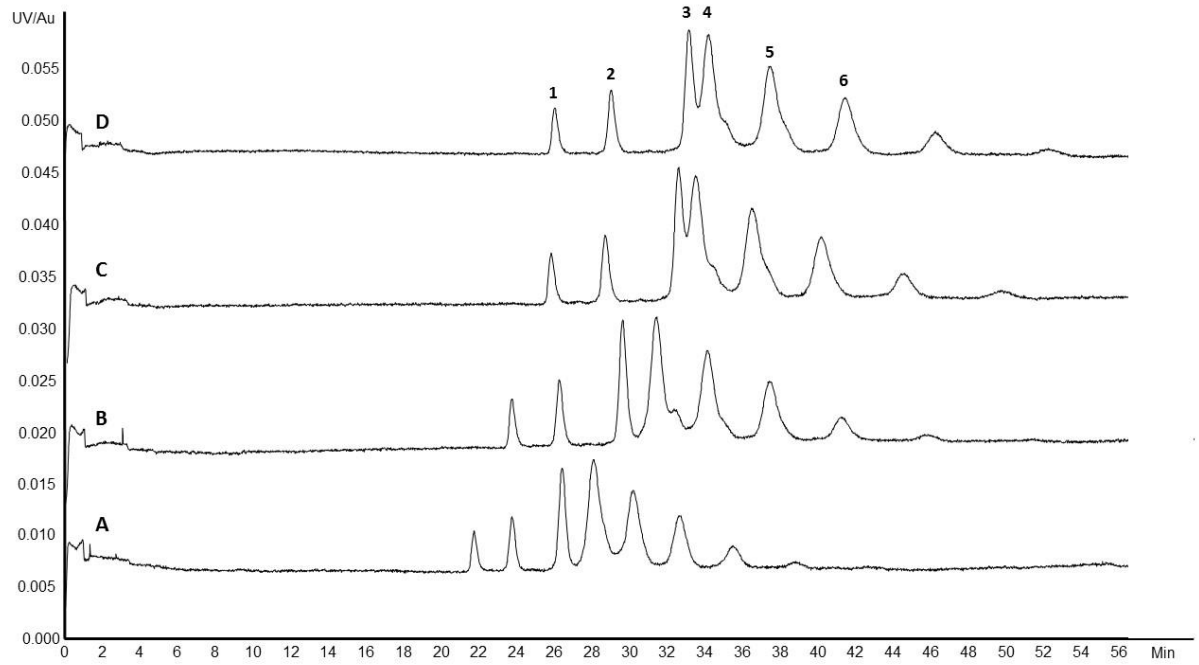

Figure 3. Impact of ammonium acetate concentration in outlet BGE on electrophoretic resolution. Inlet BGE composed by a mixture of EACA 200 mM and acetate ammonium 25 mM pH 5.70 and outlet BGE by ammonium acetate (A) 12.5 mM, (B) 25 mM, (C) 50 mM and (D) 100 mM, at pH 5.70. Peak 1-3 correspond to Fc/2 variants and peak 4-6 correspond to F(ab')2 variants. Experimental conditions: HPC-coated capillary, total/effective length 82/75.5 cm x 75 µm i.d.; voltage, 20 kV; UV absorbance at 200 nm; sample injection 0.5 psi 50s.

To optimize the separation, different concentration of ammonium acetate from 12.5 to 100 mM have been tested as outlet BGE. Results are presented in Figure 3. Electropherogram C and D show a loss of electrophoretic resolution between peak 3 and peak 4 corresponding to the last peak of Fc/2 variants and the first peak of F(ab')2 variants. Electropherogram B corresponding to outlet BGE composed by ammonium acetate 25mM pH 5.70, show the best separation with electrophoretic resolution up to 2.10 for peak 1-3 corresponding to Fc/2 variants and an average electrophoretic resolution of 1.05 for peak 4-6 corresponding to F(ab')2 variants. These results show no loss of electrophoretic resolution using electropherogram B conditions as compared to the Figure 2 experiments. This demonstrates that the presence of high concentration of EACA in the outlet BGE is not necessary to achieve the separation. As the aim was to verify the impact of the absence of EACA in the outlet BGE, we performed a MALDI-MS detection of the different fractions corresponding to the six major peaks. First of all, we observed a good crystallization of sample with DHB matrix (Figure S-1 part B) which confirms no contamination of EACA during deposit process. Figure 4 emphasizes mass spectrum of each peak. Moreover, the CZE-MALDI-MS (Fig. 4b) confirms the good

agreement between the UV detection and the deposition time in terms of Fc/2 and F(ab')2 separation.

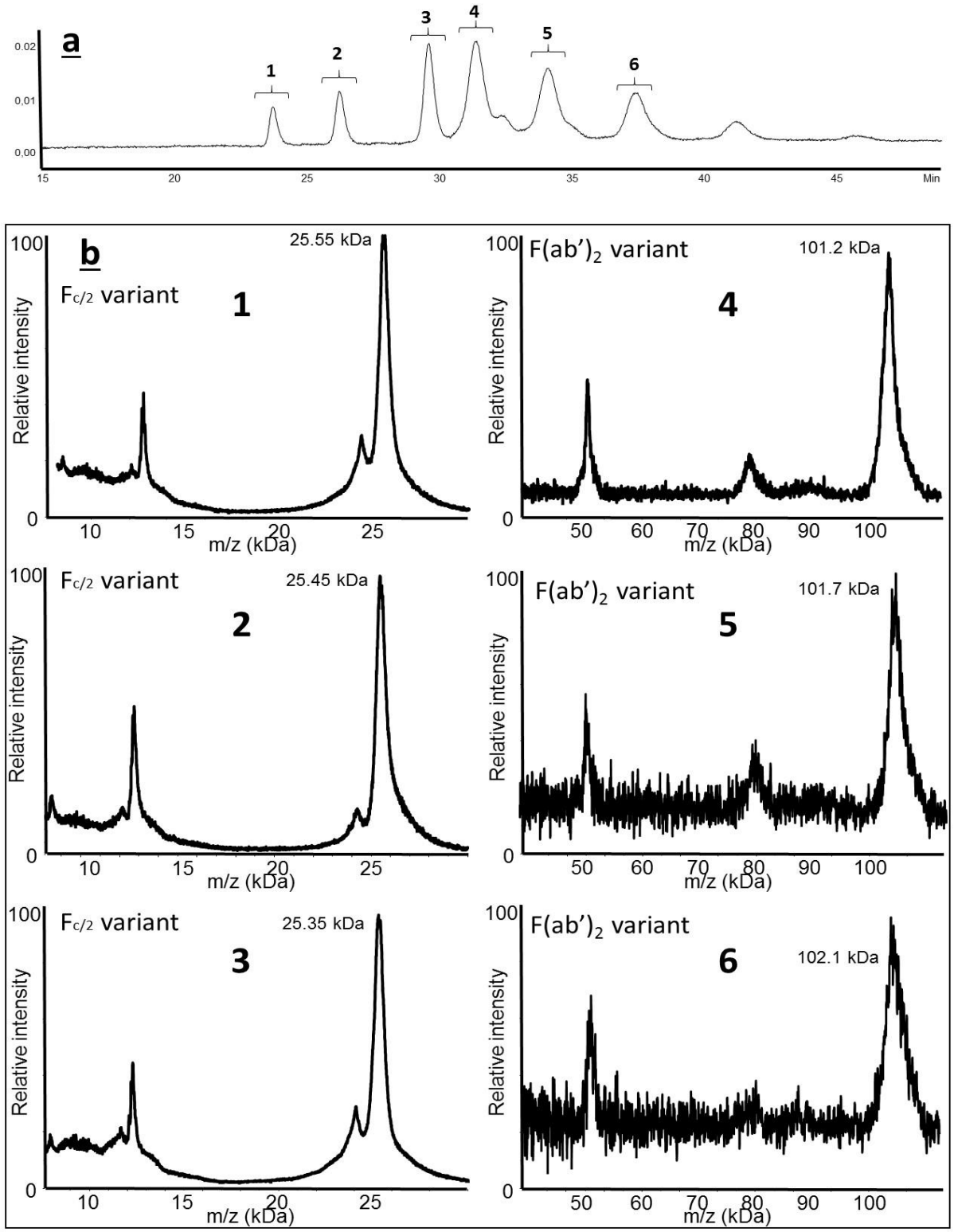

Figure 4. CZE-UV off-line coupled to MALDI-MS via fractionation for a middle-up characterization of Cetuximab charge variants. (a) CZE-UV electropherogram. (b) Analysis of CZE-UV fractions by MALDI-MS. Mass spectra of the six major peaks. CE Experimental conditions: Inlet BGE: EACA 200 mM acetate ammonium 25 mM pH 5.70 and outlet BGE: ammonium acetate 25 mM pH 5.70; other conditions: See Figure 2, MS Experimental conditions: See Experimental section; sample, IdeS digest of cetuximab (5 µg/µL); sample injection 0.5 psi 50s.

This also confirms the absence of carryover effect and diffusion phenome-non. Unfortunately, limitation of MALDI-MS resolution for molecules up to 20 kDa does not allow to measure the exact mass of the charge variants and then to characterize thoroughly these glycoproteins. Thus, to reach higher MS resolution at comparable separation efficiency, we switched to ESI-MS detection using CESI as nanoESI infusion platform. This strategy of collection fraction followed by nanoESI infusion opens the way for a deeper characterization of Fc/2 and F(ab')2 domains without sample treatment and with minimum sample volume (2 µL).

Middle-up characterization of cetuximab

Cetuximab is a chimeric mouse-human IgG1 known to bear 2 N-glycosylation sites on each heavy chain (HC) [11]. Moreover, cetuximab has a large number of microheterogeneities such as PTMs including methionine oxidation, asparagine deamidation or isomerization of aspartic acid. Furthermore, this mAb also has one C-terminal lysine truncation. These features make cetuximab an ideal sample for the evaluation of a middle-up approach using CE-MS coupling. In order to simplify the location of N-glycosylation, IdeS enzymatic reaction (Figure 1) cleaves cetuximab in the middle of HC to obtain two types of fragments each carrying one N-glycosylation site. Direct ESI-MS infusion of IdeS fragments of cetuximab proved equal signal abundances for C-terminal lysine variants which involves two distinct profiles for the Fc/2 domains, and a complex glycoprofiling for the F(ab')2 domain (Figure S-2). With this knowledge, we can expect to observe few peaks corresponding to Fc/2 fragments due to a loss of a +1 charge associated with C-terminal lysine truncation. In addition, we also expect few peaks for F(ab')2 fragments because of the ability of CZE to resolve glycoforms with acidic glycans, due to a reduction in net charge by addition of negatively charged sialic acid compound. Preliminary MALDI-MS results presented in Figure 4 confirmed these predictions with peak 1-3 corresponding to Fc/2 fragments and at least peak 4-6 corresponding to F(ab')2 fragments.

To enhance MS resolution, the same experiment was performed to collect fractions of each peaks and to analyze them by nanoESI-MS using CESI interface as infusion platform. After deposition, each fraction was collected and evaporated. No special treatment was made on collected fractions. Before infusion, dry samples were reconstituted in 2µL of 50%/49%/1%, acetonitrile/water/formic acid (v/v/v) in order to enhance sample ionization. Thus, each fraction was infused into the MS using a flow rate of 100 nL/min in order to enhance sensitivity. The total volume of sample infused to perform one acquisition was about 600 nL which represents a 100 times reduction in terms of sample consumption as compared to standard ESI-MS. This point is critical as the volume of each fraction does not expect few µL to avoid dilution effect.

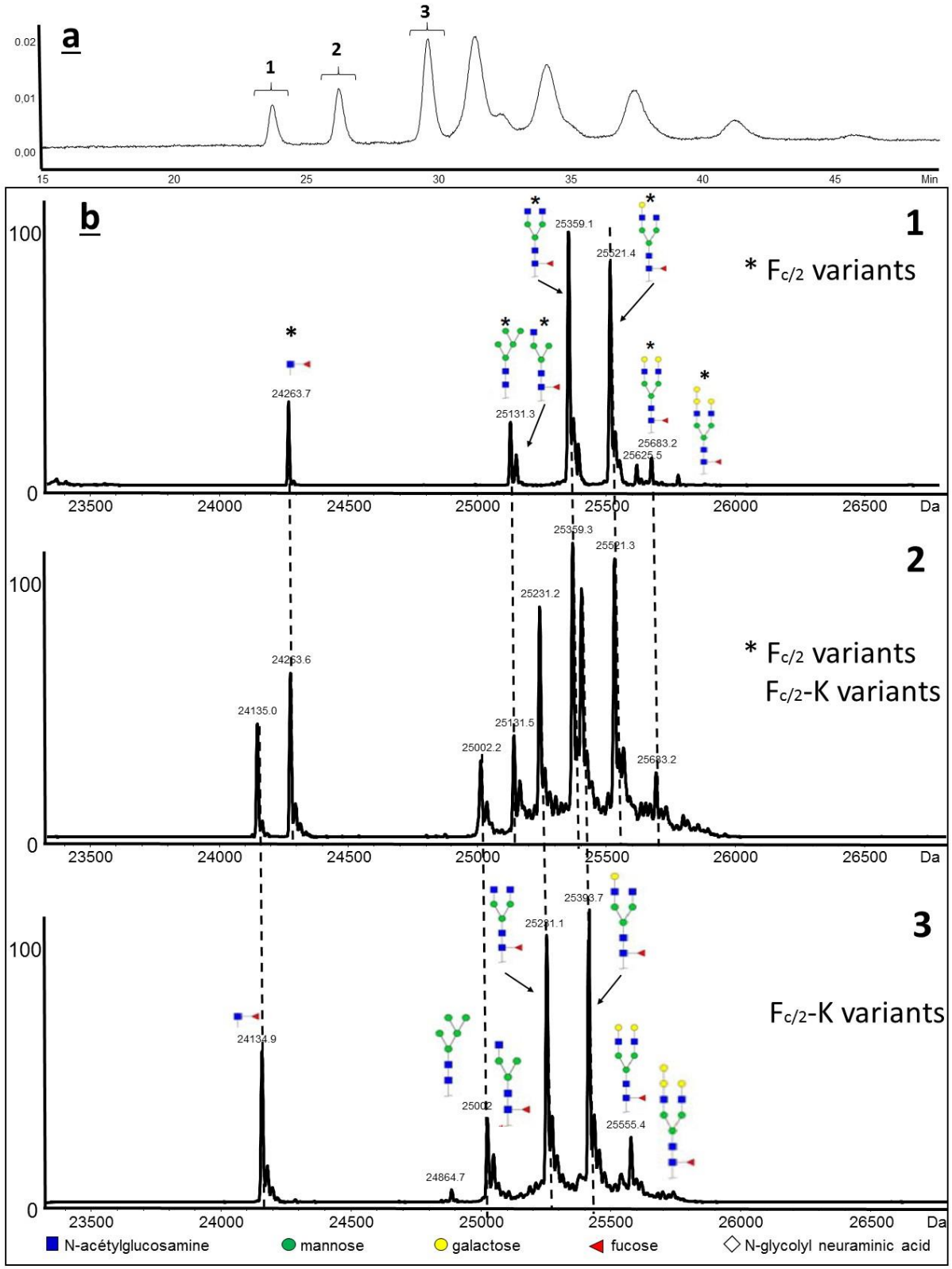

Figure 5. (a) Off line CZE-UV/ESI-MS separation of middle-up cetuximab charge variants. (b) Deconvoluted mass spectra for each Fc/2 variant. The MS peak were labeled with the correspondent glycoform. Experimental conditions: Inlet BGE: EACA 200 mM ammonium acetate 25 mM pH 5.70 and outlet BGE: ammonium acetate 25 mM pH 5.70; other conditions: See Figure 2, MS Experimental conditions: See Experimental section; sample, IdeS digest of cetuximab (5 µg/µL); sample injection 0.5 psi 50s.

Deconvoluted charge variant mass spectrum of fractions from peak 1 to peak 3 generated multiple masses corresponding to the expected presence of neutral glycosylation variants of Fc/2 fragments (Figure 5). In addition, the comparison with the direct infusion experiment (Figure S-2) shown us that the two Fc/2 fragments corresponding to intact and with one lysine truncation are detected and separated. Deconvoluted mass spectra corresponding to peak 1 and peak 3 show a difference of 128 Da between mass peak 25359.1 Da corresponding to the most abundant Fc/2 glycoform (G0F) and mass peak 25231.1 Da corresponding to G0F for Fc/2-K. Thus, we obtained the complete separation of these two fragments (Figure 5-b 1 and 3). In agreement with Ayoub et al and with the comparison of direct infusion of cetuximab IdeS subunits (Figure S-2), only one possibility of C-terminal lysine loss was expected. Regarding, deconvoluted mass spectra of peak 2, it shows a nearly equal abundance mixture of Fc/2 and Fc/2-K variants. These peaks deconvolute to the same MWs founded to peak 1 and peak 3 and could be interpreted by an overlap during the deposit process. However, regarding the MS signal of raw data used to generate deconvoluted spectra of peak 1, peak 2 and peak 3, and considering a potential equivalent ionization efficiency of all Fc/2 variants, intensity of peak 2 is higher than that of peak 1 and less than that of peak 3 (Figure S-3) following the CZE separation pattern. This observation should eliminates a potential carryover effect (peak broadening, excessive diffusion) which should show a concomitant decrease of peak 1 MS intensity and increase of peak 2 MS intensity. In addition, to exclude carryover effect, we performed infusion of fractions corresponding to peak 4 and to a period of 2 min before detection of peak 1. No ESI-MS spectra have been detected corresponding to Fc/2 fragment (data not shown). Cetuximab, as other mAbs, is known to have a large number of microheterogeneities commonly found in proteins which are reflected by differences in mass, charge or conformation and then differences in effective mobilities. Moreover, it is plausible that several cumulative modifications could lead to a variation of conformation without mass shifts. Although further work is required to identify these modifications.

An important class of modification that can significantly impact the immunogenic properties of cetuximab in terms of PK and PD is glycosylation [10]. Cetuximab contains two sites of glycosylation on the HC: one is located in the Fc/2 domain (Asn299) and the second is located in the F(ab')2 domain on Asn88 [7]. Cetuximab has been widely described in terms of glycan characterization [4,39]. In 2013, Janin-Bussat et al used a middle-up approach similar to the one described here but using HPLC-MS method, and identified 17 different glycoforms distributed on the two glycosylation sites [40]. In 2013, team of Dr. Beck characterized 24 different glycosylation using the combination of intact, middle-up, middle-down and bottom-up ESI and MALDI-MS techniques [11]. More recently Gahoual et al characterized 15 glycopeptides using a bottom-up approach by CESI-MS/MS [13]. In our work, we could identify

a total of 15 glycans divided among 7 for the Fc/2 N-glycosylation and 8 for the F(ab')2 N-glycosylation (Table S-1 and Table S-2). Regarding Fc/2 domain, branched glycan structures are primarily composed of fucose, N-acetylglucosamine, and mannose but can differ in the number of terminal galactose, mannose or N-acetylglucosamine residues. Glycan structure with the addition of galactose moieties involves a mass increase of 162 Da but does not induce a change in net charge [41,42]. Gahoual and coworkers demonstrated that particular glycopeptides having a difference of one galactose could be baseline separated13 whereas Redman et al did not observed mobility shifts between intact mAb glycoforms due to the low impact of 162 Da on the mass of intact mAb (≈0.1%) [28]. Our work demonstrated that concerning glycoform with a difference of one galactose residue, middle-up approach does not allow to observe baseline separations. The mass of galactose moiety accounts for ≈0.6% of the Fc/2 domain and ≈0.2% of the F(ab')2 domain. Following the conclusion of Redman et al, to induce effective mobility shift, the mass increase would have to be much greater or induce a significant change in conformation. However, CE has the potential to separate glycoforms bearing acidic residues, due to a reduction in net charge from the addition of negatively charged sialic acid [41]. In this point, cetuximab appears to be an ideal sample.

Indeed, regarding F(ab')2 domain, Ayoub et al identified glycosylations with 41% glycans bearing N-glycolyl neuraminic acid (NGNA) in their structures. NGNA glycoforms result from the SP2/0 murine expression system used to produce cetuximab [43]. When produced in CHO cells, biobetter version exhibit N-acetylneuraminic acid (NANA) glycoforms [9]. Despite the high proportion of the NGNA glyclan, no separation of F(ab')2 variants was obtained using HPLC-MS [11]. As shown in Figure 6, electrophoretic peaks from 4 - 6 correspond to different F(ab')2 variants. Deconvolution of each charge variant mass spectrum generated multiple masses, corresponding for peak 4 to the expected presence of neutral glycosylation variants. Following the theory developed previously for the Fc/2 variants, neutral residues did not allow effective mobility shift of F(ab')2 glycoforms due to the small impact of galactose mass residue in the F(ab')2 mass. The mass ascribed to the most intense mass peak 101924 Da (Figure 6b-4) correspond to the most abundant F(ab')2 glycoform (G2FGal2/ G2FGal2). Four other less abundant F(ab')2 glycoforms are labeled each one bearing galactose residues (Figure 6b-4). Deconvoluted mass spectra presented in Figure 6b-5 correspond to the F(ab')2 glycoforms contained in peak 5. In this fraction the most abundant mass peak 102069 Da present a mass difference of 145 Da compared to GF2Gal2/ GF2Gal2 which corresponds to the presence of a NGNA residue instead of a galactose residue. This mass peak corresponds to the second most abundant F(ab')2 glycoform (GF2Gal2/ GF2GalNGNA). Another mass peak at 102592 Da has been labeled in the mass spectra. This corresponds to a less abundant F(ab')2 glycoform also bearing one NGNA residue in the glycan structure. We note that common mass peaks are

detected in fraction 4 and 5. Indeed, the electrophoretic resolution of 1.05 does not allow the complete separation however we observe an important decrease of intensity for the glycoforms detected in fraction 4.

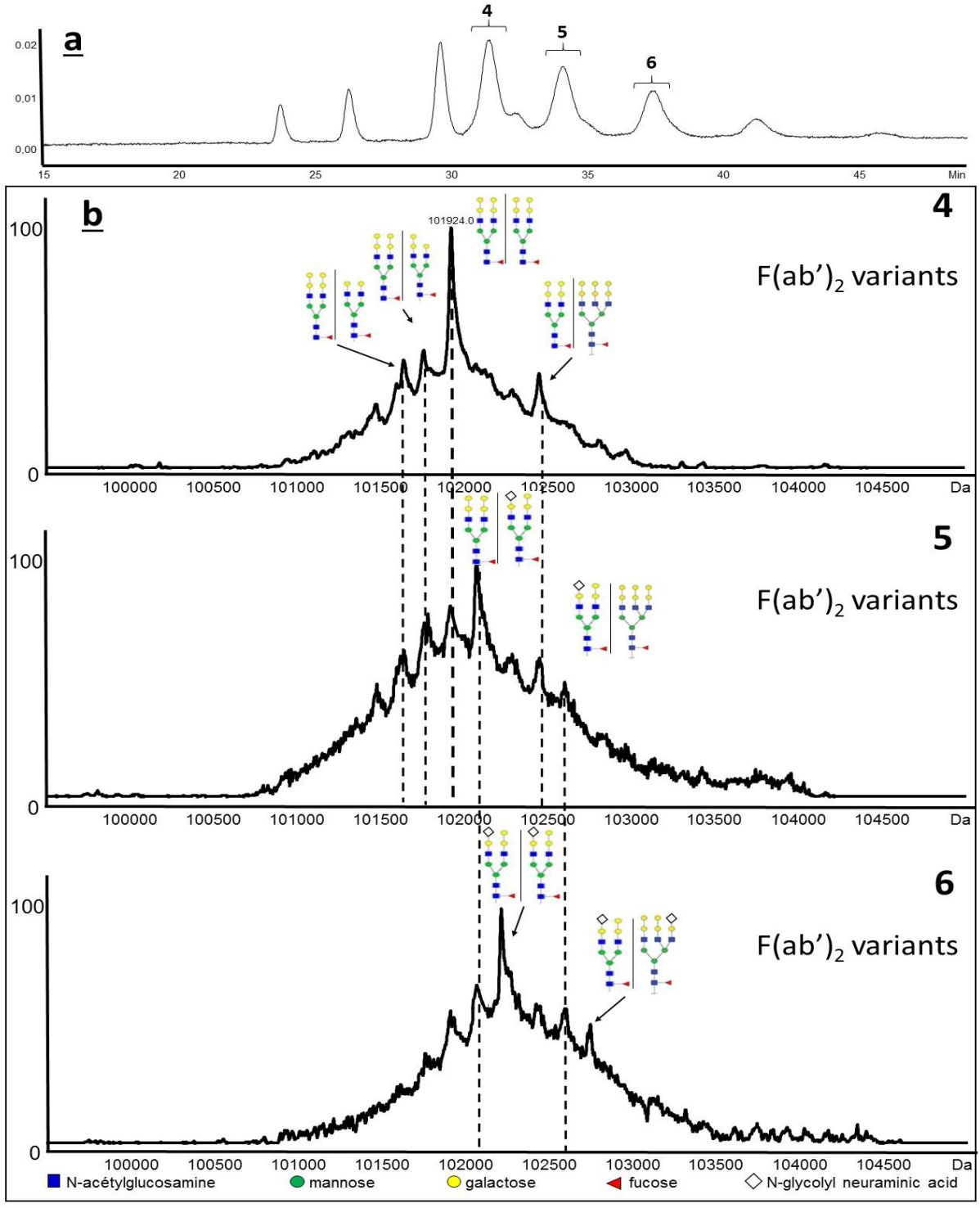

Figure 6. (a) Off line CZE-UV/ESI-MS separation of middle-up cetuximab charge variants. (b) Deconvoluted mass spectra for each F(ab')2 variant. The MS peak were labeled with the correspondent glycoform. Experimental conditions: Inlet BGE: EACA 200 mM ammonium acetate 25 mM pH 5.70 and outlet BGE: ammonium acetate 25 mM pH 5.70; other conditions: See Figure 2, MS Experimental conditions: See Experimental section; sample, IdeS digest of cetuximab (5 µg/µL); sample injection 0.5 psi 50s.

Finally, the infusion of fraction 6 corresponding to the sixth peak of the electrophoretic separation (Figure 6b-6), gives deconvoluted mass spectra corresponding to F(ab')2 glycoforms with the combination of glycan structures each bearing one NGNA residue (GF2GalNGNA/GF2GalNGNA). Indeed, the most intense peak showing a deconvoluted mass of 102217 Da also presents a difference of 145 Da compared to GF2GalNGNA/GF2GalNGNA which correspond to the presence of a second NGNA residue instead of a galactose residue. Moreover, as in fraction 5, another mass peak at 102741 Da has been detected and correspond to a less abundant in the F(ab')2 glycoforms also each bearing one NGNA residue in the glycan structures. To our knowledge, these results represent the first separation of F(ab')2 glycoforms using CZE-MS coupling. Knowing that these glycosylations have been shown to be responsible for immunogenic responses [10], separation of these glycoforms can be very helpful for the characterization of cetuximab especially to highlight batch-to-batch variations in the glycosylation profiles or for biosimilarity assessment. Furthermore, no similar results have been described using HPLC-MS coupling. This highlights the potential of CZE separation to perform middle-up approach of mAbs especially when they carry capping sialic acid carbohydrates.

Conclusion

To summarize, we have reported an original strategy bringing into play off-line CZE-UV/fraction collection method followed by nanoESI-MS infusion for the characterization of a therapeutic mAb. In order to realize mAb charge variants separation and to perform nano-ESI-MS detection, CZE conditions were optimized. First, to eliminate BGE additives, HPC coated capillary utilized to obtain residual EOF and to minimize analyte adsorption. Second, as the aim is to conserve electrophoretic resolution and complete compatibility with ESI-MS, assymetric CZE conditions involving different BGE at the ends of capillary have been described. Inlet BGE has been chosen as a mixture of EACA 200 mM/ammonium acetate 25 mM, pH 5.70 and outlet BGE composed of only ammonium acetate 25 mM, pH 5.70. MWs were measured for six charge variants detected in the CZE separation of cetuximab subunits through deconvolution of the mass spectra. The three first peaks correspond to Fc/2 variants and the last three peaks correspond to F(ab')2 variants. Separation performance obtained with these conditions shown electrophoretic resolution up to 2.10 for Fc/2 variants and an average electrophoretic resolution of 1.05 for F(ab')2 variants. Regarding Fc/2 domain, two C-terminal lysine variants were identified and separated. Moreover the separation of Fc/2 fragments allowed the glycoprofiling of the variants with the characterization of 7 glycoforms each bearing neutral residues to the glycan structure. Regarding F(ab')2 domain, deconvolution of the mass

spectra allowed the determination of 8 glycoforms detected and separated in different peaks. Glycoforms separation is due to the presence of NGNA residues in some glycan structure. Four glycosylations of F(ab')2 variants were identified by characteristic 145 Da mass shift, moreover this identification was supported by the decrease of effective mobility due to a reduction in net charge from the addition of negatively charged sialic acid. To our knowledge, this work represent the first demonstration of middle-up approach of mAb using CZE-MS and in addition the first separation of F(ab')2 glycoforms. This highlights the potential of CZE to perform separation of mAbs especially when they carry sialic acid groups. Moreover, the strategy of collection fraction followed by nanoESI infusion without sample treatment and in sample volume of 2 µL open the way to different deeper characterization with for example a top-down approach using ESI-ETD-MS most likely unraveling additional information regarding the structure of the protein.

Acknowledgments

Authors would like to thank Sciex separations Inc. for lending a CESI prototype and Dr. M. Anselme from Sciex separations Inc. for his support. The authors would like also to express their gratitude to Dr. E. Wagner-Rousset, Dr. D. Ayoub, MC. Janin-Bussat and O. Colas (Centre d'Immunologie Pierre Fabre, St Julien en Genevois, France) for helpful discussions around antibody structural characterization

References

[1] Beck, A.; Wurch, T.; Bailly, C.; Corvaia, N. Nat Rev Immunol 2010, 10, 345-352.

[2] Reichert, J. M. mAbs 2012, 4, 413-415.

[3] Arnold, J. N.; Wormald, M. R.; Sim, R. B.; Rudd, P. M.; Dwek, R. A. Annu Rev Immunol, 2007, 25, 21-50.

[4] Zhang, Z. Q.; Pan, H.; Chen, X. Y. Mass Spectrom Rev 2009, 28, 147-176.

[5] McCamish, M.; Woollett, G. Clin Pharmacol Ther 2013, 93, 315-317.

[6] Beck, A.; Diemer, H.; Ayoub, D.; Debaene, F.; Wagner-Rousset, E.; Carapito, C.; Van Dorsselaer, A.; Sanglier-Cianferani, S. Trac-Trend Anal Chem 2013, 48, 81-95.

[7] Beck, A.; Sanglier-Cianferani, S.; Van Dorsselaer, A. Anal Chem 2012, 84, 4637-4646.

[8] Beck, A.; Wagner-Rousset, E.; Ayoub, D.; Van Dorsselaer, A.; Sanglier-Cianferani, S. Anal Chem 2013, 85, 715-736.

[9] Beck, A.; Debaene, F.; Diemer, H.; Wagner-Rousset, E.; Colas, O.; Dorsselaer, A. V.; Cianférani, S. J Mass Spectrom 2015, 50, 285-297.

[10] Chung, C. H.; Mirakhur, B.; Chan, E.; Le, Q.; Berlin, J.; Morse, M.; Murphy, B. A.; Satinover, S. M.; Hosen, J.; Mauro, D.; Slebos, R. J.; Zhou, Q. W.; Gold, D.; Hatley, T.; Hicklin, D. J.; Platts-Mills, T. A. E. New Eng J Med 2008, 358, 1109-1117.

[11] Ayoub, D.; Jabs, W.; Resemann, A.; Evers, W.; Evans, C.; Main, L.; Baessmann, C.; Wagner-Rousset, E.; Suckau, D.; Beck, A. Mabs 2013, 5, 699-710.

[12] Gahoual, R.; Burr, A.; Busnel, J. M.; Kuhn, L.; Hammann, P.; Beck, A.; Franccois, Y. N.; Leize-Wagner, E. Mabs 2013, 5, 479-490.

[13] Gahoual, R.; Busnel, J.-M.; Beck, A.; François, Y.-N.; Leize-Wagner, E. Anal Chem 2014, 86, 9074-9081.

[14] Gahoual, R.; Biacchi, M.; Chicher, J.; Kuhn, L.; Hammann, P.; Beck, A.; Leize-Wagner, E.; Francois, Y. N. Mabs 2014, 6, 1464-1473.

[15] Talebi, M.; Nordborg, A.; Gaspar, A.; Lacher, N. A.; Wang, Q.; He, X. Z. P.; Haddad, P. R.; Hilder, E. F. J Chromatogr A 2013, 1317, 148-154.

[16] Mokaddem, M.; Gareil, P.; Varenne, A. Electrophoresis 2009, 30, 4040-4048.

[17] Alvarez, M.; Tremintin, G.; Wang, J.; Eng, M.; Kao, Y.-H.; Jeong, J.; Ling, V. T.; Borisov, O. V. Anal Biochem 2011, 419, 17-25.

[18] Haselberg, R.; Ratnayake, C. K.; de Jong, G. J.; Somsen, G. W. J Chromatogr A 2010, 1217, 7605-7611.

[19] Haselberg, R.; de Jong, G. J.; Somsen, G. W. Anal Chem 2013, 85, 2289-2296.

[20] Fekete, S.; Gassner, A. L.; Rudaz, S.; Schappler, J.; Guillarme, D. Trac-Trend Anal Chem 2013, 42, 74-83.

[21] Farnan, D.; Moreno, G. T. Anal Chem 2009, 81, 8846-8857.

[22] Rozhkova, A. J Chromatogr A 2009, 1216, 5989-5994.

[23] He, Y.; Lacher, N. A.; Hou, W.; Wang, Q.; Isele, C.; Starkey, J.; Ruesch, M. Anal Chem 2010, 82, 3222-3230.

[24] He, Y.; Isele, C.; Hou, W.; Ruesch, M. J Sep Sci 2011, 34, 548-555.

[25] Gassner, A. L.; Rudaz, S.; Schappler, J. Electrophoresis 2013, 34, 2718-2724.

[26] Shi, Y.; Li, Z.; Qiao, Y. B.; Lin, J. J Chromatogr B 2012, 906, 63-68.

[27] Espinosa-de la Garza, C. E.; Perdomo-Abundez, F. C.; Padilla-Calderon, J.; Uribe-Wiechers, J. M.; Perez, N. O.; Flores-Ortiz, L. F.; Medina-Rivero, E. Electrophoresis 2013, 34, 1133-1140.

[28] Redman, E. A.; Batz, N. G.; Mellors, J. S.; Ramsey, J. M. Anal Chem 2015, 87, 2264-2272.

[29] Amon, S.; Plematl, A.; Rizzi, A. Electrophoresis 2006, 27, 1209-1219.

[30] Silvertand, L. H. H.; Toraño, J. S.; de Jong, G. J.; van Bennekom, W. P. Electrophoresis 2009, 30, 1828-1835.

[31] Biacchi, M.; Bhajun, R.; Saïd, N.; Beck, A.; François, Y. N.; Leize-Wagner, E. Electrophoresis 2014, 35, 2986-2995.

[32] Gahoual, R.; Busnel, J. M.; Wolff, P.; Francois, Y. N.; Leize-Wagner, E. Anal Bioanal Chem 2014, 406, 1029-1038.

[33] Dubois, M.; Fenaille, F.; Clement, G.; Lechmann, M.; Tabet, J.-C.; Ezan, E.; Becher, F. Anal Chem 2008, 80, 1737-1745.

[34] Shen, Y. F.; Smith, R. D. J. Microcolumn Sep. 2000, 12, 135-141.

[35] Legaz, M. E.; Pedrosa, M. M. J Chromatogr A 1996, 719, 159-170.

[36] Verzola, B.; Gelfi, C.; Righetti, P. G. J Chromatogr A 2000, 868, 85-99.

[37] Punzet, M.; Ferreira, F.; Briza, P.; van Ree, R.; Malissa Jr, H.; Stutz, H. J Chromatogr B 2006, 839, 19-29.

[38] Smith, R. D.; Olivares, J. A.; Nguyen, N. T.; Udseth, H. R. Anal Chem 1988, 60, 436-441.

[39] Qian, J.; Liu, T.; Yang, L.; Daus, A.; Crowley, R.; Zhou, Q. W. Anal Biochem 2007, 364, 8-18.

[40] Janin-Bussat, M.-C.; Tonini, L.; Huillet, C.; Colas, O.; Klinguer-Hamour, C.; Corvaïa, N.; Beck, A. In Glycosylation Engineering of Biopharmaceuticals, Beck, A., Ed.; Humana Press, 2013, pp 93-113.

[41] Liu, H. C.; Gaza-Bulseco, G.; Faldu, D.; Chumsae, C.; Sun, J. J Pharm Sci 2008, 97, 2426-2447.

[42] Raju, T. S.; Scallon, B. J. Biochem Biophys Res Co 2006, 341, 797-803.

[43] Beck, A.; Wagner-Rousset, E.; Bussat, M. C.; Lokteff, M.; Klinguer-Hamour, C.; Haeuw, J. F.; Goetsch, L.; Wurch, T.; Van Dorsselaer, A.; Corvaia, N. Curr Pharm Biotechno 2008, 9, 482-501.

3. Conclusion

Lors de ce travail, nous avons développé une stratégie originale mettant en jeu le couplage indirect de la CZE et de la CESI MS, utilisée comme plateforme d'infusion nanoESI, pour la caractérisation de mAbs. Afin de réaliser la séparation des variants de charges des mAbs, les conditions électrophorétiques ont été optimisées. Dans un premier temps, dans le but d'éliminer les phénomènes d'adsorption et de réduire l'EOF, un greffage neutre HPC du capillaire a été réalisé. Ensuite, des conditions asymétriques de BGE basées sur la mise en place de différents BGE aux extrémités du capillaire ont été développées afin de conserver la résolution électrophorétique et de permettre une totale compatibilité avec une détection ESI-MS. Le BGE à l'entrée du capillaire composé de 200 mM d'acide ε-aminocaproïque et de 25 mM d'acétate d'ammonium pH 5,70, permet de conserver les performances de la séparation électrophorétique alors que le BGE à la sortie du capillaire composé de 25 mM d'acétate d'ammonium pH 5,70, permet la compatibilité avec la détection ESI-MS. Les résultats sur la séparation électrophorétique de variants de cetuximab a permis d'obtenir six pics correspondant à des masses moléculaires distinctes. Les trois premiers pics de la séparation correspondent à des variants de Fc/2 alors que les trois suivants correspondent à des variant de F(ab')2. Les performances observées de la séparation ont données une résolution allant jusqu'à 2,10 pour les variants de Fc/2 et 1,05 pour les variants de F(ab')2. Concernant le domaine Fc/2, deux variants de masse basée sur la perte d'une lysine en C-terminal de la HC ont été identifiés et séparés. De plus, cette séparation a permis d'observer le profil et la caractérisation de sept glycosylations porté par le fragment Fc/2. Concernant le domaine F(ab')2, la déconvolution des spectres de masse obtenue pour chaque pic a permis de déterminer la structure de 8 glycoformes directement impliqués dans la séparation. La séparation de ces glycoformes s'explique par la présence de résidus d'acide sialique de type NGNA présent dans différents sucres. En effet, quatre glycosylations du domaine F(ab')2 ont été identifiés par une différence de masse de 145 Da caractéristique d'un groupement NGNA. Cette séparation est donc due à la diminution de la mobilité électrophorétique des variants à cause de l'addition de la charge négative de l'acide sialique. Ce travail représente la première séparation et identification de glycoformes de mAbs à l'aide d'un couplage CZE-MS. Cela met en évidence le potentiel de la CZE pour la séparation de mAbs intacts principalement quand ceux-ci portent des groupements acides sialiques. De plus, cette stratégie de couplage indirect CZE/nanoESI-MS ouvre la voie à d'autres applications telles qu'une approche top down à l'aide d'une fragmentation ETD pouvant donner des informations complémentaires sur la structure des protéines.

II - Caractérisation de dimères de fragments $F_{c/2}$ de cetuximab par Off-line CZE-UV/ESI-MS

1. Introduction

Les résultats présentés dans cette partie sont directement liés à l'étude précédemment décrit dont l'intitulé est : I - Séparation et caractérisation de glycoformes de Cetuximab par Off-line CZE-UV/ESI-MS.

Depuis 1986 et l'approbation du murumonab-CD3 par l'administration de régulation des aliments et des médicaments américaine (FDA), les anticorps monoclonaux (mAbs) représentent la classe de molécule dont le développement est en constante croissance dans l'industrie pharmaceutique. Les mAbs sont des glycoprotéines complexes potentiellement porteuses de nombreuses micro-hétérogénéités moléculaires dont les sources sont variées (procédés de production inadaptée, conservation prolongée). Le domaine Fc des mAbs est une partie commune à de nombreux dérivés tels que les biosimilaires, les anticorps conjugués ou les protéines fusions. Cette région de la protéine est particulièrement responsable de fonctions effectrices telles que la cytotoxicité dépendante des anticorps. De plus, elle facilite la purification et peut également contribuer à améliorer la solubilité et la stabilité de composé. La présence d'une région Fc a des conséquences sur les fonctions effectrices de médiation qui pourrait être souhaitable pour des applications thérapeutiques.

L'un des principaux problèmes lors de la prise par le patient de ce type de médicaments, est la formation d'agrégats. En effet, les agrégats se formant dans des poches de produits thérapeutiques peuvent causer des réponses auto-immunes non désirées et/ou la perte d'activité biologique. Un des sites pouvant impliquer la formation d'agrégats est situé sur le domaine constant C_H2 de la région Fc. Ainsi, la complexité de ces protéines et leur utilisation en tant que traitement thérapeutique justifient la nécessité d'améliorer les méthodologies analytiques de caractérisation. Ceci de manière à obtenir une caractérisation rapide et précise de ce type de molécules. La spectrométrie de masse (MS) a su rapidement s'imposer comme une méthode de choix pour la caractérisation des mAbs, notamment en raison de son excellente spécificité, sensibilité mais aussi car elle permet d'obtenir des informations structurales dans certain cas. Comme il a été détaillée dans la partie précédente (I - Séparation et caractérisation de glycoformes de Cetuximab par Off-line CZE-UV/ESI-MS), nous avons développé la première séparation et identification de glycoformes de mAbs à l'aide d'un couplage CZE-MS. Cependant, aucune caractérisation de phénomènes d'agrégations n'a été mise en évidence dans cette étude.

Dans cette étude, nous avons cherché à comprendre la nature de la séparation des variants de charges correspondant au domaine Fc/2 obtenu dans le précédent chapitre. En effet, l'étude MS de la séparation électrophorétique montrant 3 pics de variant de Fc/2, a permis de caractériser deux variants de masse basés sur la perte d'une lysine en C-terminal de la HC du cetuximab. Cependant la présence d'un troisième pic correspondant à un mélange de ces deux variants de masse implique plusieurs questions sur l'origine de ce phénomène : diffusion lors du dépot, rôle des glycosylations rôle de la perte de lysine ou formation d'agrégats. Les résultats détaillés dans cette partie permettront donc d'apporter toutes les réponses à ces questions, notamment sur la formation de dimères de Fc/2.

Les résultats présentés dans cette partie ont fait l'objet d'une publication dans le journal scientifique *Analytica Chimica Acta (*Elsevier 2016).

2. Characterization of Cetuximab $F_{c/2}$ aggregates by Off-line CZE-MS

Yannis-Nicolas François, Michael Biacchi, Nassur Said, Charly Renard, Alain Beck, Rabah Gahoual, Emmanuelle Leize-Wagner, ***Analytical Chemistry*** 2016, 908, 168-176.

Abstract

Monoclonal antibody (mAb) therapeutics attract the largest concern due to their strong therapeutic potency and specificity. The Fc region of mAbs is common to many new biotherapeutics as biosimilar, antibody drug conjugate or fusion protein. Fc region has consequences for Fc-mediated effector functions that might be desirable for therapeutic applications. As a consequence, there is a continuous need for improvement of analytical methods to enable fast and accurate characterization of biotherapeutics. Capillary zone electrophoresis-Mass spectrometry couplings (CZE-MS) appear really attractive methods for the characterization of biological samples. In this report, we used CZE-MS systems developed in house and native MS infusion to allow precise middle-up characterization of Fc/2 variant of cetuximab. Molecular weights were measured for three Fc/2 charge variants detected in the CZE separation of cetuximab subunits. Two Fc/2 C-terminal lysine variants were identified and separated. As the aim is to understand the presence of three peaks in the CZE separation for two Fc/2 subunits, we developed a strategy using CZE-UV/MALDI-MS and CZE-UV/ESI-MS to evaluate the role of N-glycosylation and C-terminal lysine truncation on the CZE separation. The chemical structure of N-glycosylation expressed on the Fc region of cetuximab does not influence CZE separation while C-terminal lysine is significantly influencing separation. In addition, native MS infusion demonstrated the characterization of Fc/2 dimers at pH 5.7 and 6.8 and the first separation of these aggregates using CZE-MS.

Introduction

Although monoclonal antibodies (mAbs) were introduced as treatments against disease in the late 1980, they currently represent the most rapidly growing category of therapeutic molecule [1-3]. MAbs are highly complex glycoproteins potentially displaying many naturally-occurring molecular micro-heterogeneities [4, 5]. The Fc region of mAbs is common to many new platforms as biosimilar, antibody drug conjugate or fusion protein. This region of the protein is conditioning the antibody isotype and particularly is responsible for effector functions such as antibody-dependent cytotoxicity (ADCC), complement-dependent cytotoxicity (CDC) like phagocytosis [6]. Additionnally it facilitates purification and can also contribute to improved solubility and stability [7]. The presence of an Fc region has consequences for Fc-mediated effector functions that might be desirable for therapeutic applications. For successful delivery of biotherapeutic to patient, one of the major issue is aggregates formation. Aggregates in formulated therapeutic products can cause unwanted immune responses and loss of biological activity [8, 9]. Literature suggests that, depending on the nature of mAbs, mAbs aggregation in human might be mediated primarily by unfolding of CH2 domain on the Fc region [10, 11] and Fab domain [10, 12]. As a consequence, there is a continuous need for improvement of analytical methods to enable fast and accurate characterization. Mass spectrometry (MS) generally coupled with separation methods plays a key role in the characterization of therapeutic mAbs [13] . Depending on the protein structure level considered, characterization is performed following different strategies as intact molecular weight (MW) measurement, top-down, middle-up and bottom-up approaches [13-17]. Capillary zone electrophoresis (CZE) has been demonstrated to be an useful and powerful separation method for the characterization of intact proteins [18-20]. In a previous work, we developed the first analysis of intact mAb charge variant by CZE using a matrix-assisted laser desorption/ionization-MS (MALDI) detection [21]. In 2015, Redman et al published a remarkable paper on characterization of intact mAb variants using microfluidic CZE-ESI devices [22]. However, every studies described mAbs characterization without any aggregation phenomenon occurring. Recently, we developed a method for the middle-up characterization of cetuximab glycoforms using off-line CZE-UV/ESI-MS [23].

In this report, we used CZE-MS systems developed in house and native MS infusion to allow precise middle-up characterization of Fc/2 variant. The mAb selected was cetuximab which is human/murine chimeric IgG-1 directed against the epidermal growth factor receptor (EGFR) overexpressed in advanced-stage EGFR positive colorectal cancer. Cetuximab was approved in the US and EU in 2004 and 2005. Cetuximab contains two glycosylation sites both on the heavy chain (HC) [13]. Moreover, it has a large number of micro-heterogeneities such as PTMs including methionine oxidation, asparagine deamidation or isomerization of aspartic acid. This

mAb also has one C-terminal lysine truncation. These cumulated features make cetuximab an ideal sample for the characterization of mAbs variants by a middle-up approach. In order to cleave the protein in relatively large fragments and distribute N-glycosylation sites over different fragments generated, IdeS proteolysis was employed. IdeS is a cysteine endopeptidase enzyme naturally secreted by Streptococcus pyogenes [24], IdeS enzymatic reaction cleaves cetuximab between the two consecutive glycine residues present in the hinge region to obtain two types of fragments (Fc/2 and F(ab')2) each one carrying a N-glycosylation site. IdeS has demonstrated an exceptional specificity in IgG proteolysis on contrary to papain for instance, therefore it has been consequently used to ease and improve mAbs characterization [25]. IdeS characteristics regarding specificity alongside to proteolysis kinetics are explaining the success of this enzyme in mAbs characterization [26]. We developed a strategy using CZE-UV/MALDI-MS and CZE-UV/ESI-MS to evaluate the role of N-glycosylation and C-terminal lysine truncation on the CZE separation. In addition, this approach demonstrated the characterization of Fc/2 dimers at pH 5.7 and 6.8 and the first separation of these aggregates using CZE-UV off-line hyphenated to ESI-MS after Fraction collection. A better understanding of these phenomenon could improve the characterization of Fc region and their potential non-covalent interactions [27].

Materials and methods

Chemicals. Methanol (HPLC gradient grade) and acetic acid (100%) were obtained from VWR (Radnor, PA, USA). Ammonium acetate (>98%), sodium hydroxide, ε-Amino-caproic acid (>98%), hydroxypropylcellulose Mw 100000 (HPC), carboxypeptidase B (CP-B) and formic acid (>98%) were purchased from Sigma-Aldrich (Saint Louis, MO, USA). Water used to prepare buffers and sample solutions was obtained using an ELGA purelab UHQ PS water purification system (Bucks, UK). IdeS (Immunoglobulin-degrading enzyme of Streptococcus pyogenes) also named FabRICATOR and IgGZERO were purchased from Genovis (Lund, Sweden). Cetuximab (Erbitux®, Merck KGaA, Darmstadt, Germany) is a sterile, preservative-free solution for intravenous infusion containing 5 mg/mL of cetuximab. Additionally to the mAbs it contains sodium chloride, glycine, polysorbate 80, citric acid monohydrate, sodium hydroxide and water for injections.

Middle-up sample preparation. Cetuximab was cleaved in the hinge region using limited proteolysis by IdeS (FabriCATOR, Genovis) to obtain two Fc/2 fragments (theoretical pI 7.74) and one F(ab')2 fragment (theoretical pI 7.78). Sample was diluted using 147.25µL of 50 mM sodium phosphate, 150 mM NaCl, pH 6.60, to a final concentration of 1 µg/µL. A volume of 2.25 µL of IdeS (67 units/µL) was added to the sample which was left at 37°C for 30 min. After

digestion completion, sample was desalted using Amicon centrifugal filters (cut off = 10,000 Da) in pure water at 10°C and 14,000 g for 20 min. After desalting step, sample volume recovered was about 10 µL. Sample was finally diluted to a final concentration of 5 µg/µL in a total volume of 30 µL of pure water.

Carboxypeptidase B (CP-B) treatment. Cetuximab was diluted using 145.75 µL with 10 mM phosphate buffer, 150 mM NaCl, pH 7.4, to a concentration of 1 µg/L. Volumes of 2 µL of CP-B (1mg/mL) and 2.25 µL of IdeS (67 units/µL) were added to the sample which was left at 37°C for 30 min. After digestion completion, sample was treated as described in the middle-up sample preparation.

Deglycosylation protocol. Cetuximab was diluted using 141.75 µL with 10 mM phosphate buffer, 150 mM NaCL, pH 7.4, to a concentration of 1 µg/µL. Volumes of 6 µL of IgGZERO (20 units/µL) and 2.25 µL of IdeS (67 units/µL) were added to the sample which was left at 37°C for 30 min. After digestion completion, sample was treated as described in the middle-up sample preparation.

Capillary electrophoresis. The CZE experiments were carried out on a P/ACE MDQTM CE system from Sciex Separation (Brea, CA) equipped with a UV detection, a temperature controlled autosampler and a power supply able to deliver up to 30 kV. A 32 KaratTM 8.0 (Sciex Separation, Brea, CA) was used for instrument control, data acquisition and data handling. Polymicro bare fused-silica capillaries of 75 µm i.d., 375 o.d. (75.5 cm effective length, 82 cm total length) were obtained from Photonlines (St-Germain-en-Laye, France). New capillaries were conditioned by successive flushes with 1M and 0.1M NaOH and then with water under a pressure of 30 psi for 10 min each. The temperature in the capillary cartridge and autosampler was set to 25 °C. The acquisition rate was 10 points/s. Capillaries were rinsed with water and dried by air when not in use. UV absorbance remained fixed at 200 nm. Voltage applied at 20 kV with a ramp duration of 0.17 min and injection sample conditions were 0.5 psi for 50s. Concerning modified capillaries, capillaries were coated in house with hydroxypropylcellulose Mw 100000 (HPC) following the protocol described by Shen et al [28]. 5% HPC in pure water (w/v) was prepared to apply capillary coating. Coating durability is around 20 runs without recoating step. For cetuximab separation, inlet BGE 200 mM ε-Amino-caproic acid (EACA)-ammonium acetate 25 mM pH 5.70 and outlet BGE ammonium acetate 25 mM pH 5.70 have been used as separation conditions. Injection volumes have been calculated using CEToolbox application (Pansanel, GooglePlay).

CZE/Fraction Collection Interface. This interface have been described in a previous study [21]. Briefly, automated off-line coupling of CZE to MS was set up by using a homemade modified automatic spotting device Proteineer FC (Bruker Daltonics, Bremen, Germany) for

the sheath flow-assisted spotting from the CZE capillary outlet onto a fraction collection target. The original set up of the UV cell in the P/ACE MDQTM (Sciex Separation, Brea, CA) was delocalized outside the CE instrument in order to allow simultaneous UV detection and fractions collection. Hystar 3.2 (Bruker Daltonics, Bremen, Germany) was used for Proteineer FC control.

CESI-MS as nanoESI infusion platform. The infusion experiments were carried out with a PA 800 plus capillary electrophoresis (CE) system from Sciex Separation (Brea, CA). Prototype fused-silica capillaries (total length 95 cm; 30 µm i.d.) whose outlet end (about 3 cm) was etched with hydrofluoric acid were provided by Sciex Separation (Brea, CA, USA). New capillaries were initially conditioned by flushing them for 10 min with MeOH, 10 min with 0.10 M sodium hydroxide, 10 min with 0.10 M hydrochloric acid and finally with water for 20 min each flushing step being conducted at 75 psi (5.17 bar).

MALDI-TOF-MS. The matrix was prepared by dissolving 2,5-dihydroxybenzoic acid (DHB) (2g/L) in 0.1% trifluoroacetic acid/acetonitrile (TFA/ACN) (30/70)(v/v). Fraction collection was realized using Ground Steel MALDI target (Bruker Daltonics, Bremen, Germany). Mass spectra of the CZE fractions were recorded using an Autoflex II MALDI-TOF (Bruker Daltonics, Bremen, Germany), operating in reflector mode and with FlexControl software. Positively charged ions were detected and sums of 1500 single-shot spectra were acquired automatically from each sample by using the AutoXecute software. Data processing was performed with FlexAnalysis 3.0 (Bruker Daltonics, Bremen, Germany). All spectra were calibrated by external calibration using Protein calibration standard I (Bruker Daltonics, Bremen, Germany) for intact protein separation.

ESI-TOF-MS. For nanoESI-MS infusion experiments the PA800 plus system was coupled to a maxis 4G (Bruker Daltonics, Bremen, Germany). MS transfer parameters were optimized using the actual sample directly infused via the CE system using a pressure of 5 psi (340 mbar). MS parameters were optimized so that high m/z ions could be properly transferred to the TOF mass analyzer while avoiding fragmentation. Ion funnels RF were set at values of 300 Vpp and 400 Vpp. The electrospray voltage (capillary voltage) was typically ranging from -1.2 to -1.8 kV. Dry gas was set at 1.5 L/min and source temperature at 180°C. Data processing was performed with DataAnalysis 4.0 (Bruker Daltonics, Bremen, Germany). Deconvolution of the mass spectra was performed based on maximum entropy analysis using ESI Compass 1.3 Maximum Entropy Deconvolution Option in DataAnalysis 4.0. All spectra were calibrated by external calibration using Pepmix (Bruker Daltonics, Bremen, Germany) and cesium iodide (CsI) from Sigma-Aldrich (Saint-Louis, MO, USA).

Results and discussions

Cetuximab is a chimeric mouse-human IgG1 known to bear 2 N-glycosylation sites on each HC [26]. As every mAbs, cetuximab can exhibit a wide variety of micro-heterogeneities particularly due to PTMs. Furthermore, this mAb can also experience HC C-terminal lysine truncation. In a preliminary experiment, direct nanoESI-MS infusion of cetuximab IdeS digest was performed. Results proved equivalent signal abundances for Fc/2 C-terminal lysine variants which involves two distinct profiles for the Fc/2 domains (Fig. S-1). With this information, we expected to detect at least two peaks corresponding to Fc/2 fragments due to a loss of a +1 charge associated with C-terminal lysine truncation.

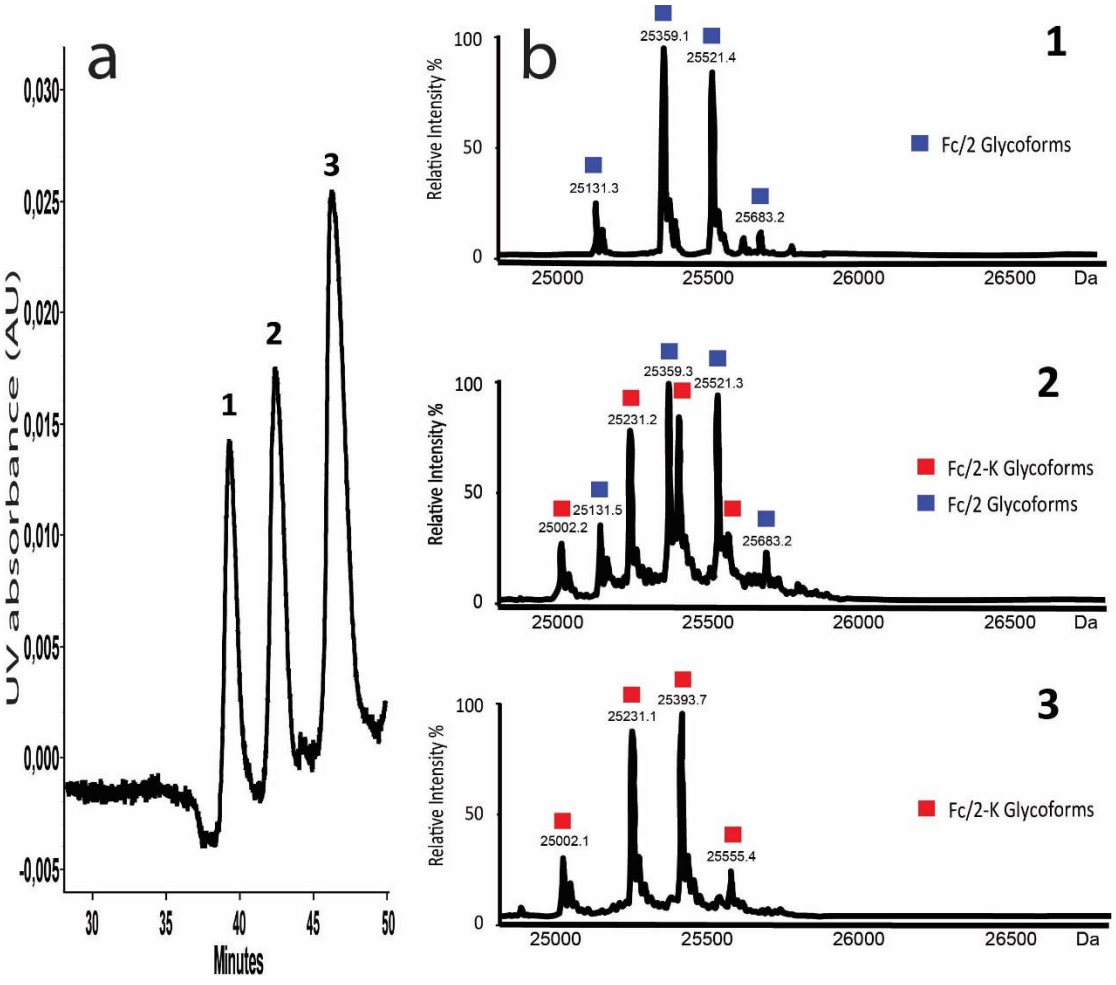

Figure 1. (a) Off line CZE-UV/ESI-MS separation of middle-up cetuximab charge variants. (b) Deconvoluted mass spectra for each Fc/2 variant. The MS peak were labeled with the correspondent glycoform. Experimental conditions: Inlet BGE: EACA 200 mM ammonium acetate 25 mM pH 5.70 and outlet BGE: ammonium acetate 25 mM pH 5.70; HPC-coated capillary, total/effective length 82/75.5 cm x 75 µm i.d.; voltage, 20 kV; UV absorbance at 200 nm; MS condition: See Experimental section; sample, IdeS digest of cetuximab (5 µg/µL); sample injection 0.5 psi 50s.

Cetuximab IdeS digest was separated and characterized by offline CZE-UV/ESI-MS using an instrumental setting developed in house, described in a previous work [23]. Briefly, assymetric

CZE conditions was based on inlet background electrolyte (BGE) composed of ε-amino-caproic acid (EACA) 200 mM/ammonium acetate 25 mM, pH 5.7 and outlet BGE composed of ammonium acetate 25 mM, pH 5.7. Hydroxypropylcellulose (HPC) coating was used to minimize analyte adsorption on the capillary wall and reduce the electroosmotic flow (EOF). After deposition, each fraction was collected and evaporated. Before infusion, dry samples were reconstituted in 2µL of ACN/H20/FA (50/49/1, v/v) in order to enhance sample ionization. Thus, each fraction was infused into the MS using a flow rate of 10 nL/min in order to enhance sensitivity.

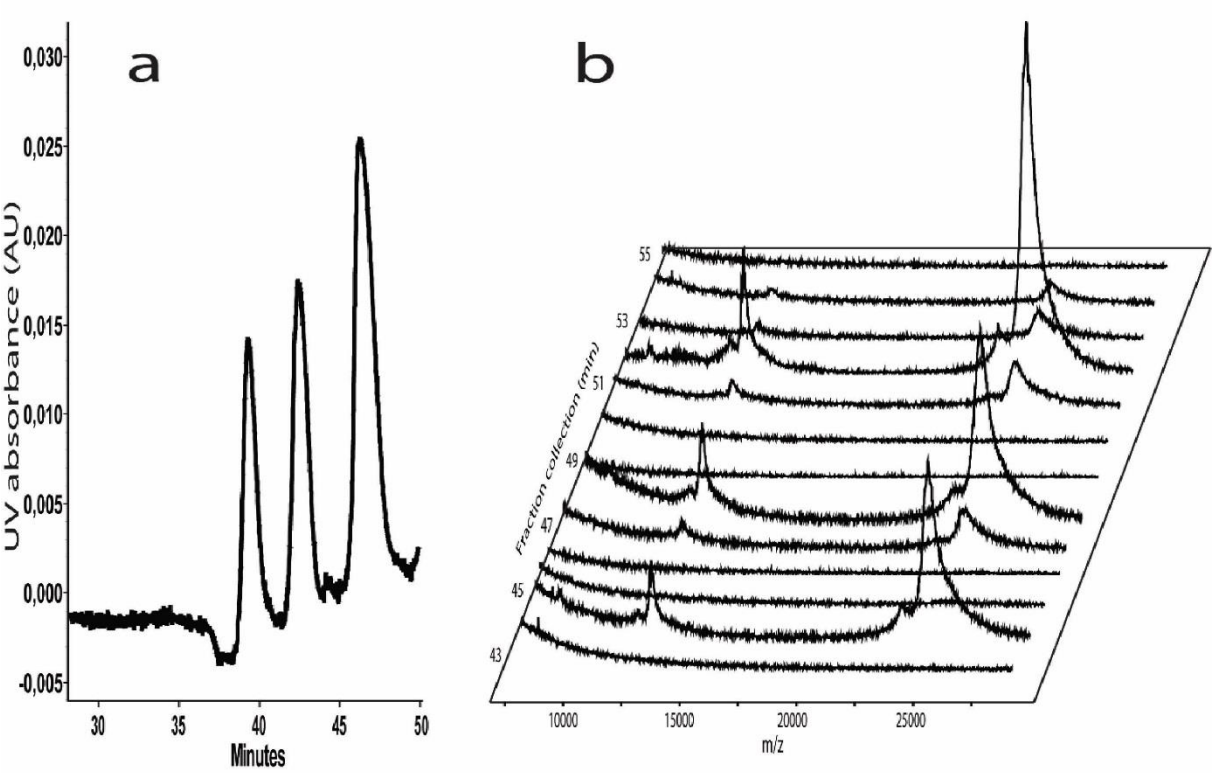

Figure 2. CZE-UV off-line coupled to MALDI-MS via fractionation for a middle-up characterization of Cetuximab $F_{c/2}$ variants. (a) CZE-UV electropherogram. (b) Analysis of CZE-UV fractions by MALDI-MS. CE Experimental conditions: See Fig. 1; MS condition: See Experimental section; sample, IdeS digest of cetuximab (5 µg/µL)

As emphasized in Fig. 1-b, CZE separation obtained shows three consecutive peaks completely separated corresponding to the Fc/2 region. Deconvoluted mass spectra of each fraction generated multiple masses corresponding to the expected presence of different neutral glycosylation on Fc/2 fragments (Fig. 1-b). In addition, the comparison with the direct infusion experiment (Fig. S-1) demonstrates that the two Fc/2 fragments corresponding to intact and with one lysine truncation (Fc/2 and Fc/2-K) are detected and separated. Deconvoluted mass spectra corresponding to peak 1 and peak 3 show a difference of 128 Da between mass peak 25359.1 Da corresponding to the most abundant Fc/2 glycoform (G0F) and mass peak 25231.1 Da corresponding to G0F for Fc/2-K. Thus, we demonstrate the complete separation of these

two fragments (Fig. 1-b). Regarding deconvoluted mass spectra of peak 2 (Fig. 1-b 2), it shows a nearly equal abundance mixture of Fc/2 and Fc/2-K variants. Deconvoluted mass spectra corresponding to peak 2 lead to the same MWs as peak 1 and peak 3 and could be attributed to either an overlap during the fraction collection process or by cumulative PTMs which could lead to a variation of conformation without mass shifts. Moreover, these phenomenon could also be explained by a possible aggregation of Fc/2 fragments.

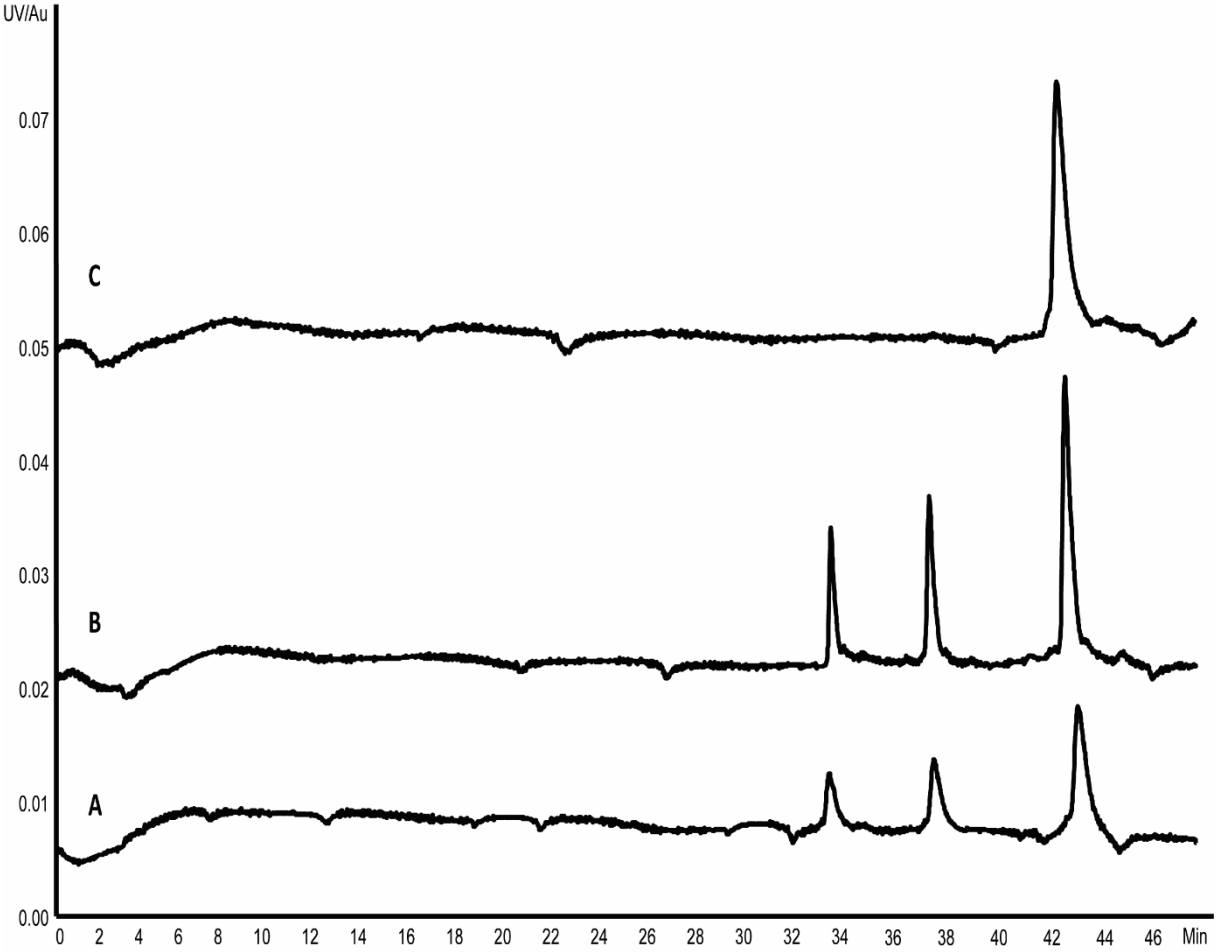

Figure 3. Impact of glycosylation and lysine truncation on electrophoretic resolution. Electrophoretic separation of (A) cetuximab with IdeS treatment, (B) cetuximab with IgGZERO+IdeS treatment and (C) cetuximab with CP-B+IdeS treatment. Inlet BGE composed by a mixture of EACA 200 mM and ammonium acetate 25 mM pH 5.70 and outlet BGE by ammonium acetate 25 mM at pH 5.70.

Validation of the CZE-UV/ESI-MS fraction collection process.

The fraction collection process developed in this study based on the apparent mobility of each peak, allows theoretically to recover the total quantity of each variant present in the peak. However the offline coupling strategy developed do not permit to represent the separation using MS spectra and then to observe possible diffusion. Regarding raw MS spectra (Fig. S-2) and based on the hypothesis of an equivalent ionization efficiency of all Fc/2 variants, potential carryover effect due to peak broadening or excessive diffusion should be eliminated.

However, in the aim to confirm the absence of overlap during the fraction collection, experiments using CZE-UV/MALDI-MS coupling have been performed. Same experimental conditions described previously have been implemented. Only deposition process has been changed to allow a fraction collection directly on a MALDI target plate with a deposition time interval of 1 minute between each fraction. Fraction collection was performed between 43 and 55 minutes according to the electropherogram represented in Fig. 2-a. Fig. 2-b emphasizes the reconstruction of the separation using MS spectra of each fraction. First of all, the CZE/MALDI-MS reconstruction exhibits three distinct peaks corresponding to the CZE separation. MS spectra of each peak in CZE show a value of 25 kDa corresponding to the expected Fc/2 variants. However due to presence of other sources of micro-heterogeneities and the limited resolution of the MALDI-MS instrument used, it was impossible to deduce the average mass of the charge variants confidently. In the other hand, the CZE/MALDI-MS profile (Fig. 2-b) confirms the good agreement between the UV detection time and the deposition time. Based on the difference of protein apparent mobilities described by the equation:

$$\mu_{app} = \frac{\mathrm{L\,l}}{\mathrm{t}_m\,\mathrm{V}}$$

with μ_{app} the apparent mobility which is the sum of effective mobility and residual electroosmotic flow, L and l the total capillary length and length to detection window, respectively, V the applied voltage and t_m the migration time. Deposition time t_d is then calculated by the equation:

$$t_d = \frac{L^2}{\mu_{app} V}$$

The localization of each Fc/2 variants shows a time shift between the UV detections and the spotting process due to UV localization at 6.5 cm from the capillary outlet. These results confirm without any remaining ambiguity the absence of carryover effect and diffusion phenomenon during the fraction collection process, demonstrating the accuracy and robustness of the instrumental setting developed. As a consequence, characteristics of the analytical strategy developed to characterize Fc/2 fragments cannot be doubted. Therefore it is necessary to focus on cetuximab structure and potential modifications on the fragments generated during IdeS proteolysis in order to conclude on the CZE separation described previously (Fig. 1-a).

Influence of N-glycosylation on Fc/2 variants separation

An important class of modification that can significantly impact the immunogenic properties of cetuximab in terms of pharmacokinetic and pharmacodynamic is glycosylation [26]. Cetuximab

contains two sites of glycosylation on the HC: one is located in the Fc/2 domain (Asn299) and the second is located in the F(ab')2 domain on Asn88 [13]. Cetuximab glycosylation profile has been extensively described in particular as it is expressing glycoforms responsible for immunogenic reaction [26]. Regarding Fc/2 domain, branched glycan structures are primarily composed of fucose, N-acetylglucosamine, and mannose but can differ in the number of terminal galactose, mannose or N-acetylglucosamine residues. Glycan structure with the addition of galactose moieties involves a mass increase of 162 Da but does not induce a change in net charge [29, 30]. In a previous work we demonstrated that concerning glycoform with a difference of one galactose residue, middle-up approach does not allow to observe baseline separations [23]. The mass of galactose moiety accounts for only ≈0.6% of the Fc/2 domain and ≈0.2% of the F(ab')2 domain. Thus mass difference should be greater or involve a significant conformation change to enable a significant effective mobility shift. To confirm these theoretical considerations, analysis of deglycosylated sample of cetuximab by CZE-UV/ESI-MS was performed in order to evaluate the role of glycans in the separation process. Before IdeS digestion of cetuximab, a deglycosylation step using IgGZERO has been added to the digestion protocol. The aim is to remove all glycans in the Asn299. IgGZERO is an endoglycosidase specific for N-glycans on the Fc domain of IgGs which are cleaved after the first Glc-NAc. Following IgGZERO and IdeS treatment, the most intense ion series is corresponding to expected cetuximab Fc/2 and Fc/2-K subunits with GlcNac-Fuc remaining (Fig. 4-b and Table S-1). CZE separation, represented by Fig. 3-B, contains the three peaks corresponding to Fc/2 subunits with the same profile in terms of migration times and resolution compared to the original separation of the sample non treated with IgGZERO enzyme (Fig. 3-A). With these results, we can conclude that glycan nature expressed on Asn299 residue does not impact significantly the CZE separation of Fc/2 variants. Effect of others microheterogeneities like asparagine deamidation or methionine oxidation are quite difficult to observe on the intact protein or using limited proteolytic digestion because of the number of these modifications, their localization and the difficulty to control the level of modification in the proteins. Indeed, some parameters like storage conditions or digestion pH can induce endogenous asparagine deamidation or methionine oxidation.

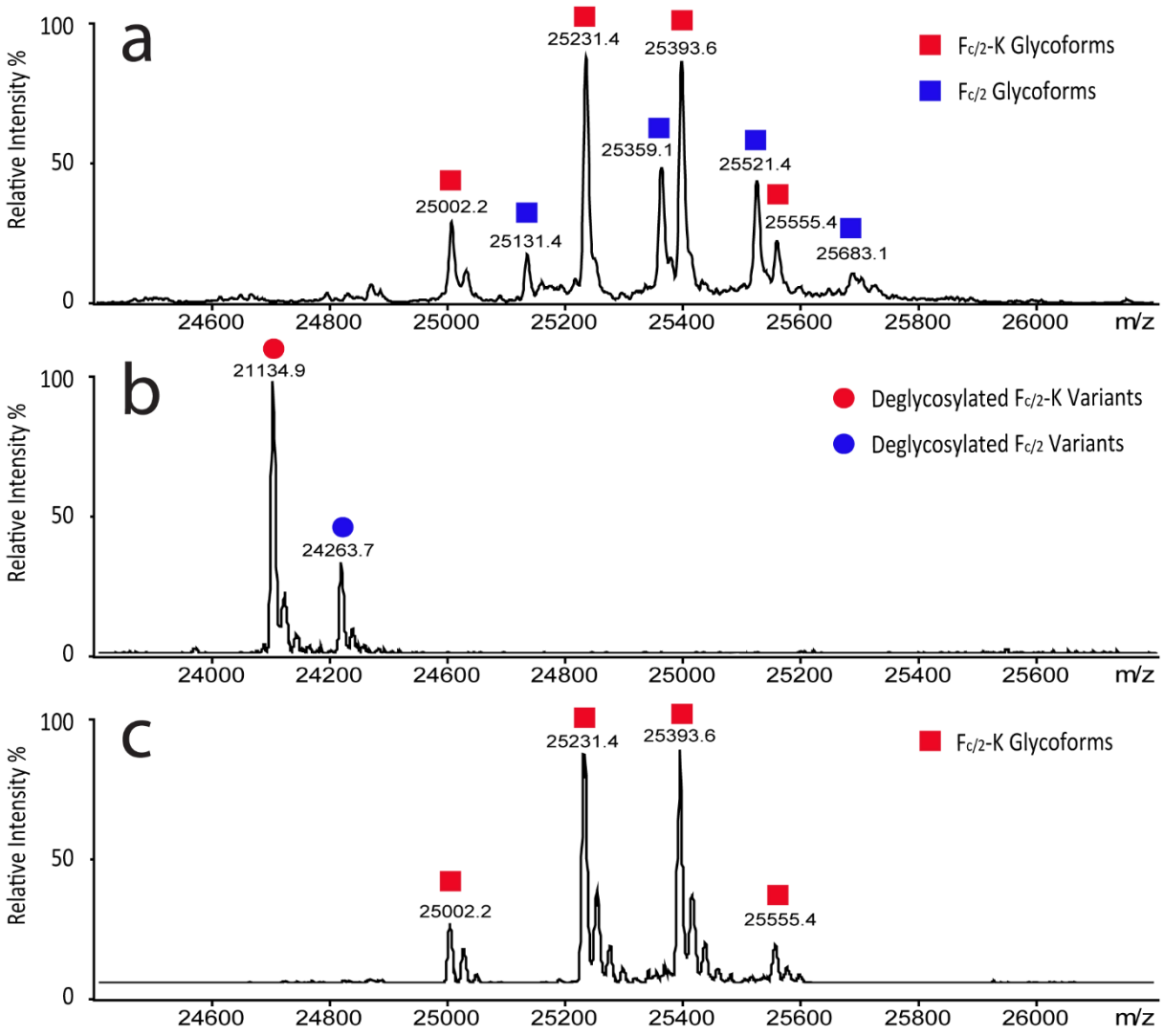

Figure 4. $F_{c/2}$ variants deconvoluted mass spectra of (a) IdeS digest cetuximab, (b) IgGZERO+IdeS digest cetuximab and (c) CP-B+IdeS digest cetuximab. Experimental conditions: 50%/49%/1%, acetonitrile/water/formic acid (v/v/v); sample concentration 1 µM.

Influence of C-terminal lysine truncation on Fc/2 variants separation

The presence of the lysine in HC C-terminal position is essential to understand the separation of Fc/2 variants. Indeed, in this work, we demonstrated the separation of the two Fc/2 fragments corresponding to intact and with one lysine truncation (Fc/2 and Fc/2-K). Peak 1 and peak 3 (Fig. 3-A) correspond to the most abundant Fc/2 glycoforms and Fc/2-K glycoforms respectively. However for peak 2, a nearly equal abundance mixture of Fc/2 and Fc/2-K have been observed. To evaluate the role of lysine truncation in the separation process, IdeS digestion protocol has been modified to perform a total truncation of C-terminal lysine of cetuximab. Sample was treated by carboxypeptidase B (CP-B), removing systematically the C-terminal lysine residues from the two HCs of the antibody [31], it was added prior to IdeS

digestion step. After CP-B and IdeS treatment, the most intense ion series corresponds to the expected glycoforms of Fc/2-K variants (Fig. 4-c). Moreover, as compared to sample treated only with IdeS (Fig. 4-a), we observe in the deconvoluted mass spectrum the complete loss of Fc/2 glycoforms profile which confirms the total removing of C-terminal lysine on the HC.

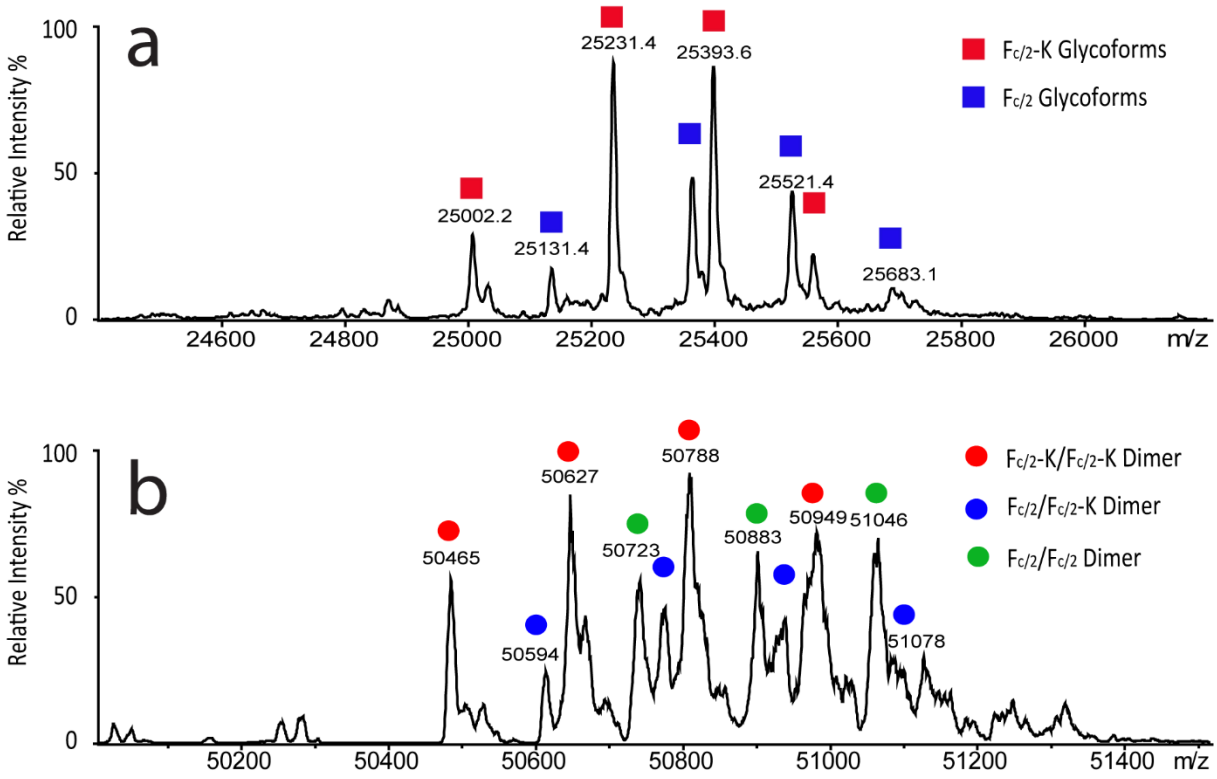

Figure 5. Cetuximab $F_{c/2}$ variants deconvoluted mass spectra in (a) denaturing and (b) CZE conditions. Experimental conditions: denaturing conditions, 50%/49%/1%, acetonitrile/water/formic acid (v/v/v); CZE conditions, ammonium acetate 25mM, pH 5.7; sample concentration 1 µM.

CZE separation, represented on Fig. 3-C, contains only one peak corresponding to Fc/2 subunits. This peak has been characterized as Fc/2-K variants. Moreover, we observe the absence of the two first peaks compared to the original separation. Despite the fact that the loss of the first peak is totally justified by the unique presence of Fc/2 variant trimmed from both HC C-terminal lysine, the loss of the second peak proves that the presence of C-terminal lysine residue plays an active role in the separation process of Fc/2 variants. These results also allow to reject hypothetic effects on effective mobilities due to several cumulative PTMs leading to modification of the mAb without MS resolved mass shift. The second peak of the separation of sample without CP-B treatment (Fig. 3A) has been characterized by a nearly equal abundance mixture of Fc/2 and Fc/2-K variants. Absence of this peak in cetuximab separation after CP-B+IdeS digestion (Fig. 3-C) and the fact that no Fc/2-K variants have been detected at that migration time involves a strong relation between Fc/2 and Fc/2-K variants.

Role of Fc/2 aggregates on CZE separation

In this work, we described a protocol based on nanoESI-MS infusion of the different fractions using CESI-MS interface as nanoESI platform. After CZE separation, each collected fraction was systematically dried by air and reconstituted in 2µL of ACN/H20/FA (50/49/1, v/v) to enhance ionization. However, infusion conditions differed from CZE conditions at pH 5.7 and are denaturing by the presence of acetonitrile. Indeed, in these experiments, noncovalent interactions between Fc/2 variants can influence CZE separation at pH 5.7 but protein-protein interaction cannot actually be observed due to the denaturing nanoESI-MS conditions used. To evaluate these phenomena, nanoESI-MS infusion of cetuximab IdeS digest have been prepared using denaturing conditions and using CZE conditions as well. For denaturing conditions, sample has been diluted in 2µL as stated previously while for CZE conditions, sample has been diluted in ammonium acetate 25mM, pH 5.7. Each sample has been reconstituted to a concentration of 1µM. Each sample was infused into the MS using a flow rate of 10 nL/min. Deconvoluted mass spectra in denaturing conditions generated multiple masses corresponding to the expected presence of neutral glycosylation variants of Fc/2 fragments (Fig. 5-a and Table S-1). Mass peak 25359.2 Da corresponding to G0F for Fc/2 and mass peak 25231.1 Da corresponding to G0F for Fc/2-K are in agreement with theoretical masses of these glycoforms (Table S-1). Concerning CZE conditions, charge variant deconvoluted mass spectrum generated multiple masses between 50.4 kDa and 51.2 kDa (Fig. 5-b). These peaks correspond to the formation of protein-protein aggregates due to noncovalent interactions between Fc/2 subunits. Indeed, mass peak 50463.4 Da corresponds to the mass of Fc/2-K subunits homodimer, mass peak 50591.8 Da corresponds to the heterodimer between Fc/2 and Fc/2-K subunits and mass peak 50721.2 Da corresponds to Fc/2 subunit homodimer (Table S-2). Moreover, other mass peaks between 50.4 kDa and 51.2 kDa correspond to the expected presence of the combination of neutral glycosylation variants of these different dimers (Table S-2). In addition, no deconvoluted mass spectrum have been observed at 25 kDa that excludes the presence of free Fc/2 variants in ammonium acetate 25mM at pH 5.7. These results allow the explanation of the presence of three different peaks corresponding to Fc/2 variants in CZE separation. Indeed, peak 1 (Fig. 1-a) corresponds to Fc/2 homodimer, peak 2 corresponds to Fc/2/Fc/2-K heterodimer and peak 3 correspond to Fc/2-K homodimer. To our knowledge, this is the first time that aggregates of cetuximab Fc/2 subunits have been separated and characterized. Finally, a last study consisted to infuse the same sample in condition classically described in the literature as native conditions [32-35]. To be in agreement with CZE condition, 25 mM ammonium acetate at pH 6.8 was selected. Obtained MS profiles were exactly the same as those observed using CZE conditions. This confirms the presence of Fc/2 dimers in native conditions and exclude a possible major role of

EACA in the protein-protein aggregation process. Results obtained from CZE-UV/MALDI-MS and CZE-UV/ESI-MS experiments clearly demonstrate the formation of Fc/2 fragments following IdeS proteolytic digestion of mAbs. This observation is particularly important from an analytical chemistry point of view. As the outstanding selectivity of IdeS has favored its wide use among the scientific community and the formation of aggregates, even if reversible, it should be taken into account in the development of innovative analytical methods.

Conclusions

To summarize, we report here the use of off-line CZE-UV/ESI-MS for the middle-up characterization of Fc/2 variant of cetuximab. To validate the method, developed CZE offline instrumental setting was implemented to perform direct deposition of collected fractions on a MALDI target and used to characterize IdeS proteolytic digest fragments by MALDI-TOF MS. Results obtained demonstrated successful CZE-UV/MALDI-MS characterization. MS spectra were in complete agreement with the separation observed by the intermediate of the UV detection. Additionally data generated allowed to exclude overlapping during the collection and deposition process validating the instrumental approach regarding accuracy and robustness. In order to investigate further the separation of three different species corresponding to Fc/2 variants, CZE-UV/ESI-MS experiments were performed on cetuximab treated consequently with igGZERO and IdeS enzymes to permit the removal of N-glycosylation present in the Fc region of the protein. Results showed that chemical structure of N-glycosylation expressed on the Fc region of cetuximab does not influence CZE separation as both revealed similar UV electropherograms. A similar analysis was then realized on cetuximab treated successively with CP-B and IdeS in order to systematically remove C-terminal lysine from the Fc/2 fragments generated. Results from the CZE separation exhibited the absence of two peaks compared to the same separation using cetuximab solely treated with IdeS. As expected from the CZE conditions employed, C-terminal lysine is significantly influencing separation. Finally direct nanoESI-MS infusion of the collected fractions from CZE separation were performed using different buffering systems, one which is similar to CZE background electrolyte composed of ammonium acetate 25 mM, pH 5.70, and in parallel in denaturing conditions favoring ionization efficiency and signal sensitivity. In similar conditions as CZE separation, MS allowed to identify without ambiguity the nature of each analyte separated by CZE. MS spectra obtained revealed the first specie to be Fc/2 homodimer while the second specie detected was Fc/2/Fc/2-K heterodimer and the last one being Fc/2-K homodimer. Results demonstrated to be in agreement regarding the influence of C-terminal lysine on the separation but additionally showed the formation of Fc/2 aggregates as a consequence of IdeS proteolytic

digestion. Results from the different experiments demonstrated the relevance of this CZE-UV/ESI-MS analytical method developed to perform the comprehensive middle-up characterization of mAbs. Indeed it was possible to conclude of the formation of Fc/2 aggregates generated consequently from cetuximab IdeS treatment. As this enzyme recently tends to be widely used, due its exceptional specificity, for mAbs characterization, aggregation related to sample treatment appears as a new insight which should be taken into account during method development. The implementation of an electrophoretic separation prior to the MS analysis showed its interests in this work as separation efficiency provided by CZE allowed to completely separate the different types of aggregates while showing the possibility to maintain these aggregates. Moreover, the Fc region of mAbs is common to many new platforms and these results could improve the characterization of Fc region of biotherapeutics and their potential non-covalent interactions.

Acknowledgments

Authors would like to thank Sciex separations Inc. for lending a CESI8000 system and Dr. M. Anselme from Sciex for his support. The authors would like also to express their gratitude to Dr. E. Wagner-Rousset, MC. Janin-Bussat and O. Colas (Centre d'Immunologie Pierre Fabre, St Julien en Genevois, France) for helpful discussions around antibody structural characterization.

Reference

[1] A. Beck, T. Wurch, C. Bailly, N. Corvaia, Strategies and challenges for the next generation of therapeutic antibodies, Nat. Rev. Immunol. 10 (2010) 345-352.

[2] J.M. Reichert, Marketed therapeutic antibodies compendium, mAbs 4 (2012) 413-415.

[3] A. Beck, H. Diemer, D. Ayoub, F. Debaene, E. Wagner-Rousset, C. Carapito, A. Van Dorsselaer, S. Sanglier-Cianferani, Analytical characterization of biosimilar antibodies and Fc-fusion proteins, Trac-Trends Anal. Chem. 48 (2013) 81-95.

[4] J.N. Arnold, M.R. Wormald, R.B. Sim, P.M. Rudd, R.A. Dwek, The impact of glycosylation on the biological function and structure of human immunoglobulins, Annu. Rev. Immunol. 25 (2007) 21-50.

[5] Z.Q. Zhang, H. Pan, X.Y. Chen, Mass spectrometry for structural characteization of therapeutic antibodies, Mass Spectrom. Rev. 28 (2009) 147-176.

[6] J.V. Ravetch, J.P. Kinet, Fc Receptors, Annu. Rev. Immunol. 9 (1991) 457-492.

[7] M. Torres, A. Casadevall, The immunoglobulin constant region contributes to affinity and specificity, Trends Immunol. 29 (2008) 91-97.

[8] A. Braun, L. Kwee, M. Labow, J. Alsenz, Protein Aggregates Seem to Play a Key Role Among the Parameters Influencing the Antigenicity of Interferon Alpha (IFN-α) in Normal and Transgenic Mice, Pharm. Res. 14 (1997) 1472-1478.

[9] H. Schellekens, Bioequivalence and the immunogenicity of biopharmaceuticals, Nat. Rev. Drug. Discov. 1 (2002) 457-462.

[10] H. Wu, R. Kroe-Barrett, S. Singh, A.S. Robinson, C.J. Roberts, Competing aggregation pathways for monoclonal antibodies, FEBS Lett. 588 (2014) 936-941.

[11] C.B. Andersen, M. Manno, C. Rischel, M. Thorolfsson, V. Martorana, Aggregation of a multidomain protein: A coagulation mechanism governs aggregation of a model IgG1 antibody under weak thermal stress, Prot. Sci. 19 (2010) 279-290.

[12] R.K. Brummitt, D.P. Nesta, L.Q. Chang, S.F. Chase, T.M. Laue, C.J. Roberts, Nonnative Aggregation of an IgG1 Antibody in Acidic Conditions: Part 1. Unfolding, Colloidal Interactions, and Formation of High-Molecular-Weight Aggregates, J. Pharm. Sci. 100 (2011) 2087-2103.

[13] A. Beck, S. Sanglier-Cianferani, A. Van Dorsselaer, Biosimilar, Biobetter, and Next Generation Antibody Characterization by Mass Spectrometry, Anal. Chem. 84 (2012) 4637-4646.

[14] A. Beck, E. Wagner-Rousset, D. Ayoub, A. Van Dorsselaer, S. Sanglier-Cianférani, Characterization of Therapeutic Antibodies and Related Products, Anal. Chem. 85 (2013) 715-736.

[15] A. Beck, F. Debaene, H. Diemer, E. Wagner-Rousset, O. Colas, A.V. Dorsselaer, S. Cianférani, Cutting-edge mass spectrometry characterization of originator, biosimilar and biobetter antibodies, J. Mass Spectrom. 50 (2015) 285-297.

[16] R. Gahoual, M. Biacchi, J. Chicher, L. Kuhn, P. Hammann, A. Beck, E. Leize-Wagner, Y.N. Francois, Monoclonal antibodies biosimilarity assessment using transient isotachophoresis capillary zone electrophoresis-tandem mass spectrometry, Mabs 6 (2014) 1464-1473.

[17] R. Gahoual, J.-M. Busnel, A. Beck, Y.-N. François, E. Leize-Wagner, Full Antibody Primary Structure and Microvariant Characterization in a Single Injection Using Transient Isotachophoresis and Sheathless Capillary Electrophoresis–Tandem Mass Spectrometry, Anal. Chem. 86 (2014) 9074-9081.

[18] X. Han, Y. Wang, A. Aslanian, M. Bern, M. Lavallee-Adam, J.R. Yates, Sheathless Capillary Electrophoresis-Tandem Mass Spectrometry for Top-Down Characterization of Pyrococcus furiosus Proteins on a Proteome Scale, Anal. Chem. 86 (2014) 11006-11012.

[19] L.L. Sun, M.D. Knierman, G.J. Zhu, N.J. Dovichi, Fast Top-Down Intact Protein Characterization with Capillary Zone Electrophoresis-Electrospray Ionization Tandem Mass Spectrometry, Anal. Chem. 85 (2013) 5989-5995.

[20] Y. Zhao, N.M. Riley, L. Sun, A.S. Hebert, X. Yan, M.S. Westphall, M.J.P. Rush, G. Zhu, M.M. Champion, F.M. Medie, P.A.D. Champion, J.J. Coon, N.J. Dovichi, Coupling Capillary Zone Electrophoresis with Electron Transfer Dissociation and Activated Ion Electron Transfer Dissociation for Top-Down Proteomics, Anal. Chem. 87 (2015) 5422-5429.

[21] M. Biacchi, R. Bhajun, N. Saïd, A. Beck, Y.N. François, E. Leize-Wagner, Analysis of monoclonal antibody by a novel CE-UV/MALDI-MS interface, Electrophoresis 35 (2014) 2986-2995.

[22] E.A. Redman, N.G. Batz, J.S. Mellors, J.M. Ramsey, Integrated Microfluidic Capillary Electrophoresis-Electrospray Ionization Devices with Online MS Detection for the Separation and Characterization of Intact Monoclonal Antibody Variants, Anal. Chem. 87 (2015) 2264-2272.

[23] M. Biacchi, R. Gahoual, N. Said, A. Beck, E. Leize-Wagner, Y.-N. François, Glycoform Separation and Characterization of Cetuximab Variants by Middle-up Off-Line Capillary Zone Electrophoresis-UV/Electrospray Ionization-MS, Anal. Chem. 87 (2015) 6240-6250.

[24] K. Wenig, L. Chatwell, U. von Pawel-Rammingen, L. Bjorck, R. Huber, P. Sondermann, Structure of the streptococcal endopeptidase IdeS, a cysteine proteinase with strict specificity for IgG, P. Nat. Acad. Sci. USA 101 (2004) 17371-17376.

[25] B. Vincents, U. von Pawel-Rammingen, L. Bjorck, M. Abrahamson, Enzymatic characterization of the streptococcal endopeptidase, IdeS, reveals that it is a cysteine protease with strict specificity for IgG cleavage due to exosite binding, Biochem. 43 (2004) 15540-15549.

[26] D. Ayoub, W. Jabs, A. Resemann, W. Evers, C. Evans, L. Main, C. Baessmann, E. Wagner-Rousset, D. Suckau, A. Beck, Correct primary structure assessment and extensive glyco-profiling of cetuximab by a combination of intact, middle-up, middle-down and bottom-up ESI and MALDI mass spectrometry techniques, Mabs 5 (2013) 699-710.

[27] H. Wu, K. Truncali, J. Ritchie, R. Kroe-Barrett, S. Singh, A.S. Robinson, C.J. Roberts, Weak protein interactions and pH- and temperature-dependent aggregation of human Fc1, mAbs, (2015) DOI: 10.1080/19420862.2015.1079678.

[28] Y.F. Shen, R.D. Smith, High-resolution capillary isoelectric focusing of proteins using highly hydrophilic-substituted cellulose-coated capillaries, J. Microcolumn Sep. 12 (2000) 135-141.

[29] H.C. Liu, G. Gaza-Bulseco, D. Faldu, C. Chumsae, J. Sun, Heterogeneity of monoclonal antibodies, J. Pharm. Sci. 97 (2008) 2426-2447.

[30] T.S. Raju, B.J. Scallon, Glycosylation in the Fc domain of IgG increases resistance to proteolytic cleavage by papain, Biochem. Bioph. Res. Co. 341 (2006) 797-803.

[31] B. Antes, S. Amon, A. Rizzi, S. Wiederkum, M. Kainer, O. Szolar, M. Fido, R. Kircheis, A. Nechansky, Analysis of lysine clipping of a humanized Lewis-Y specific IgG antibody and its relation to Fc-mediated effector function, J. Chromatogr. B, 852 (2007) 250-256.

[32] R. Gahoual, J.M. Busnel, P. Wolff, Y.N. Francois, E. Leize-Wagner, Novel sheathless CE-MS interface as an original and powerful infusion platform for nanoESI study: from intact proteins to high molecular mass noncovalent complexes, Anal. Bioanal. Chem. 406 (2014) 1029-1038.

[33] E.J. Nettleton, M. Sunde, Z.H. Lai, J.W. Kelly, C.M. Dobson, C.V. Robinson, Protein subunit interactions and structural integrity of amyloidogenic transthyretins: Evidence from electrospray mass spectrometry, J. Mol. Biol. 281 (1998) 553-564.

[34] J.L.P. Benesch, B.T. Ruotolo, D.A. Simmons, N.P. Barrera, N. Morgner, L.C. Wang, R. Helen, C.V. Robinson, Separating and visualisiing protein assemblies by means of preparative mass spectrometry and microscopy, J. Struct. Biol. 172 (2010) 161-168.

[35] D. Touboul, L. Maillard, A. Grasslin, R. Moumne, M. Seitz, J. Robinson, R. Zenobi, How to Deal with Weak Interactions in Noncovalent Complexes Analyzed by Electrospray Mass Spectrometry: Cyclopeptidic Inhibitors of the Nuclear Receptor Coactivator 1-STAT6, J. Am. Soc. Mass Spectrom. 20 (2009) 303-311.

3. Conclusion

Lors de ce travail, nous avons utilisé notre stratégie développée dans la partie précédente (I - Séparation et caractérisation de glycoformes de Cetuximab par Off-line CZE-UV/ESI-MS) mettant en jeu le couplage indirect de la CZE-UV/nanoESI-MS, pour la caractérisation de variant de charges du domaine Fc/2 du cetuximab. La séparation électrophorétique correspondant au domaine Fc/2 du cetuximab comporte trois pics. Cependant, seuls deux variants de masses dus à la perte d'une lysine en C-terminal de la HC ont été caractérisés en MS. Afin d'expliquer la présence de ces trois pics électrophorétiques, nous avons dans un premier temps, validé notre stratégie instrumentale en réalisant la même séparation à l'aide du couplage CE-UV/MALDI-MS. La reconstruction des spectres MALDI-MS issue de la séparation CZE ont permis d'obtenir trois pics correspondant à des masses de 25 kDa. Ceci certifie l'absence de phénomènes de diffusion durant le processus de déposition de l'échantillon confirmant la robustesse du couplage. Dans le but de vérifier le potentiel rôle de la structure des N-glycosylations ainsi que de la perte de lysine en C-terminal de la HC sur la séparation électrophorétique, deux séparations CZE-UV/ESI-MS d'un échantillon de cetuximab respectivement déglycosylé et débarrassé de leur lysine en C-terminal, ont été réalisées. Concernant les glycosylations, aucune différence entre les électrophérogrames des échantillons glycosylés et déglycosylé n'a été observée. Ceci permet de conclure sur la non-influence de la nature des glycosylation sur la séparation CZE. Concernant l'absence de lysine en C-terminal de la HC, les résultats montrent la perte des deux premiers pics de la séparation CZE démontrant une influence importante de cette propriété de la molécule. La caractérisation MS basée sur l'infusion nanoESI de chaque fraction collectée de la séparation MS, ont été réalisé à l'aide de différents électrolytes : une solution d'infusion similaire aux conditions de CZE (25 mM acétate d'ammonium pH 5,7) et une solution d'infusion classiquement utilisée en condition dénaturante (eau/acétonitrile/acide formique 49/50/1). Dans les conditions de la séparation CZE, la MS a permis d'identifier sans ambiguïté la nature de chaque analyte séparé. En effet, les spectres MS ont mis en évidence la présence d'homodimères de $F_{c/2}$ correspondant au premier pic de la séparation, la présence d'hétérodimères de $F_{c/2}$/$F_{c/2}$-K pour le second pic et d'homodimères de $F_{c/2}$-K pour le dernier pic. Ces résultats confirment l'influence de la présence ou non de lysine en C-terminal de la HC sur la séparation CZE. De plus la mise en œuvre d'une séparation électrophorétique avant l'analyse MS a montré son intérêt notamment pour séparer des complexes non-covalents tout en montrant la possibilité de maintenir ces agrégats en solution.

PARTIE IV : Caractérisation structural d’anticorps conjugués par CESI-MS à différents niveaux : intact, middle-up et peptide.

I - Caractérisation à différents niveaux d'anticorps conjugués par la combinaison d'approches MS native, middle-up et bottom-up protéomique à l'aide du couplage CESI-MS

1. Introduction

Les anticorps conjugués (ADCs) sont une catégorie émergeante de molécules biothérapeutiques dérivés des anticorps monoclonaux (mAbs). Ils combinent la forte activité anti-cancereuse des médicaments cytotoxiques ainsi que la sélectivité et les propriétés pharmacologiques favorables des mAbs. La synthèse des ADCs est obtenue par une liaison chimique entre le médicament et le mAbs. Aujourd'hui, deux ADCs (brentuximab vedotin et trastuzumab emtansine) ont été approuvés et mis sur le marché par la FDA, et plus de 50 candidats sont actuellement en cours d'essais cliniques ce qui montre l'intérêt croissant de l'industrie biopharmaceutique pour ces composés. Le principe d'un ADCs est basé sur l'utilisation de la forte spécificité du mAbs pour cibler un type de cellules tumorales et ainsi permettre par relargage, l'application de médicaments extrêmement puissants. Cela limite l'exposition des cellules saines et donc permet d'optimiser les traitements notamment avec la diminution des effets secondaires.

La conjugaison du médicament sur l'ADCs peut être réalisée par différentes réactions chimiques dépendant du site de conjugaison c'est à dire la nature de l'acide aminé ou de la glycosylation concernée. Aujourd'hui, les deux ADCs commercialisés sont issus d'une conjugaison du médicament soit sur une lysine, soit sur une cystéine après réduction des ponts disulfures. De telles réactions chimiques amènent à la production d'un mélange de différentes espèces pouvant porter un nombre fluctuant de médicaments sur la protéine. La complexité de ce mélange peut être retranscrite à l'aide de certain paramètres permettant de caractériser un ADCs : la distribution de conjugués sur le mAbs, la quantité de mAbs non conjugués ainsi que le ratio médicament sur anticorps (DAR). Ces trois paramètres sont considérés comme attributs critiques de qualité (CQAs) pour les ADCs car ils ont un fort impact sur la sécurité, la pharmacocinétique et la pharmacodynamique de la molécule.

L'hétérogénéité des ADCs, générée par la variation du nombre et la position de médicaments conjugués sur le mAbs durant sa production, implique une réponse immunitaire différente suivant la valeur du DAR. Tout comme les mAbs, les ADC sont des glycoprotéines complexes potentiellement porteuses de nombreuses micro-hétérogénéités moléculaires dont les sources

sont variées (procédés de production inadaptée, conservation prolongée). La complexité de ces protéines et leur utilisation en tant que traitement thérapeutique justifient la nécessité d'améliorer les méthodologies analytiques de caractérisation. Actuellement, un large panel de techniques analytiques est utilisé pour répondre à caractérisation structurale fine des protéines thérapeutiques incluant les ADCs. La spectrométrie de masse (MS) a su rapidement s'imposer comme une méthode de choix pour la caractérisation des protéines thérapeutiques, notamment en raison de son excellente spécificité, sensibilité mais aussi car elle permet d'obtenir des informations structurales dans certain cas. Les trois parties précédentes décrites dans ce manuscrit ont démontré l'intérêt du couplage CZE-MS pour la caractérisation structurale de mAbs et l'évaluation de la biosimilarité.

Dans ce chapitre, nous avons donc appliqué une approche globale mettant en œuvre l'utilisation du couplage CESI-MS en tant que plateforme d'infusion nanoESI et méthode séparative pour obtenir la caractérisation structurale à différents niveaux d'un ADCs. L'ADCs sélectionné comme modèle standard pour cette étude est le brentuximab vedotin (BV) qui est l'un des deux ADCs approuvés par la FDA (2011) pour le traitement du lymphome d'Hodgkin. La structure de cet ADCs est basée sur la conjugaison d'un mAbs chimérique cAC10 par un inhibiteur de la tubuline monométhylique auristatine E (MMAE) via réduction des ponts disulfures inter-chaines. La présence de cystéines liées aux MMAE ne perturbe pas la totalité des ponts disulfures entre les chaines HC et LC ce qui permet de maintenir la structure de la protéine et donc sa haute spécificité pour l'antigène ciblé.

Le premier niveau de caractérisation décrit dans ce chapitre consiste à évaluer le couplage CESI-MS en tant que plateforme nanoESI en MS native pour la détermination du DAR moyen du BV et pour obtenir la distribution de conjugués sur le mAbs. Ensuite, dans le but d'approfondir la caractérisation du BV en utilisant le même système, nous avons réalisé une approche middle-up par digestion IdeS. La réaction enzymatique de l'ADCs avec l'IdeS suivi de son passage en MS permet d'obtenir deux types de fragments distincts (Fc/2 et F(ab')2) ainsi que de nombreuses informations de structure. Enfin, une méthode CESI-MS basée sur une approche bottom-up protéomique a été développée dans le but d'obtenir la structure primaire de l'ADCs ainsi que les peptides liés aux médicaments. En effet, due aux propriétés d'hydrophobicités du médicament, nous avons dans un premier temps optimisé les conditions du protocole de digestion trypsique. Puis, à l'aide du couplage CESI-MS, nous avons réalisé la caractérisation simultanée de la séquence en acide aminé de l'ADCs ainsi que son profil de glycosylation et l'identification des peptides conjugués. Ceci nous a permis d'obtenir des informations primordiales de structure concernant l'efficacité de conjugaison du médicament sur le mAbs.

Les résultats présentés dans cette partie ont fait l'objet d'une publication dans le journal scientifique *Analytica Chimica Acta (*Elsevier 2016).

2. Structural characterization of antibody drug conjugate by a combination of intact, middle-up and bottom-up techniques using sheathless capillary electrophoresis – tandem mass spectrometry as nanoESI infusion platform and separation method

Nassur Said, Rabah Gahoual, Lauriane Kuhn, Alain Beck, Yannis-Nicolas Francois, Emmanuelle Leize-Wagner, ***Analytica Chimica Acta*** 2016 28, 50-59.

Abstract

Antibody-drug conjugates (ADCs) represent a fast growing class of biotherapeutic products. Their production leads to a distribution of species exhibiting different number of conjugated drugs overlaying the inherent complexity resulting from the monoclonal antibody format, such as glycoforms. ADCs require an additional level of characterization compared to first generation of biotherapeutics obtained through multiple analytical techniques for complete structure assessment. We report the development of complementary approaches implementing sheathless capillary electrophoresis-mass spectrometry (sheathless CE-MS) to characterize the different aspects defining the structure of brentuximab vedotin. Native MS using sheathless CE-MS instrument as a nanoESI infusion platform enabled accurate mass measurements and estimation of the average drug to antibody ratio alongside to drug load distribution. Middle-up analysis performed after limited IdeS proteolysis allowed to study independently the light chain, Fab and F(ab')2 subunits incorporating 1, 0 to 4 and 0 to 8 payloads respectively. Finally, a CZE-ESI-MS/MS methodology was developed in order to be compatible with hydrophobic drug composing ADCs. From a single injection, complete sequence coverage could be achieved. Using the same dataset, glycosylation and drug-loaded peptides could be simultaneously identified revealing robust information regarding their respective localization and abundance. Drug-loaded peptide fragmentation mass spectra study demonstrated drug specific fragments reinforcing identification confidence, undescribed so far. Results reveal the method ability to characterize ADCs primary structure in a comprehensive manner while reducing tremendously the number of experiments required. Data generated showed that sheathless CZE-ESI-MS/MS characteristics position the methodology developed as a relevant alternative for comprehensive multilevel characterization of these complex biomolecules.

Introduction

Antibody-drug conjugates (ADCs) are an emerging category of biotherapeutic products, derived from monoclonal antibodies (mAbs)[1], they combine the high anti-cancer potency of cytotoxic drugs to the exquisite selectivity and favorable pharmacological properties of mAbs and are obtained through chemical linkage [2]. In addition to the recent approval by the US Food and Drugs Administration (FDA) of brentuximab vedotin (Adcentris, Seattle Genetics) and trastuzumab emtansine (Kadcyla, Genentech/ Roche), more than 50 ADCs are currently investigated in clinical trials indicating the broad interest of the biopharmaceutical industry for this type of treatments [3, 4]. ADCs incorporate highly potent warheads which can be delivered using the specificity provided by mAbs to target tumor cells and limit exposition of healthy cells therefore optimizing treatment potency and limit the occurrence of side effects [5].

Drug conjugation can be implemented through different chemical reactions depending on the amino acid residue or glycan concerned [6]. Especially, drug coupling using the lysine residue's amino group and cysteine residues through thiol group after reduction of disulphide bridges have been reported[7, 8]. Drug conjugation reaction leads to the production of a mixture of different species showing various numbers of drugs conjugated to the protein. While for a single species, drugs can still be distributed in various manner on the different conjugation sites. This complexity is reflected by parameters such as the drug loading and distribution, the amount of naked antibody (D0) and the average number of drug to antibody (DAR). These three parameters are critical quality attributes (CQAs) for ADCs because they impact on safety, pharmacokinetics (PK) and pharmacodynamics (PD) [9].

ADCs heterogeneity generated from different number and position of conjugated drugs during production is combining to the intrinsic variability coming from conjugation with mAbs. mAbs are complex glycoproteins exhibiting a consequent number of micro-heterogeneities which can influence significantly antibody-dependent cell-mediated cytotoxicity (ADCC) and complement-dependent cytotoxicity of the considered mAb [10, 11]. Currently a large panel of analytical techniques are required in order to perform a comprehensive structural characterization of biotherapeutic proteins including ADCs [12]. Mass spectrometry (MS) has gradually taken a decisive position in structure characterization of biopharmaceutical products [13-15], likewise ADCs, mainly due to an outstanding selectivity allowing the identification of the different species composing ADCs samples and the opportunity to obtain structural information [16]. MS still often needs to be used in concomitance with separation methods mainly including liquid chromatography[17, 18]. Recently, we described a method using capillary zone electrophoresis (CZE) coupled to tandem MS with a sheathless interface to perform the simultaneous characterization of several aspects of a protein in one injection

including amino acid sequence, glycosylation characterization and other types of posttranslational modifications (PTMs) [19, 20].

In this work, we report the development of a combined approach implementing a sheathless capillary electrophoresis (CE) interface used as nanoESI infusion platform or as separation method, hyphenated to tandem MS (MS/MS) for the detailed structural characterization of different traits composing the primary structure of ADCs. The ADC considered is brentuximab vedotin (BV) which has been approved in 2011 by the FDA and consequently by the European Medicine Agency (EMA) for the treatment of Hodgkin Lymphoma and systemic anaplastic large cell lymphoma [7, 21]. It is composed of the tubulin inhibitor monomethyl auristatin E (MMAE) conjugated to the chimeric mAb cAC10 via a protease cleavable maleimidocaproyl-valine-citrulline-((*p*-aminobenzyl)oxy)carbonyl linker (MC-PV-PABC) by reduction of the interchain disulfide bonds [7, 22]. The presence of the cysteine linked MMAE does not disrupt the interchain disulfide bonds between the heavy chain (Hc) and (Lc) of the mAb which allows to maintain the higher order structure of the protein and therefore its specificity for the targeted antigen [23, 24].

As the first level of characterization, we evaluated the sheathless CE-MS system as nanoESI infusion platform for native MS determination of the average DAR of BV and to obtain the drug load distribution. In a second step, to further investigate the characterization of BV using the instrument setting, we realized the middle-up analysis of BV by IdeS proteolysis. IdeS enzymatic reaction cleaves ADC between the two consecutive glycine residues present in the hinge region to obtain two types of fragments ($F_{c/2}$ and F(ab')2). Finally, a sheathless CE-MS/MS method derived from bottom-up proteomic analysis was developed in order to be fully compatible to the requirement of the analysis of ADCs. Analysis of BV was performed by sheathless CZE-ESI-MS/MS enabling the simultaneous characterization of the amino acid sequence of the ADC alongside to the characterization of glycoforms as well as drug-loaded peptides giving precious information regarding the drug effectively conjugated to the mAb.

Materials and Methods

Materials. Chemicals used were of analytical grade or high purity grade and purchased from Sigma-Aldrich (Saint Louis, MO, USA). Water used to prepare buffers and sample solutions was obtained using an ELGA purelab UHQ PS water purification system (Bucks, UK). RapiGest SF surfactant was purchased from Waters (Milford, MA, USA). IdeS (immunoglobulin degrading enzyme of Streptococcus pyogenes) also named FabRICATOR was purchased from Genovis (Lund, Sweden). Brentuximab vedotin (Adcetris®) was produced by Millenium Pharmaceuticals/Takeda (London, UK).

Intact ADC sample preparation. Brentuximab vedotin (BV) was buffer exchanged with 200 mM ammonium acetate buffer, pH 7.0 using Amicon centrifugal filters (cut off = 10,000 Da) (Merck Millipore, Darmstadt, Germany). Six desalting step were performed at 10°C and 14,000×g for 20 min each step. Ammonium acetate allows the native structure of proteins to be preserved and be compatible to ESI-MS.

ADC IdeS digestion. BV was cleaved in the hinge region using proteolysis by IdeS (FabriCATOR, Genovis) to obtain two Fc/2 fragments and one F(ab')2 fragment. Sample was diluted using 147.25µL of 50 mM sodium phosphate, 150 mM NaCl, pH 7.4, to a final concentration of 1 µg/µL. A volume of 2.25 µL of IdeS (67 units/µL) was added to the sample which was left at 37°C for 30 min. After digestion completion, sample was buffer exchanged with 200 mM ammonium acetate buffer, pH 7.0 using Amicon centrifugal filters (cut off = 10,000 Da) at 10°C and 14,000×g for 20 min.

In solution protein digestion. Prior to tryptic digestion, BV was cleaved by IdeS following the ADC IdeS digestion protocol. After IdeS digestion completion, sample was buffer exchanged using three desalted cycles (Amicon centrifugal filters, 10,000 cutoff) with 50 mM ammonium bicarbonate buffer, pH 8.0 at 4°C and 14 000g for 20 min. Theoretical concentration of 12.5 µg/µL was obtained by reducing sample volume to 10 µL. Sample were then diluted using 0.1% RapiGest surfactant to a final concentration of 6.25 µg/µL and heated to 40°C during 10 min. Dithiothreitol (DTT) was added to the sample in order to obtain a final concentration of 25 mM. Samples were then heated to 80°C during 10 min. After being cooled down to room temperature (RT), 10% of acetonitrile was added to the sample. A volume of 2 µL of trypsin (0.5µg/µL) was added to the sample which was left at room temperature for 3h, another volume of 2 µL was added afterward. Digestion was performed overnight at 37°C. A second reduction step was performed by addition of DTT to a final concentration of 35 mM and heated for 45 min at 56°C. Then 40% of isopropanol was added to the sample. After digestion completion, formic acid (FA) was added to a final concentration of 1% (v/v) in order to cleave the surfactant. Sample was left at RT for 2h. Sample was finally diluted to a final concentration in protein of 2.2 µM using ammonium acetate 50 mM, pH 4.0.

Capillary electrophoresis. CE experiments and NanoESI infusion experiments were performed with a CESI8000 capillary electrophoresis system from Sciex Separation (Brea, CA, USA). Bare fused-silica capillaries (total length 100 cm; 30 µmi.d.) with characteristic porous tip on its final end on 3 cm, a second capillary (total length 80 cm; 50 µmi.d.) filled during experiments with BGE allows electric contact. New capillaries were flushed for 10 min at 75 psi (5.17 bar) with methanol, then 10 min with 0.1 M sodium hydroxide, followed 10 min with 0.1 M hydrochloric acid and water for 20 min also at 75 psi. Finally, the capillary was flushed

10 min at 75 psi with BGE which was 10% acetic acid. For CE experiments, hydrodynamic injection (10 psi for 1 min) corresponding to a total volume of 90 nL of sample injected was used. Separations were performed using a voltage of +20 kV. For infusion experiments, sample was flushed into the capillary at 50 psi for 3 min. The infusion was carried out by flushing sample at 5 psi for 5 min corresponding to a total volume of 245 nL of sample.

Mass spectrometry. For ADC peptide mapping characterization, the CESI8000 system was hyphenized to a Triple TOF 6600 mass spectrometer (Sciex, Darmstadt, Germany). The TripleTOF 6600 MS is equipped with a hybrid analyzer composed of a quadrupole followed by a time-of-flight (TOF) analyzers. ESI source parameters was set as follow: ESI voltage -1.45kV while Gas supplies (GS1 and GS2) were deactivated, source heating temperature 150°C and curtain gas value 4. Mass/charge (m/z) range was 100-2000 in MS and 50-2000 in MS/MS. External calibration was performed using a digest of β-galactosidase from Sigma-Aldrich (Saint-Louis, MO, USA). For ADC NanoESI infusion experiments, the CESI8000 system was coupled to a maxis 4G (Bruker Daltonics, Bremen, Germany). MS transfer parameters were optimized using the actual sample directly infused via the CE system using a pressure of 5 psi (340 mbar). Ion funnels were set at values of 300 and 400 Vpp. The electrospray voltage (capillary voltage) typically ranged from −1.2 to −1.8 kV. Dry gas was set at 1.5 L/min and source temperature at 180 °C. Data processing was performed with DataAnalysis 4.0. Deconvolution of the mass spectra was performed on the basis of maximum entropy analysis using ESI Compass 1.3 Maximum Entropy Deconvolution Option. All spectra were calibrated by external calibration using CsI from Sigma-Aldrich (Saint-Louis, MO, USA).

MS/MS data analysis. Data obtained from the sheathless CE-MS/MS experiments were analyzed using Peakview and BioParmaView software (Sciex, San Francisco, CA). The mass tolerance allowed for search algorithm identification, for precursor ions was set to ± 5 ppm and ±0.05 Da for fragmentation ions.

Drug-to-Antibody Ratio (DAR) determination. Average DAR values are calculated from the relative peak area of each peak from the charge state deconvoluted mass spectrum. Average DAR value is obtained following the equation:

$$DAR = \frac{\sum nA_{Drug}}{\sum A_{Drug}}$$

With n the number of drug loads ranging from 0 to 8.

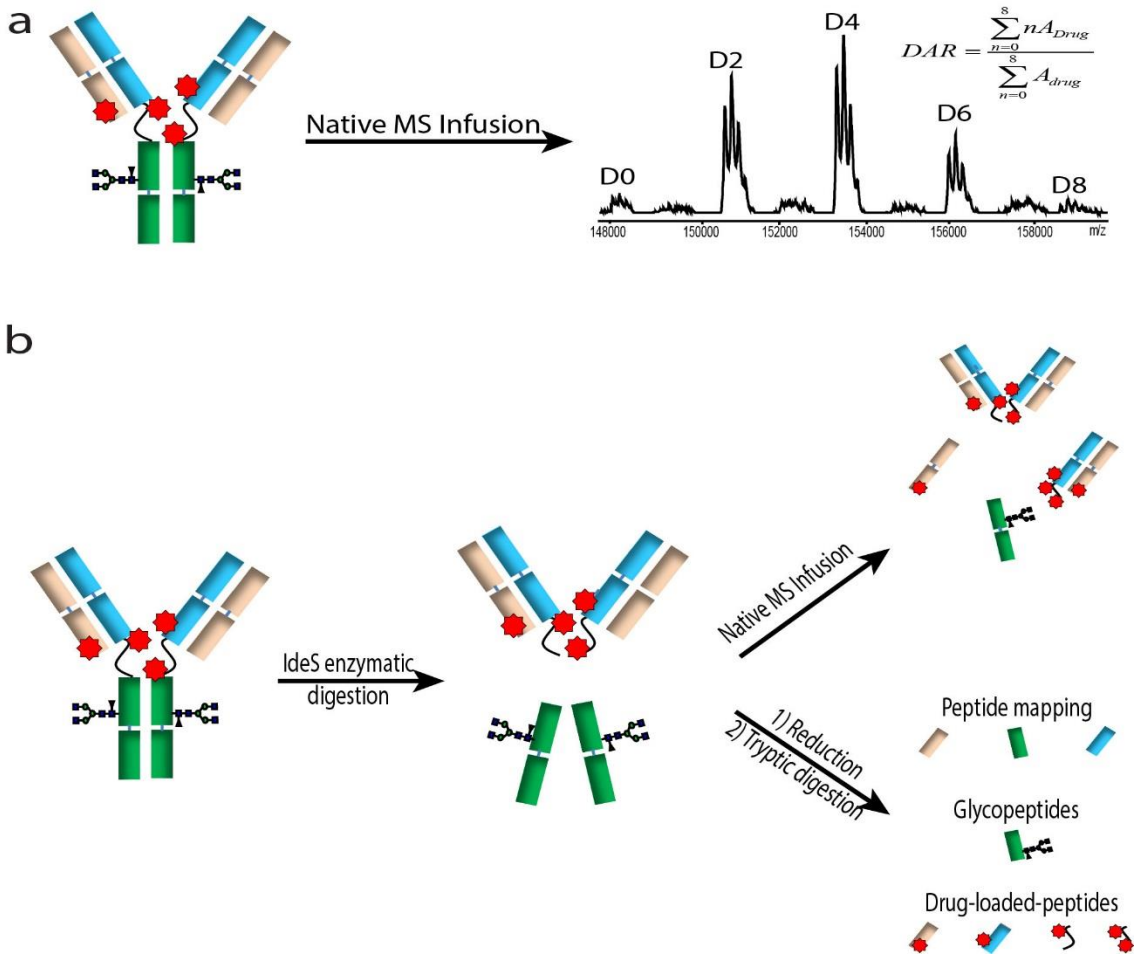

Figure 1. Overview of brentuximab vedotin structural characterization using sheathless CE-MS. (a) Native MS infusion for average DAR determination and drug loaded distribution assessment, (b) middle-up and bottom-up analysis.

Results and Discussion

Average DAR determination using sheathless CE-MS Interface as a nanoESI emitter

For the first level of the characterization of ADCs, we determined the average DAR of BV and we assessed the drug load distribution by native MS on a high resolution Q-TOF instrument using sheathless CE-MS interface as a nanoESI emitter (Figure 1-a). In a recent work, we demonstrated the suitability of sheathless CE-MS system to be used as nanoESI infusion platform [25]. Benefits of nanoESI as the lower source temperature and spraying voltage, the increase of sensitivity and the decrease of ion suppression phenomenon make this interface a particularly suitable system for biotherapeutic characterization. To achieve optimal MS resolution and sensitivity, BV was systematically desalted before infusion removing any components in the formulation buffer that can affect the stability of the spray. Within a single run of 5 min at a flow rate of 50 nL/min which represent an infused sample volume of 250 nL,

we obtained a distribution of baseline resolved species of BV (Figure 1-a) with masses corresponding to the intact mAb linked with zero to eight payloads (Table S-1). Mass accuracies on intact BV using Q-TOF instrument is in total agreement with results reported in the literature [15, 24]. As expected, no detection of ions corresponding to dissociated BV, referring to free drug loaded Lc or Hc formed during the ionization or the ion transfer inside the mass spectrometer, were observed confirming that the combination of sheathless CE-MS interface with native MS enhances the detection of drug load species and allows to obtain the drug-load distribution. Based on charge state deconvoluted mass spectrum of intact and deglycosylated BV, average DAR value of 3.8 and 3.9 respectively were calculated using the relative quantification from the peak area of each peak and the corresponding number of drugs loaded. Obtained average DAR values are in good agreement with values described in the literature [15, 24] [23]. It demonstrated the ability of the sheathless CE-MS system used as nanoESI infusion platform to get structural information of intact BV within a single run and consuming in total less than of 3µg of ADC samples.

Middle-up analysis of brentuximab vedotin

Dozens of recent reports from the FDA, Industry and academic clearly show that to extensively characterize large biopharmaceuticals such as mAbs, ADC, biosimilars to name a few, consistent data must be reported at multiple levels [26-34] . To investigate deeper the structure of BV using the same system, we performed the middle-up analysis of BV by IdeS proteolysis. Middle-up refers to the mass measurements of large fragments of proteins after proteolysis [27]. IdeS is a cysteine endopeptidase enzyme naturally secreted by Streptococcus pyogenes [35], IdeS enzymatic reaction cleaves specifically immunoglobulin G (IgG) between the two consecutive glycine residues present in the hinge region (Figure 1-b). IdeS cleavage results in two fragments (Figure 1-b) : Fc/2 domain carrying a N-glycosylation site in the Asn 297 with a mass of approximately 25 kDa (Table S-1) depending on the glycoform and F(ab)'2 domain potentially linked by zero to eight payloads with masses of approximately 97 kDa to 108 kDa (Table S-1). Following the same protocol used for intact ADC, within a single run of 5 min at a flow rate of 50 nL/min, we obtained a raw MS data with three main charge envelope profiles (Figure 2-a). The resulting most intense charge envelope between m/z 2000 and 3300 (Figure 2-b) corresponds to the light chain (Lc) with the incorporation of one payload (Mw = 25040.1 ± 0.1 Da). Of note, two glycations corresponding to a difference of mass of 162 Da were characterized (Figure 2-b1). The second part of the raw data matching to the most intense charge envelope between m/z 3500 and 4800 corresponds to two different subunits of the BV. First, as shown in Figure 2-b2, the measured masses correspond to the Fab domain linked by zero to four payloads (Table S-1). This confirms that, even in native condition, IdeS reaction affects BV with the partial loss of noncovalent interaction certainly due to the hydrophobicity of

the drug whose the chemical impact is greater on 100 kDa subunits as compared to intact ADC.

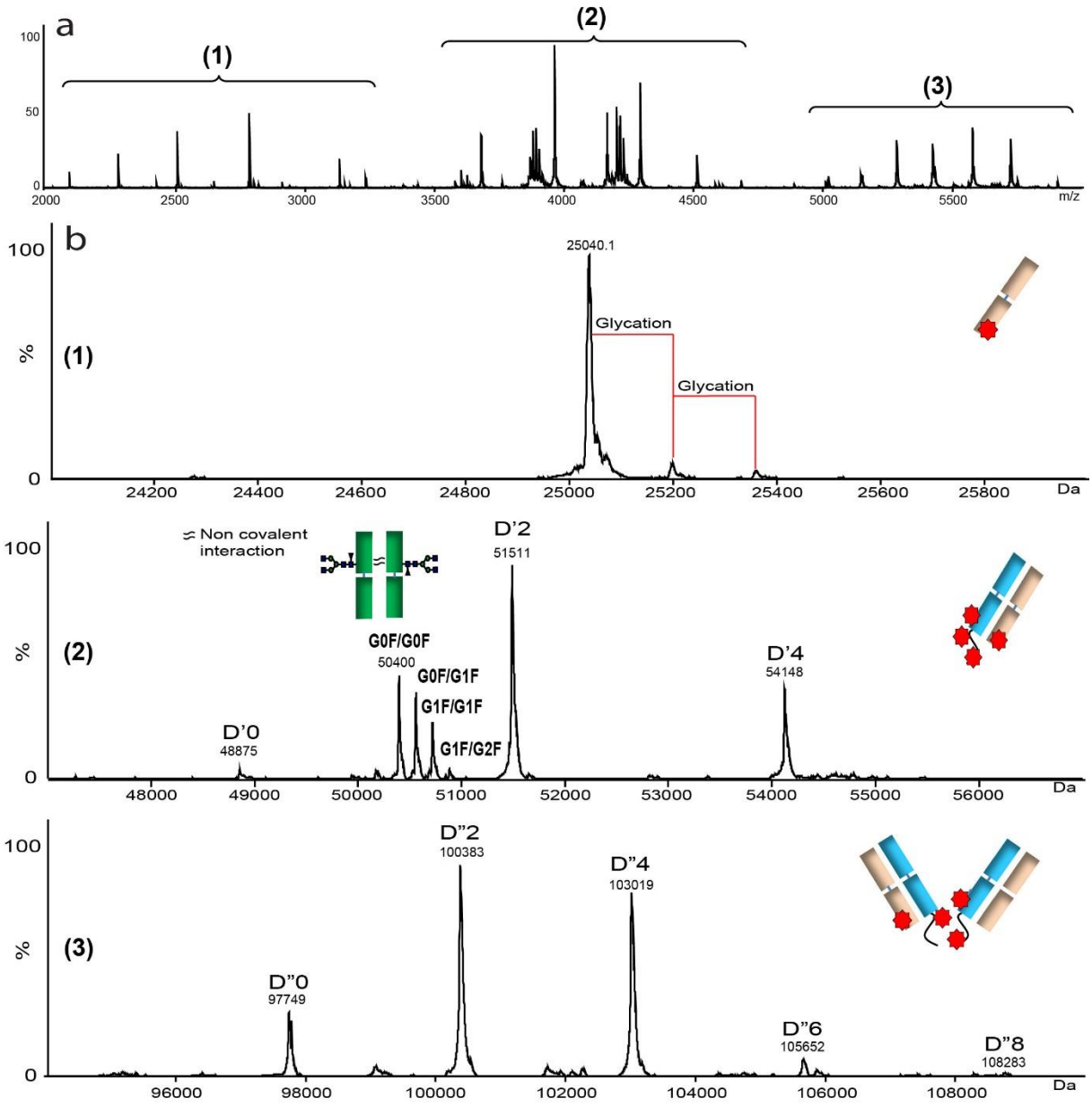

Figure 2. (a) MS spectra corresponding to Native MS NanoESI infusion of middle-up brentuximab vedotin. (b) Charge state deconvoluted mass spectra of (1) Lc-drug conjugated subunit, (2) Fab subunits with the incorporation of 0 to 4 payloads and $F_{c/2}$ homodimers and (c) F(ab)'2 subunits with the incorporation of 0 to 8 payloads.

In second way, charge variant deconvoluted mass spectra also generated multiple masses between 50.4 kDa to 50.8 kDa (Figure 2-b2). These peaks correspond to the formation of protein-protein aggregates due to noncovalent interactions between Fc/2 subunits. Indeed, mass peak 50400 ± 1 Da correspond to the mass of Fc/2 subunits homodimer bearing two

G0F glycosyl-ation (Table S-2). Other mass peaks between 50.4 kDa and 50.8 kDa correspond to the expected presence of the combination of neutral glycosylation variants of this dimer (Table S-1). In addition, no deconvoluted mass spectrum have been observed at 25 kDa that excludes the presence of free Fc/2 variants in ammonium acetate 200mM at pH 7.0. To our knowledge, this is the first time that non-covalent dimers of ADC Fc/2 subunits have been observed and characterized. Finally, the third profile of the raw data (Figure 2-a) corresponding to the most intense charge envelope between m/z 5000 to 6000 is in agreement with the theoretical masses of F(ab')2 subunits with the incorporation of 0, 2, 4, 6 and 8 payloads (Figure 2-c3 and Table S-1). The profile obtained for the middle-up analysis shows that IdeS digestion generates for the same part of the protein different kind of fragments which is illustrated for example by the detection of free Lc and F(ab')2 incorporating also the Lc of the ADC. Therefore the mixture generated from the limited digestion did not allow to estimate a satisfying DAR value justifying the intact analysis described previously. However IdeS proteolytic digestion still remains relevant in order to study specific traits defining the structure of therapeutic proteins due to the lowered and distinctive m/z range, especially to evaluate the efficiency of the coupling reaction for each proteolytic fragment. Therefore MS resolution and mass accuracy is improved supported by the increase of sensitivity and the decrease of ion suppression phenomenon provided by the nanoESI infusion demonstrated the ability of the sheathless CE-MS system to obtained structural information of BV after proteolysis.

Primary structure characterization of brentuximab vedotin.

MS based bottom-up analysis is commonly used to determine protein amino acid sequence and to locate or quantify PTMs. For primary structure characterization (Figure 1-c), this methodology is particularly crucial in early development of therapeutic antibodies as well as during long-term life cycle management of the biopharmaceutical products [36]. As part of this work, one of the objectives was to investigate the possibility to perform robust and confident peptide mapping of BV by sheathless CZE-ESI-MS/MS.

BV was characterized in a bottom-up proteomic adapted strategy, samples were digested by trypsin using an in solution digestion protocol. Recently, we described a method using sheathless CZE-ESI-MS to perform the simultaneous characterization of several aspects of a mAb in one injection including amino acid sequence and in comitance several types of PTMs [19, 37]. However, for the BV analysis, the intensity of some peptide signals on the electropherogram was low or non-existent due to the presence of hydrophobic drug-loaded-peptides. More recently, Janin-Bussat et al. described the first Lys-C digestion protocol for the characterization of drug-loaded-peptides of ADC by LC-MS analysis. They improved the classical protocol by several steps of reduction, fractionation, enrichment and addition of

organic solvent during the digestion protocol to maintain the drug-loaded peptides in solution [17].

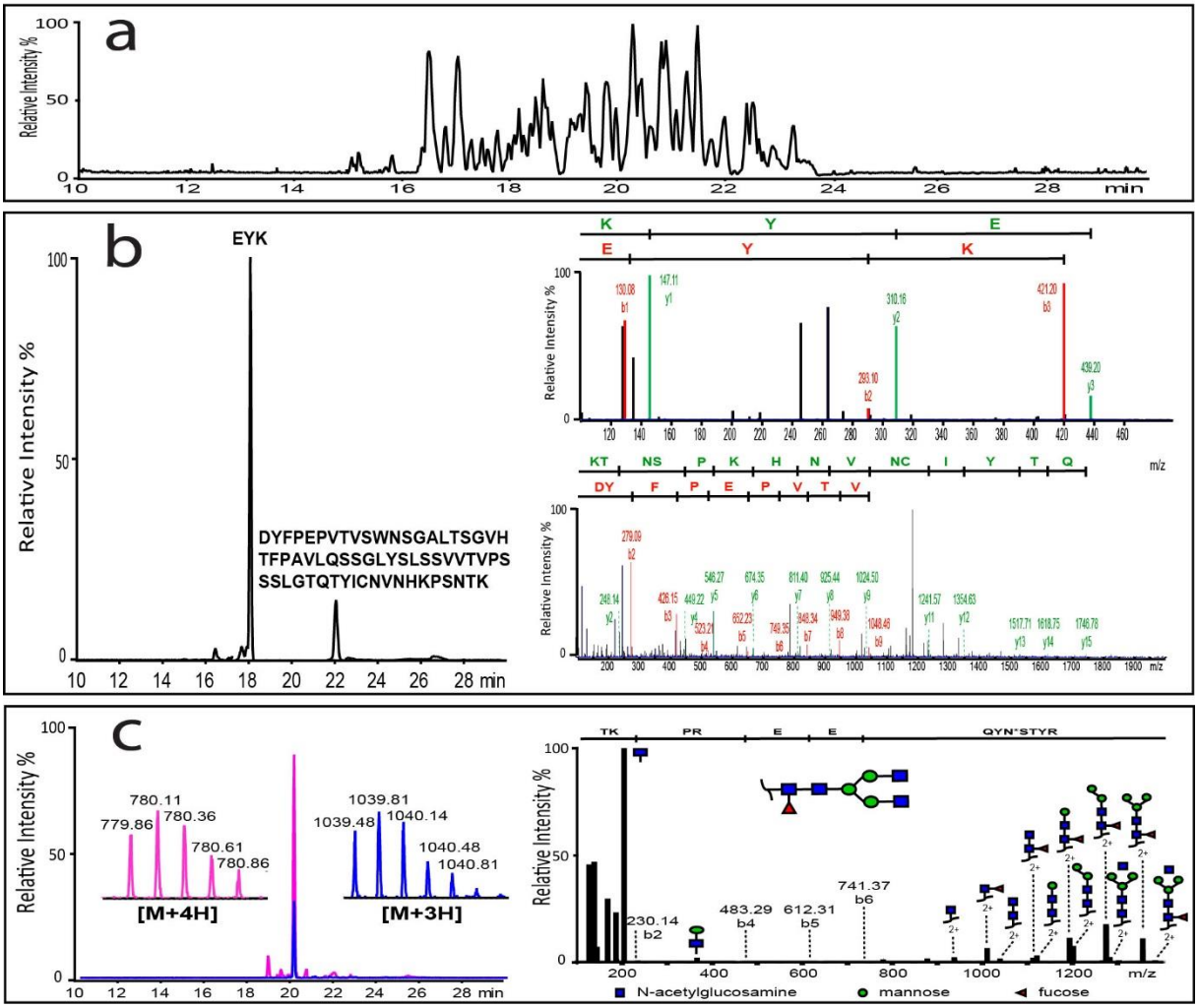

Figure 3. (a) Base Peak Electropherogram corresponding to the analysis by CESI-MS of brentuximab tryptic digest. (b) Extracted Ion Electrophorogram (EIE) corresponding to the m/z ratios of [EYK] and [DYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTK]. MS/MS fragmentation spectra for both peaks (right-hand side). (c) EIE of m/z ratios 780.11 and 1039.81 ([TKPREEQYNSTYR] + G0F) and corresponding MS/MS fragmentation spectra (right-hand side).

Taking that into account and to enhance the level of characterization especially regard-ing ADCs amino acid sequence, glycoforms and drug-loaded-peptides, the sample treatment was adapted for trypsin digestion while still compatible with the electrophoretic separation, and modified with the elimination of the enrichment and LC collection steps. Briefly, prior to trypsin digestion protocol, IdeS enzymatic reaction was performed to cleave BV between the two consecutive glycine residues present in the hinge region to obtain the Hc peptide [THTCPPCPAPELLG] which can be conjugated by one or two payloads. After IdeS digestion

completion, sample was buffer exchanged against 50 mM ammonium bicarbonate buffer, pH 8.0. Before trypsin digestion, 10% of acetonitrile was added to the sample and 40% of isopropanol after digestion [17]. Sample was finally diluted to a final concentration in protein of 2.2 µM using ammonium acetate 50 mM, pH 4.0. BV digests were analyzed through sheathless CZE-ESI-MS/MS with an injection volume corresponding to a quantity of 200 fmol of digested peptides.

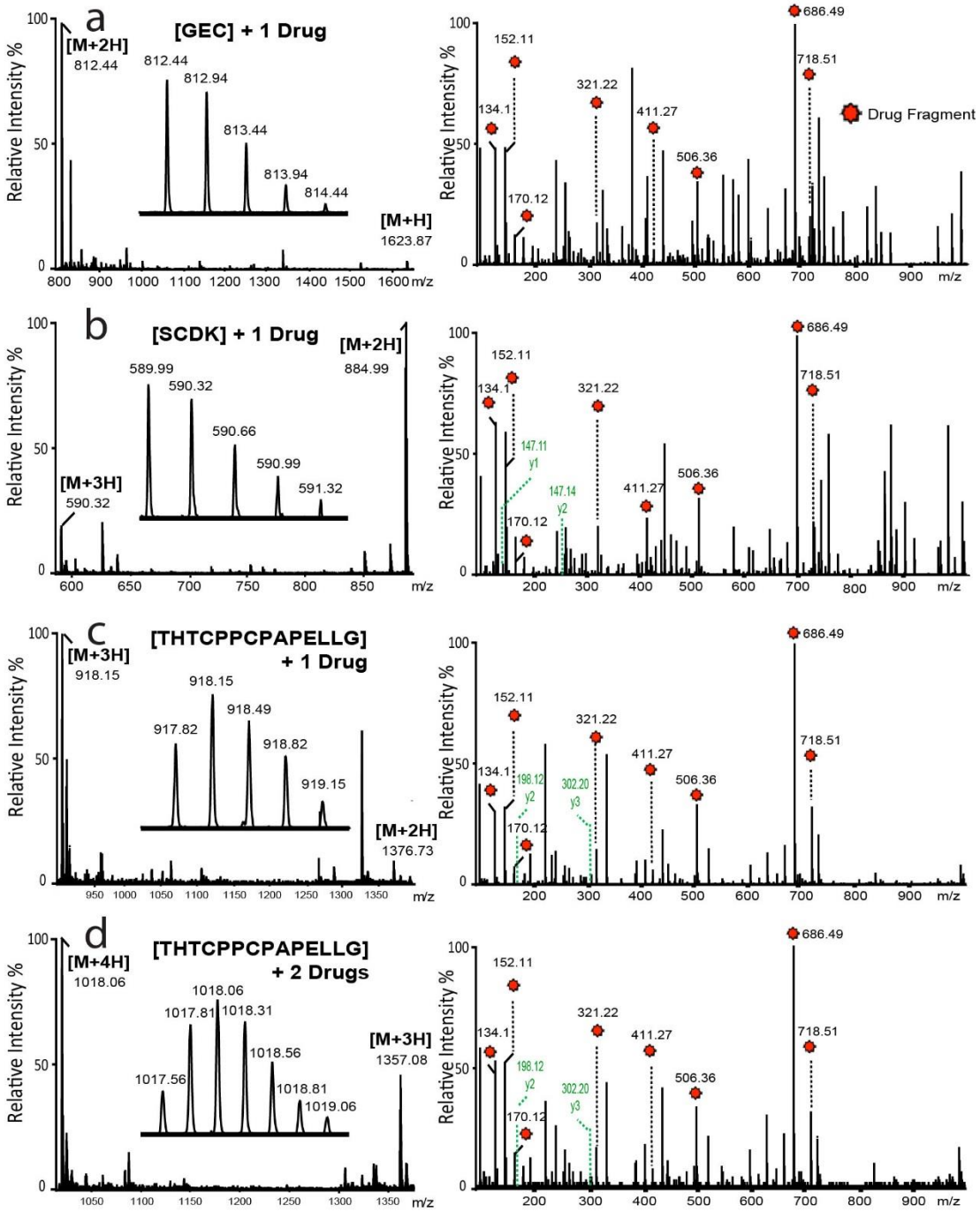

Figure 4. MS and MS/MS spectra of drug-loaded peptides. (a) [GEC] + 1 payload, (b) [SCDK] + 1 payload, (c) [THTCPPCPAPELLG] + 1 payload and (d) [THTCPPCPAPELLG] + 2 payloads

The total duration of the separation was less than 35 min and the resulting electropherogram showed that all peptides were detected between 14 to 28 min (Figure 3-a). The tremendous efficiency of CZE separation allows peak widths at half height ranging from 0.15 min down to 0.8 min representing a number of theoretical plates ranging from 200000 up to 600000 plates. Digested peptides were analyzed by MS/MS enabling to characterize in a single injection the complete sequence coverage of BV. These results confirm the absence of any detrimental effect due to the modification of the enzymatic proteolysis protocol on the amino acid sequence characterization compared to previously published results [19]. The presence of organic solvent during reduction and digestion steps did not prevent the electrophoretic separation and detection of tryptic peptides regardless of their properties. Indeed, tryptic digestion produces a heterogeneous mixture of peptides, especially with regard to their number of amino acids. The identification during the analysis of very small peptide of 3 amino acids are observed alongside to a 63 amino acids peptide having a molecular mass of 6656.29 Da (Figure 3-b). Additionally, peptide identification was based on both peptide mass measurement and fragmentation products using MS/MS spectra in order to enable unambiguous characterization. Results demonstrate the possibility to perform characterization of ADCs amino acid sequence using the sheathless CZE-ESI-MS/MS method developed advanced and using only a single injection of 200 fmol of digested BV peptide mixture.

Characterization of brentuximab vedotin N-glycosylation.

ADCs are glycosylated proteins and those glycans are naturally incorporated in the protein during secretion into the extracellular environment [6]. BV is N-glycoslylated on the Asn 297 in the CH2 region. The excellent ionization efficiency provided by sheathless CZE-ESI-MS allows detection of glycopeptides with significant intensities, generating good quality MS/MS identification of even lowly abundant species. N-glycosylation characterization of BV was performed using the same data sets used for peptide mapping. In the case of BV, the extracted ion electropherograms (XIE) of the m/z ratio matching to the peptide TKPREEQYN297STYR bearing 11 glycosylations were observed (Table S-2). Identifications were based on accurate mass measurement in MS1 (Figure 3-c) provided by high resolution MS and additionally fragmentation spectra. MS/MS spectra corresponding to the fragmentation of the different glycopeptides were subsequently extracted from the data. As expected with low energy collision induced dissociation (CID), product ions observed in MS/MS mainly correspond to the fragmentation of the glycan moiety. Very low intense b ions fragment can be observed but the peptide backbone is usually preserved (Figure 3-c). All detected glycopeptides have a charge state of 3 and 4. The similarity of the glycoforms characterization of BV in terms of number of glycosylation and MS/MS quality spectra as compared to the mAb characterization previously described [19], allows to conclude that the modification of the enzymatic protocol, especially

the presence of acetonitrile and isopropanol did not affect the characterization of glycopeptides.

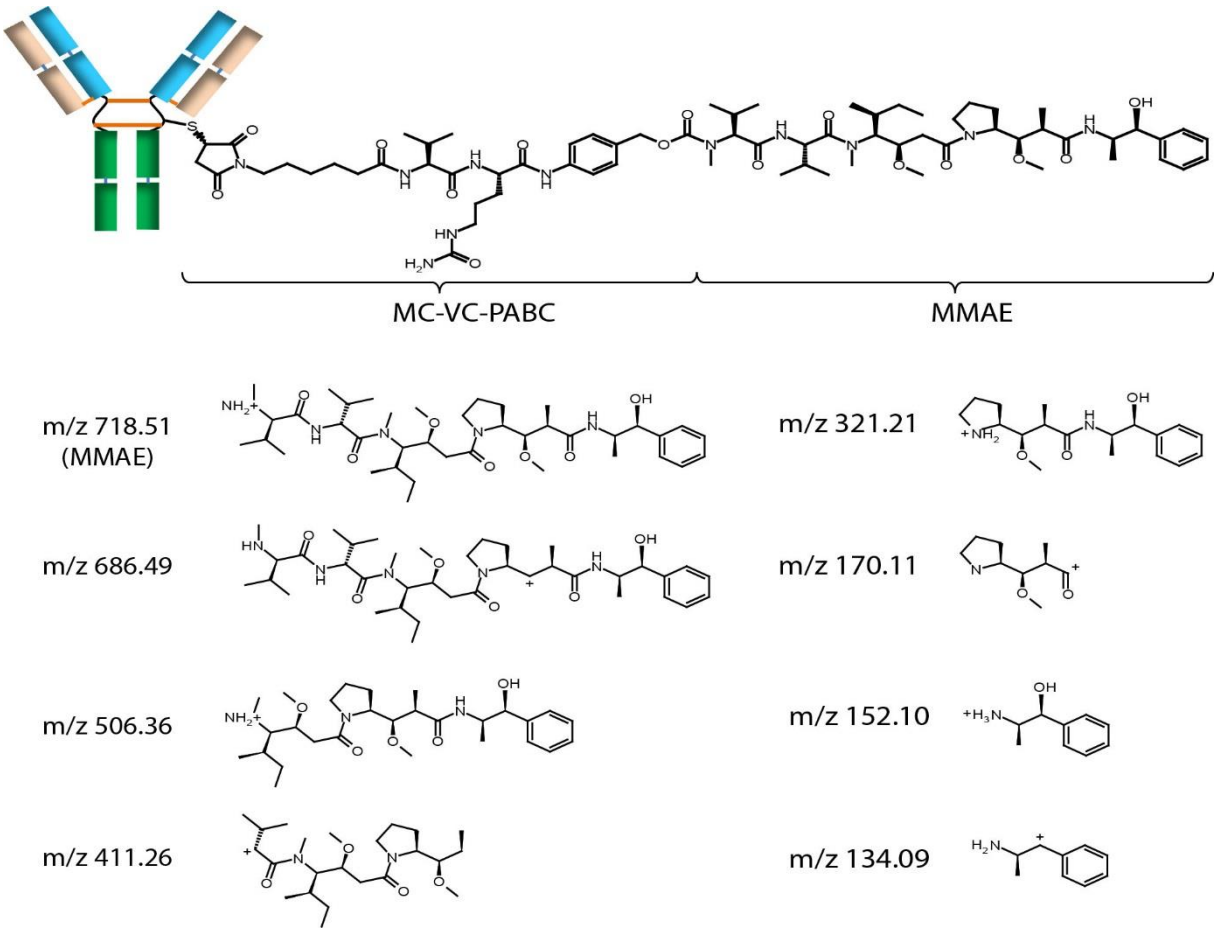

Figure 5. Schematic structures of brentuximab vedotin and characteristic fragments of MMAE payload

Identification of drug-loaded-peptide of BV

The final aim of this work was to propose a methodology using CZE-ESI-MS/MS allowing to obtain structural information of ADCs especially regarding drug-loaded-peptides. The major issue concerning the characterization of drug-loaded-peptide is based on their hydrophobic properties because of the presence of the conjugated drug [24]. Using the protocol we developed in a previous work for the characterization of mAbs based on an aqueous solution trypsin digestion [19], no or very low intensity of these modified peptides was observed due to a possible precipitation during the reduction or digestion steps. Hydrophobic drug-loaded peptides generated from the tryptic digestion of BV could be successfully maintained in solution by incorporating acetonitrile and isopropanol in the sample preparation. These peptides could be successfully identified (Figure 4). Each peptide were characterized based on accurate mass measurement in MS1 provided by high resolution MS and additionally by the characteristic in-

source fragmentation ions (m/z 718.51) described by Junutula et al [38] and also with seven additional diagnostic ions (m/z 686.49, 506.36, 411.26, 321.21, 170.11, 152.10, 134.09) observed in all MS/MS spectra of molecules containing MMAE (Figure 4). As glycopeptide characterization, product ions, observed in MS/MS with low energy CID, mainly correspond to the fragmentation of the MMAE moiety. Very low y ion fragment can be observed but the peptide backbone is usually preserved. However, while glycopep-tide MS/MS spectra allow to detect ions corresponding to fragments of glycopeptide, drug-loaded-peptide MS/MS spectra only allow to detect the fragmentation of the drug. This makes it necessary to have a few signing ions on the MS/MS spectrum to characterize those peptides without ambiguity. The seven diagnostic ions have been identified by a molecular modeling study using Mass Spec Calculator Pro (BCP Instrument, France). Figure 5 shows the structures of MMAE fragments matching with the theoretical masses detected in all drug-loaded-peptides MS/MS spectra.

Brentuximab vedotin drug localization

Peptide	Unmodif	+ 1 payload	+ 2 payload
GEC	31.2 ± 3.4	68.8 ± 3.4	
SCKD	89.8 ± 6.3	10.2 ± 6.3	
THTCPPCPAPELLG	88.8 ± 2.5	10.4 ± 2.2	0.9 ± 0.5

Table 1. Table summarizing the study regarding drug-loaded-peptides occurrence levels for brentuximab vedotin (distribution %intact/%modification) were determined through signal intensity for each targeted drug-loaded-peptide.

Using the same set of data, we identified four different drug-loaded-peptides. Peptides presented a [M+2H] mass of 812.44 (Figure 4-a) and a [M+2H] of 884.99 (Figure 4-b), were corresponding the theoretical masses of the Lc peptide [GEC] and the Hc peptide [SCDK] both conjugated to 1 payload. The peptide corresponding to m/z 918.15 exhibiting a charge state of 3+ (Figure 4-c) was consistent to the Hc peptide [THTCPPCPAPELLG] conjugated to 1 MMAE molecule. Finally, the peptide presented a [M+4H] mass of 1018.06 was compatible the Hc peptide [THTCPPCPAPELLG] conjugated to 2 payloads. For [GEC] and [SCDK], Cys 218 and Cys 220 are involved in the disulfide bond with the Lc peptide and the Hc peptide respectively, while for [THTCPPCPAPELLG] Cys 226 and Cys 229 are involved in the disulfide bonds between the two Hc. Those results confirm the characterization of the expected drug-loaded-peptides linked to the interchain cysteines of the Hc and Lc [17]. The capacity to separate peptides having payload is clearly interesting because they tend to compete against each other during the ionization process, potentially interfering with relative quantification. Using the sheathless CZE-ESI-MS/MS methodology, the m/z ratios corresponding to the native peptide and the peptide incorporating 1 or 2 payloads (Table 1) migrate as two different

peaks separated by several minutes. In the context of characterization of BV, results of relative quantitation allowed to estimate site by site the yield of drug incorporation (Table 1). To evaluate the robustness of this method, the analysis was repeated six times and considered as technical replicates, allowing us to obtain equivalent results. The same data were used for MS/MS peptide mapping, glycoforms and drug-loaded-peptide characterization without requiring additional experiments. Results detailed in this section clearly shows the ability using sheathless CZE-ESI-MS/MS analysis to perform a site specific localization study of the conjugation between the mAb and the drug employed giving additional information required in the development pipeline of ADCs therefore completing the primary structure characterization of this type of therapeutic proteins.

Conclusion

To summarize, we report here the implementation of new analytical approaches based on a sheathless capillary electrophoresis (CE) interface used as nanoESI infusion platform or as separation method, hyphenated to tandem MS for the detailed and rapid structural characterization of a gold standard ADC. Starting with an intact ADC native MS study (top level) and going through middle-up and then bottom-up MS approach, we were able to characterize BV from intact mass measurement to the amino acid sequence level. At the intact ADC level, we evaluated sheathless CE-MS system used as nanoESI infusion platform for native MS determination of the average DAR and drug load distribution of BV. We show native sheathless CE-MS is applicable to interchain cysteine-linked ADCs without any detrimental effect on the noncovalent structure. Within a single run of few minutes and consuming less than of 3µg BV sample, calculated DAR value were in good agreement with already published values. To investigate deeper the characterization of BV, middle-up analysis using IdeS enzymatic digestion has been performed following the same experimental protocol used for intact BV analysis. Structural information of the expected Fc/2 subunits and F(ab')2 subunits with the incorporation of zero to eight payloads have been obtained. Additionally, other subunits as Lc conjugated to one payload and Fab conjugated to zero to four payloads have been observed and characterized. Of note, regarding native MS analysis of Fc/2, results showed the formation of Fc/2 non-covalent dimers following BV IdeS treatment. To the best of our knowledge, this is the first time that non-covalent dimers of ADC Fc/2 subunits have been observed and characterized. Finally, we report here the use of sheathless CZE-ESI-MS/MS for the bottom-up analysis of BV to enhance the level of characterization, especially regarding amino acid sequence, glycoforms and drug-loaded-peptides. Using a single injection of 200 fmol, we were able to characterize the primary structure of those conjugates in a robust

manner. We managed to perform the complete amino acid sequence characterization. Simultaneously, using the same set of data, glycosylations and drug-loaded-peptides were also characterized. Concerning glycosylation, 11 different N-glycans could be characterized for BV, confirming the advantages of using sheathless CZE-ESI-MS improved sensitivity. Regarding the drug-loaded-peptides, we identified four drug-loaded-peptides linked to the interchain cysteines of the Hc and Lc which gave precious information regarding the localization of the conjugated drugs on the studied ADC and the assessment of the yield of drug incorporation. In order to characterize unambiguously these drug-loaded-peptides, seven diagnostic ions have been identified and structurally characterized as MMAE fragments present in MS/MS fragmentation spectra. The CZE-ESI-MS/MS data reported here indicate that electrophoretic separation, combined to the highly efficient sheathless CE-MS interface become a relevant alternative to LC-ESI-MS/MS for innovative approaches in MS for the characterization of a large panel of highly complex biotherapeutics.

Acknowledgments

Authors would like to thank Sciex separations Inc. for lending a CESI8000 system and Dr. M. Anselme from Sciex separations Inc. for his support. The authors would like also to express their gratitude to Dr. M. Biacchi (LSMIS, Strasbourg, France), P. Hammann and J. Chicher (Institut de Biologie Moléculaire et Cellulaire, Strasbourg, France), Dr. E. Wagner-Rousset, , MC. Janin-Bussat and O. Colas (Centre d'Immunologie Pierre Fabre, St Julien en Genevois, France) for helpful discussions around antibody structural characterization by LC-MS. This work was supported by the CNRS (UMR 7140), the University of Strasbourg and a doctoral fellowship from the General Council of Mayotte to NS.

References

[1] R.V.J. Chari, Targeted cancer therapy: Conferring specificity to cytotoxic drugs, Accounts Chem. Res., 41 (2008) 98-107.

[2] S. Ornes, Antibody-drug conjugates, P. Natl. acad. Sci. USA, 110 (2013) 13695-13695.

[3] P. Sapra, B. Shor, Monoclonal antibody-based therapies in cancer: Advances and Challenges, Pharmacol. Therapeut., 138 (2013) 452-469.

[4] A. Beck, J.M. Reichert, Antibody-drug conjugates, Mabs, 6 (2014) 15-17.

[5] R.S. Zolot, S. Basu, R.P. Million, Antibody-drug conjugates, Nat. Rev. Drug Discov., 12 (2013) 259-260.

[6] A. Beck, J.F. Haeuw, T. Wurch, L. Goetsch, C. Bailly, N. Corvaia, The Next Generation of Antibody-drug Conjugates Comes of Age, Discov. Med., 53 (2010) 329-339.

[7] A. Younes, N.L. Bartlett, J.P. Leonard, D.A. Kennedy, C.M. Lynch, E.L. Sievers, A. Forero-Torres, Brentuximab Vedotin (SGN-35) for Relapsed CD30-Positive Lymphomas, New Engl. J. Med., 363 (2010) 1812-1821.

[8] S. Verma, D. Miles, L. Gianni, I.E. Krop, M. Welslau, J. Baselga, M. Pegram, D.Y. Oh, V. Dieras, E. Guardino, L. Fang, M.W. Lu, S. Olsen, K. Blackwell, E.S. Grp, Trastuzumab Emtansine for HER2-Positive Advanced Breast Cancer, New Engl. J. Med., 367 (2012) 1783-1791.

[9] U. Iyer, V.J. Kadambi, Antibody drug conjugates - Trojan horses in the war on cancer, J. Pharmacol. Toxicol. Meth., 64 (2011) 207-212.

[10] S. Krapp, Y. Mimura, R. Jefferis, R. Huber, P. Sondermann, Structural analysis of human IgG-Fc glycoforms reveals a correlation between glycosylation and structural integrity, J. Mol. Biol., 325 (2003) 979-989.

[11] R. Jefferis, Recombinant antibody therapeutics: the impact of glycosylation on mechanisms of action, Trends Pharmacol. Sci., 30 (2009) 356-362.

[12] A. Beck, F. Debaene, H. Diemer, E. Wagner-Rousset, O. Colas, A.V. Dorsselaer, S. Cianférani, Cutting-edge mass spectrometry characterization of originator, biosimilar and biobetter antibodies, J Mass Spectrom., 50 (2015) 285-297.

[13] M. Biacchi, R. Gahoual, N. Said, A. Beck, E. Leize-Wagner, Y.-N. François, Glycoform Separation and Characterization of Cetuximab Variants by Middle-up Off-Line Capillary Zone Electrophoresis-UV/Electrospray Ionization-MS, Anal. Chem., 87 (2015) 6240-6250.

[14] D.R. Stoll, D.C. Hannes, J. Danforth, E. Wagner, D. Guillarme, S. Fekete, A. Beck, Direct Identification of Rituximab Main Isoforms and Subunit Analysis by Online Selective Comprehensive Two-Dimensional Liquid Chromatography-Mass Spectrometry, Anal. Chem., 87 (2015) 8307-8315.

[15] F. Debaene, A. Boeuf, E. Wagner-Rousset, O. Colas, D. Ayoub, N. Corvaia, A. Van Dorsselaer, A. Beck, S. Cianferani, Innovative Native MS Methodologies for Antibody Drug Conjugate Characterization: High Resolution Native MS and IM-MS for Average DAR and DAR Distribution Assessment, Anal. Chem., 86 (2014) 10674-10683.

[16] A. Beck, E. Wagner-Rousset, D. Ayoub, A. Van Dorsselaer, S. Sanglier-Cianférani, Characterization of Therapeutic Antibodies and Related Products, Anal. Chem., 85 (2013) 715-736.

[17] M.-C. Janin-Bussat, M. Dillenbourg, N. Corvaia, A. Beck, C. Klinguer-Hamour, Characterization of antibody drug conjugate positional isomers at cysteine residues by peptide mapping LC-MS analysis, J. Chromatogr. B, 981 (2015) 9-13.

[18] D. Firth, L. Bell, M. Squires, S. Estdale, C. McKee, A rapid approach for characterization of thiol-conjugated antibody–drug conjugates and calculation of drug–antibody ratio by liquid chromatography mass spectrometry, Anal. Biochem., 485 (2015) 34-42.

[19] R. Gahoual, J.-M. Busnel, A. Beck, Y.-N. François, E. Leize-Wagner, Full Antibody Primary Structure and Microvariant Characterization in a Single Injection Using Transient Isotachophoresis and Sheathless Capillary Electrophoresis–Tandem Mass Spectrometry, Anal. Chem., 86 (2014) 9074-9081.

[20] R. Gahoual, M. Biacchi, J. Chicher, L. Kuhn, P. Hammann, A. Beck, E. Leize-Wagner, Y.N. Francois, Monoclonal antibodies biosimilarity assessment using transient isotachophoresis capillary zone electrophoresis-tandem mass spectrometry, Mabs, 6 (2014) 1464-1473.

[21] J.A. Francisco, C.G. Cerveny, D.L. Meyer, B.J. Mixan, K. Klussman, D.F. Chace, S.X. Rejniak, K.A. Gordon, R. DeBlanc, B.E. Toki, C.L. Law, S.O. Doronina, C.B. Siegall, P.D. Senter, A.F. Wahl, cAC10-vcMMAE, an anti-CD30-monomethyl auristatin E conjugate with potent and selective antitumor activity, Blood, 102 (2003) 1458-1465.

[22] E. Oflazoglu, K.M. Kissler, E.L. Sievers, I.S. Grewal, H.P. Gerber, Combination of the anti-CD30-auristatin-E antibody-drug conjugate (SGN-35) with chemotherapy improves antitumour activity in Hodgkin lymphoma, Brit. J. Haematol., 142 (2008) 69-73.

[23] J.F. Valliere-Douglass, W.A. McFee, O. Salas-Solano, Native Intact Mass Determination of Antibodies Conjugated with Monomethyl Auristatin E and F at Interchain Cysteine Residues, Anal. Chem., 84 (2012) 2843-2849.

[24] J. Chen, S. Yin, Y. Wu, J. Ouyang, Development of a Native Nanoelectrospray Mass Spectrometry Method for Determination of the Drug-to-Antibody Ratio of Antibody–Drug Conjugates, Anal. Chem., 85 (2013) 1699-1704.

[25] R. Gahoual, J.M. Busnel, P. Wolff, Y.N. Francois, E. Leize-Wagner, Novel sheathless CE-MS interface as an original and powerful infusion platform for nanoESI study: from intact proteins to high molecular mass noncovalent complexes, Anal. Bioanal. Chem., 406 (2014) 1029-1038.

[26] B. Wang, A.C. Gucinski, D.A. Keire, L.F. Buhse, M.T. Boyne, Structural comparison of two anti-CD20 monoclonal antibody drug products using middle-down mass spectrometry, Analyst, 138 (2013) 3058-3065.

[27] Z.Q. Zhang, H. Pan, X.Y. Chen, mass spectrometry for structural characterization of therapeutic antibodies, Mass Spectrom. Rev., 28 (2009) 147-176.

[28] A. Wakankar, Y. Chen, Y. Gokarn, F.S. Jacobson, Analytical methods for physicochemical characterization of antibody drug conjugates, mAbs, 3 (2011) 161-172.

[29] S. Fekete, D. Guillarme, P. Sandra, K. Sandra, Chromatographic, Electrophoretic, and Mass Spectrometric Methods for the Analytical Characterization of Protein Biopharmaceuticals, Anal. Chem., (2016) 10.1021/acs.analchem.5b04561.

[30] Z. Kalman-Szekeres, M. Olajos, K. Ganzler, Analytical aspects of biosimilarity issues of protein drugs, J. Pharm. Biomed. Anal., 69 (2012) 185-195.

[31] L. Fornelli, D. Ayoub, K. Aizikov, A. Beck, Y.O. Tsybin, Middle-Down Analysis of Monoclonal Antibodies with Electron Transfer Dissociation Orbitrap Fourier Transform Mass Spectrometry, Anal. Chem., 86 (2014) 3005-3012.

[32] A. Beck, G. Terral, F. Debaene, E. Wagner-Rousset, J. Marcoux, M.-C. Janin-Bussat, O. Colas, A. Van Dorsselaer, S. Cianférani, Cutting-edge mass spectrometry methods for the multi-level structural characterization of antibody-drug conjugates, Expert Rev. Proteomics, (2015) 10.1586/14789450.2016.1132167.

[33] A. Resemann, W. Jabs, A. Wiechmann, E. Wagner, O. Colas, W. Evers, E. Belau, L. Vorwerg, C. Evans, A. Beck, D. Suckau, Full validation of therapeutic antibody sequences by middle-up mass measurements and middle-down protein sequencing, mAbs, (2016) 10.1080/19420862.2015.1128607.

[34] E.A. Redman, J.S. Mellors, J.A. Starkey, J.M. Ramsey, Characterization of Intact Antibody Drug Conjugate Variants Using Microfluidic Capillary Electrophoresis–Mass Spectrometry, Anal. Chem., (2016) 10.1021/acs.analchem.5b03866.

[35] K. Wenig, L. Chatwell, U. von Pawel-Rammingen, L. Bjorck, R. Huber, P. Sondermann, Structure of the streptococcal endopeptidase IdeS, a cysteine proteinase with strict specificity for IgG, P. Natl. acad. Sci. USA, 101 (2004) 17371-17376.

[36] S. Flatman, I. Alam, J. Gerard, N. Mussa, Process analytics for purification of monoclonal antibodies, J. Chromatogr. B, 848 (2007) 79-87.

[37] R. Gahoual, A. Burr, J.M. Busnel, L. Kuhn, P. Hammann, A. Beck, Y.N. Franccois, E. Leize-Wagner, Rapid and multi-level characterization of trastuzumab using sheathless capillary electrophoresis-tandem mass spectrometry, Mabs, 5 (2013) 479-490.

[38] J.R. Junutula, H. Raab, S. Clark, S. Bhakta, D.D. Leipold, S. Weir, Y. Chen, M. Simpson, S.P. Tsai, M.S. Dennis, Y. Lu, Y.G. Meng, C. Ng, J. Yang, C.C. Lee, E. Duenas, J. Gorrell, V. Katta, A. Kim, K. McDorman, K. Flagella, R. Venook, S. Ross, S.D. Spencer, W. Lee Wong, H.B. Lowman, R. Vandlen, M.X. Sliwkowski, R.H. Scheller, P. Polakis, W. Mallet, Site-specific conjugation of a cytotoxic drug to an antibody improves the therapeutic index, Nat. Biotech., 26 (2008) 925-932.

3. Conclusion

Dans ce chapitre, nous avons appliqué une approche globale mettant en œuvre l'utilisation du couplage CESI-MS en tant que plateforme d'infusion nanoESI et méthode séparative afin obtenir la caractérisation structurale à différents niveaux du brentuximab vedotin (BV). Le premier niveau de caractérisation a consisté à évaluer le couplage CESI-MS en tant que plateforme nanoESI en MS native pour la détermination du DAR moyen du BV et pour obtenir la distribution de conjugués sur le BV. Nous avons démontré que l'utilisation de l'interface CESI-MS est totalement compatible pour l'infusion en condition native d'ADCs de type « cycteine-linked » sans effet destructeur sur la structure non-covalente du BV. En une seule infusion MS de quelques minutes représentant moins de 3 µg d'échantillon, nous avons déterminé une valeur du DAR moyen de 3,9 ± 0,1 totalement en accord avec la littérature. Ensuite, dans le but d'approfondir la caractérisation du BV en utilisant le même système, nous avons réalisé une approche middle-up par digestion IdeS. Des informations structurales attendues sur les fragments Fc/2 et F(ab')2 avec l'incorporation de zéro à huit conjugués ont été obtenues. De plus, des informations de structure complémentaires concernant la chaine légère conjuguée à un médicament et le fragment Fab avec l'incorporation de zéro à quatre médicaments, ont été observées et caractérisées. A noté, nous avons confirmé les résultats décrit dans la partie III (II - Caractérisation de dimères de fragments Fc/2 de cetuximab par Off-line CZE-UV/ESI-MS) sur la mise en évidence d'interactions non-covalentes expliquant la formation de dimères de Fc/2. Enfin, une méthode séparative CESI-MS basée sur une approche bottom-up protéomique a été développée dans le but d'améliorer le niveau de caractérisation du BV particulièrement au niveau de la séquence d'acides aminés, du profil de glycosylations et des peptides liés aux médicaments. En analysant les résultats obtenus sur une injection de 200 fmol d'un digeste de BV, nous avons été capables de caractériser la totalité de la séquence d'acide aminé du BV. Sur le même set de données, nous avons aussi pu caractériser un grand nombre de glycosylations et la totalité des peptides liés aux médicaments. En effet, 11 N-glycosylations différentes ont été caractérisées pour le BV confirmant les performances du couplage CESI-MS sur la caractérisation de glycoprotéines thérapeutiques (Partie II : Caractérisation de la structure primaire d'anticorps monoclonaux et évaluation de biosimilarité à l'aide du couplage CESI-MS). Concernant les peptides liés aux médicaments, nous avons identifiés quatre peptides modifiés correspondant aux peptides portant une ou plusieurs cystéines inter-chaines du BV. Cette identification permet d'obtenir des informations sur la localisation des médicaments sur le BV ainsi que des estimations de rendement de réaction de l'incorporation de la drogue par site de modifications. De plus, dans le but d'identifier sans aucune ambiguïté la structure des peptides liés à un médicament, nous

avons caractérisé sept ions diagnostiques correspondant à des fragments de MMAE sur les spectres MS/MS. Cette étude confirme donc que le couplage CZE-ESI-MS mettant en jeu l'interface CESI, devient une bonne alternative au couplage LC-ESI-MS pour la caractérisation de protéines thérapeutiques complexes.

Projets de Recherche

Nos projets de recherche se situent en prolongation directe des résultats que nous avons obtenus jusqu'à présent. Ils consistent notamment en l'amélioration et l'exploitation maximale des possibilités du couplage de l'électrophorèse capillaire et de la spectrométrie de masse pour l'étude de protéines thérapeutique ainsi que pour d'autres applications tels que l'étude de protéines exprimées par des organismes dans le cadre de l'analyse protéomique classique ou la caractérisation de modifications post-transcriptionnelles des ARNs.

1. Caractérisation de protéines entières par enrichissement d'échantillon et stratégie Top Down à l'aide du Couplage CE-UV/MALDI-MS

Nous avons développé et évalué le premier couplage CE-UV/MALDI-MS/MS à interface coaxiale avec système de distribution et de dépôt de matrice entièrement automatisé. Grâce à sa conception robuste et entièrement paramétrable, nous avons démontré la répétabilité et la polyvalence du couplage sur une séparation de protéines entières. De plus, l'analyse de mAbs entiers a confirmé la possibilité d'utiliser le système CE-UV/MALDI-MS/MS pour la séparation de protéines de haut poids moléculaire. Cependant, l'absence de résolution du MALDI-MS apparait comme une limitation importante à la caractérisation fine de protéines par cette méthode. Partis de ce constat et des alternatives en termes de stratégie pour l'analyse de protéines entières, nous avons le projet de réaliser la séparation et la caractérisation de protéines entières par une stratégie Top Down protéomique.

Aujourd'hui, la caractérisation d'une protéine en MS peut se faire selon deux approches complémentaires : une approche dite Bottom Up et une approche dite Top Down. L'approche Bottom Up largement décrite dans les chapitres précédents est à ce jour la méthode la plus utilisée en protéomique. Elle consiste à soumettre une digestion protéolytique aux protéines puis d'identifier par spectrométrie de masse les peptides formés. Au cours de ces dernières années, l'approche Bottom Up a pris beaucoup d'ampleur grâce aux nouveaux enzymes de digestion de plus en plus spécifiques et efficaces et grâce aux avancées technologiques des analyseurs de masse. Toutefois, les différentes étapes nécessaires à la protéolyse sont susceptibles de conduire à des pertes d'informations. C'est pourquoi des stratégies d'analyses de protéines entières, ont vu le jour réduisant les étapes de préparation de l'échantillon. C'est pour cette raison que l'approche Top Down s'est développée. Elle consiste à caractériser une ou plusieurs protéines entières sans digestion préalable. Grâce à la fragmentation directe de la protéine entière par MS, il est possible d'obtenir des spectres de masse de la protéine fragmentée d'acides aminés en acides aminés. Et c'est à partir de l'analyse des spectres de fragmentation que l'on peut remonter à la séquence de la protéine, et aux éventuelles modifications post traductionnelles. De nos jours, la stratégie Top Down connait un développement considérable notamment dans la recherche et la caractérisation de biomarqueurs, de micro-organismes ou d'anticorps monoclonaux. La plupart de ces études mettent en jeu des spectromètres de masse à sources d'ionisation ESI pour l'analyse Top Down. Cependant, lors d'analyses de mélanges complexes, la production d'ions multi chargés complexifie les spectres MS/MS. C'est la raison pour laquelle la source MALDI et sa production d'ions majoritairement mono chargés est plus aisée à mettre en œuvre pour la stratégie Top Down des protéines entières. Néanmoins, cette stratégie nécessite une grande pureté et une haute quantité (> 1 pM) d'analytes.

Le laboratoire a démontré la capacité du couplage CE-MALDI-MS à séparer des anticorps intacts avec une haute résolution. C'est cette résolution qui permet de séparer des analytes avec une grande pureté. Cependant, concernant la quantité de l'échantillon déposé, la CE peut difficilement fournir des fractions concentrées. En effet, dues aux propriétés de diffusion en front plat des zones d'échantillons, la capacité de chargement d'un capillaire ne peut excéder 1 % du volume total. Les volumes d'échantillons communément injectés ne dépassent donc pas 10 nL. Dans le but de trouver une alternative à l'augmentation du volume d'échantillon, des techniques de préconcentrations sont utilisées. Cependant dans certains cas, ces méthodes ne sont pas suffisantes pour la stratégie Top Down par fragmentation ISD. C'est parti de ce postulat que nous proposons d'utiliser la nouvelle interface CE/MALDI-MS afin de collecter et ainsi de faire de l'enrichissement de fractions. En effet la répétabilité tant dans la séparation que dans la collecte et le dépôt de fraction nous permettra de « purifier » les protéines d'intérêt tout en les concentrant par dépôts successifs.

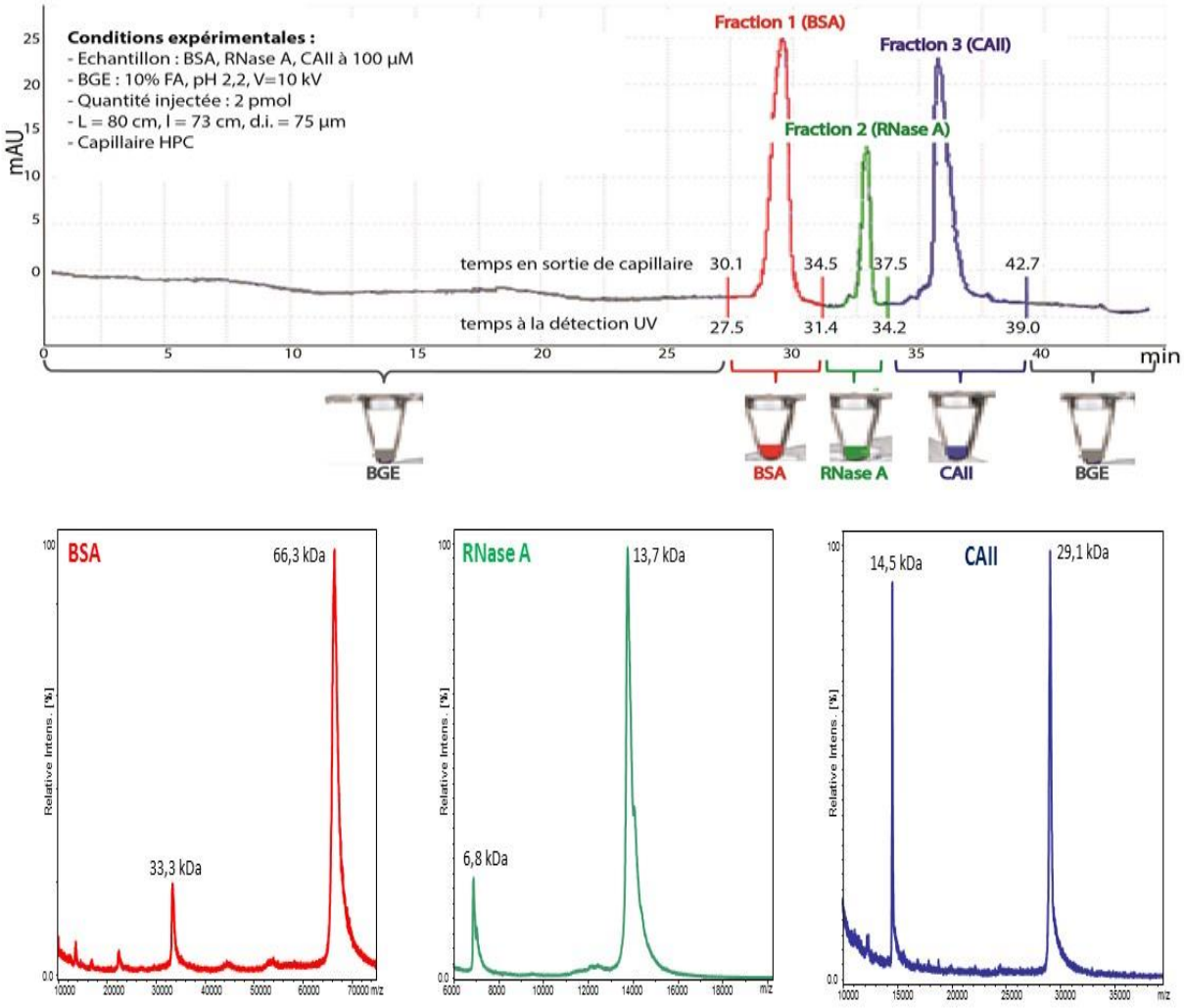

Illustration du processus de collection de fractions pour le mélange de 3 protéines entières

2. Développement du couplage Electrophorèse Capillaire-Spectrométrie de Masse pour la caractérisation des modifications post-transcriptionnelles des ARNs

Les ARNs sont impliqués dans de nombreux processus biologiques. La plupart des ARNs contiennent des modifications post-transcriptionnelles (dont les ARNs de transferts et ribosomiques) qui facilitent ou régulent leurs fonctions. Ces modifications sont extrêmement variées et dépendant de leur localisation, peuvent induire de multiples effets en modulant la structure des ARN, leur stabilité, et leur localisation cellulaire. Elles exercent en outre une fonction clé dans la traduction et le décodage de l'ARNm. Enfin ces modifications sont des signaux/signatures qui confèrent la résistance aux antibiotiques, la réponse au stress, et sont impliquées dans de nombreuses pathologies humaines. Cependant, la caractérisation des modifications dans l'ARN (144 connues) reste à ce jour un défi technique, et la spectrométrie de masse est un moyen de le relever.

Aujourd'hui, les approches de spectrométrie de masse classiques telles que la nanoLC-MSMS génèrent des informations sur la séquence des ARN, la localisation et l'identification des modifications post-transcriptionnelles et sur le niveau d'expression des ARN. A terme, comme pour les protéines, la spectrométrie de masse devrait conduire à une analyse de type RNomique complémentaire aux approches de séquençage haut débit. Les approches classiques par nanoLC-MSMS ont l'avantage d'être très standardisées et robustes, mais ne permettent pas encore d'accéder à l'ensemble des informations nécessaires à la caractérisation complète des protéines et des ARN. Dans ce contexte, notre laboratoiree 1 projette d'utiliser le couplage CESI-MSMS, en complément à l'approche nanoLC-MSMS, pour la caractérisation des modifications post transcriptionnelles des ARNs.

L'objectif principal de ce projet sera de développer et d'appliquer cette nouvelle stratégie CESI-MSMS pour identifier et quantifier les nucléotides modifiés des ARNs en complément à l'approche nanoLC-MSMS. Nous utiliserons comme système modèle, les ARNs non codant (ARNnc) de *Staphylococcus aureus*, développés en collaboration avec l'équipe du Dr. Pascale Romby, DR 1ère CNRS et directrice de l'équipe «ARN messagers et ARN régulateurs bactériens » de l'IBMC à l'université de Strasbourg.

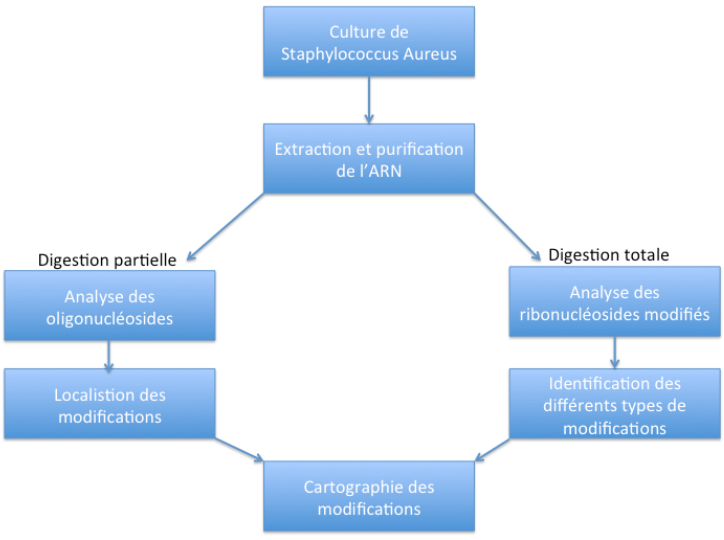

Schéma de l'approche conceptuelle et scientifique

Curriculum Vitae

DR. YANNIS FRANCOIS

Adresse personnelle : 2 Boulevard de Lyon
67000 Strasbourg - FRANCE
Date de naissance : 21 Janvier 1978
Tel.: 06.88.49.88.49

Adresse laboratoire : 1 rue Blaise Pascal
67008 Strasbourg Cedex – FRANCE
Tel.: 03.68.85.16.41
E-mail: yfrancois@unistra.fr

I. Formation

- ***Maître de Conférences*** – Janv. 2013,
  - Laboratoire de Spectrométrie de Masse des Interactions et des Systèmes, CNRS-UMR 7140, Chimie de la Matière Complexe, **Université de Strasbourg**
  - Directeur : Dr. E. LEIZE-WAGNER

- ***Maître de Conférences*** – Sept. 2007 – Déc. 2012
  - Laboratoire de Dynamique et Structure Moléculaire et par Spectrométrie de Masse, CNRS-UMR 7177, Institut de Chimie, **Université de Strasbourg**
  - Directeur : Dr. E. LEIZE-WAGNER

- ***Attaché temporaire d'enseignement et de recherche*** **(ATER)** – Sept. 2006 – Sept. 2007
  - Laboratoire Chimie et Biologie Pharmacologiques et Toxicologiques, CNRS-UMR 8601, Equipe Physico-Chimie des espèces radicalaires, **Université Paris Descartes**
  - Directeur : Prof. M. GARDES-ALBERT

- ***Doctorant es Sciences,*** **spécialité Chimie Analytique** – Sept. 2003 – Nov. 2006
  - Laboratoire d'Electrochimie et Chimie Analytique – **Ecole Nationale Supérieure de Chimie Paris (ENSCP)**
  - Directeur : Prof. P. GAREIL, Dr. A. VARENNE

	Publication	***Proceeding***	***Conf. Invitée***	***Conférence***	***Séminaire***	***Poster***
International	**18**	**9**	**6**	**7**	**16**	**23**
National			**5**	**5**	**9**	**8**

II. Publications avec comité de lecture

1. **New integrated measurement protocol using capillary electrophoresis instrumentation for the determination of viscosity, conductivity and absorbance of ionic liquid-molecular solvent mixture**

YN. Francois, K. Zhang, A. Varenne, P. Gareil, *Analytica Chimica Acta* 2006, 562, 164-170.

2. **Nonaqueous capillary electrophoretic behavior of 2-arylpropionic acids in the presence of acids in the presence of an achiral ionic liquid: a chemometric approach**
   YN. Francois, A. Varenne, E. Juillerat, A-C. Servais, P. Chiap, P. Gareil, *Journal of Chromatography A* 2007, 1138, 268-275.
3. **Evaluation of chiral ionic liquids as additives to cyclodextrins for enantiomeric separations by capillary electrophoresis**
   YN. Francois, A. Varenne, E. Juillerat, D. Villemin, P. Gareil, *Journal of Chromatography A* 2007, 1155(2), 134-141.
4. **Determination of aqueous inclusion complexation constants and stoichiometry of alkyl(methyl)methylimidazolium-based ionic liquids cations and neutral cyclodextrins by affinity capillary electrophoresis**
   YN. Francois, A. Varenne, J. Siriex, P. Gareil, *Journal of Separation Science* 2007, 30, 751-760.
5. **Rapid and multi-level characterization of trastuzumab using sheathless capillary electrophoresis-tandem mass spectrometry**
   R. Gahoual, A. Burr, JM. Busnel, L. Kuhn, P. Hammann, A. Beck, YN. François, E. Leize-Wagner, *mAbs* 2013, 5(3), 479-490.
6. **Novel sheathless CE-MS interface used as an original and powerful infusion platform for nanoESI study: from intact proteins to high molecular mass noncovalent complexes**
   R. Gahoual, JM Busnel, P. Wolff, YN. François, E. Leize-Wagner, *Analytical and Bioanalytical Chemistry* 2014, 406, 1029-1038
7. **Analysis of monoclonal antibody by a novel CE-UV/MALDI-MS interface**
   M. Biacchi, R. Bhajun, N. Saïd, YN. François, E. Leize-Wagner, *Electrophoresis* 2014, 35(20), 2986-2995.
8. **Full antibody primary structure and micro-variant characterization in a single injection using transient isotachophoresis and sheath-less capillary electrophoresis – tandem mass spectrometry**
   R. Gahoual, JM. Busnel, A. Beck, YN. François, E. Leize-Wagner, *Analytical Chemistry* 2014, 86(18), 9074-9081
9. **Monoclonal antibodies biosimilarity assessment using transient isotachophoresis capillary zone electrophoresis-tandem mass spectrometry**
   R. Gahoual, M. Biacchi, J. Chicher, L. Kuhn, P. Hammann, A. Beck, E. Leize-Wagner, YN. François, *mAbs* 2014, 6(6), 1464-1473.
10. **Glycoform separation and characterization of Cetuximab by Middle-up Off-line CZE-UV/ESI-MS**
    M Biacchi, R Gahoual, N Saïd, A Beck, E. Leize-Wagner, YN François, *Analytical Chemistry* 2015, 87(12), 6240-6250
11. **Priming the Cells for Their Final Destination: Microenvironment Controlled Cell Culture by a Modular ECM-mimicking Feeder Film**

J. Barthes, E. Vrana, H. Ozcelik, R. Gahoual, Y. Francois, J. Bacharouche, G. Francius, J. Hemmerlé, M-H. Metz-Boutigue, P. Schaaf, P. Lavalle, *Biomaterials Science* 2015, 3(9), 1302-1311.

12. **Comprehensive multi-level characterization of biologics using sheathless capillary electrophoresis hyphenated to tandem mass spectrometry**
    R. Gahoual, JM. Busnel, A. Beck, YN. François, E. Leize-Wagner, *LCGC North America* 2015, 13(4), 8-15.
13. **Improvement of mitochondria extract from Saccharomyces cerevisiae characterization in shotgun proteomics using sheathless capillary electrophoresis coupled to tandem mass spectrometry**
    M. Ibrahim, R. Gahoual, L Enkler, H. Becker, J. Chicher, P. Hammann, Y.N. Francois, L. Kuhn, E. Leize-Wagner, *Journal of Chromatographic Science* 2016, 54(4), 653-663
14. **Independent highly sensitive characterization of asparagine deamidation and aspartic acid isomerization by sheathless CZE-ESI-MS/MS**
    R. Gahoual, A. Beck, Y-N. Francois, E. Leize-Wagner, *Journal of Mass Spectrometry* 2016, 51(2), 150-158
15. **Characterization of cetuximab Fc/2 dimers by off-line CZE-MS**
    Y.N. Francois, M. Biacchi, N. Said, C. Renard, A. Beck, R. Gahoual, E. Leize-Wagner, *Analityca Chimica Acta* 2016, 908, 168-176.
16. **Iminosugar-Cyclopeptoid Conjugates Raise Multivalent Effect in Glycosidase Inhibition at Unprecedented High Levels**
    ML. Lepage, JP. Schneider, A. Bodlenner, A. Meli, F. De Riccardis, M. Schmitt, C. Tarnus, NT. Nguyen-Huynh, YN. Francois, E. Leize-Wagner, C. Birck, A. Cousido-Siah, A. Podjarny, I. Izzo, P. Compain, *Chemistry-A European Journal* 2016, 22, 5151-5155.
17. **Structural characterization of antibody drug conjugate by a combination of intact, middle-up and bottom-up techniques using sheathless capillary electrophoresis – tandem mass spectrometry as nanoESI infusion platform and separation method**
    N. Said, R. Gahoual, L. Kuhn, A. Beck, YN. François, E. Leize-Wagner, *Analytica Chimica Acta*, 2016, 28, 50-59
18. **Cutting-edge capillary electrophoresis characterization of monoclonal antibodies and related products**
    R. Gahoual, Beck, E. Leize-Wagner, YN. François, *Journal of Chromatography B*, 2016, 28, On line available

III. Communication

Conférence invitée international

1. CESI-MS as a flexible platform: from nanoESIemitter for noncovalentcomplexes characterization to peptide mapping of monoclonal antibodies
   R. Gahoual, **Y. N. Francois**, E. Leize-Wagner
   19ème International Symposium ITP2012, 30 septembre - 3 octobre 2012, Baltimore

2. Characterization of monoclonal antibodies by mass spectrometry Contribution of CE-MS
R. Gahoual, Y. N. Francois, E. Leize-Wagner
29th International Symposium of MicroScale Bioseparations, 10-14 mars 2013, Charlottesville

3. Simultaneous peptide mapping, posttranslational modifications and major N-glycosylation characterization of trastuzumab by sheathless capillary electrophoresis - tandem mass spectrometry
Rabah Gahoual, Alicia Burr, Jean-Marc Busnel, Lauriane Kuhn, Philippe Hamman, Alain Beck, **Yannis-Nicolas François**, Emmanuelle Leize-Wagner
CE Pharm 13, 6-13 octobre 2013, Arlington, USA

4. Contribution of sheathless capillary electrophoresis coupled to mass spectrometry (CESI-MS) for proteomic analysis: application to the study of yeast mitochondrial proteome
Rabah Gahoual, Jean-Marc Busnel, Johanna Chicher, Lauriane Kuhn, Phillipe Hammann, **Yannis Nicolas François**, Emmanuelle Leize-Wagner
International EUPA 2013, 14-17 octobre 2013, Saint-Malo, France

5. Complete primary structure and biosimilarity assessment of monoclonal antibodies in a single analysis using CESI-MS/MS
Rabah Gahoual, Jean-Marc Busnel, Alain Beck, Yannis Nicolas François, Emmanuelle Leize-Wagner
IMSC, 24-29 aout 2014, Geneva, Suisse

6. Multilevel Characterization of Biotherapeutics using CESI-MS: from Intact Protein to Peptide Mapping Approach
Yannis Nicolas Francois
International Biopharmaceutical Symposium, 15 octobre 2015, Billerica, USA

Conférence nationale invitée

1. CESI – MS: A major advance in CE - MS coupling
R. Gahoual**, Y. N. Francois**, E. Leize-Wagner
Journée Scientifique SFSM-SFEAP, 25 mars 2013, Paris, France

2. CESI – MS: A major advance in CE - MS coupling
R. Gahoual, Y. N. Francois, E. Leize-Wagner
Club de Chromatographie du Val de Seine, 15 novembre 2012, Rouen

3. Capillary Electrophoresis-Mass Spectometry for the Analysis of Proteins and PTMs characterization
R. Gahoual, Y.N. Francois, **E. Leize-Wagner**
Club de Chromatographie du Val de Seine, 12 novembre 2014, Rouen

4. Capillary Electrophoresis-Mass Spectometry for the Analysis of Biotherapeutic Proteins
Y. N. Francois
European Pharmcopoeia, 24 mars 2015, Strasbourg

5. CE-MS: An Important Method for Protein and Post Translational Modification Characterization
Y. N. Francois
SMAP 2015, 15-18 septembre 2015, Ajaccio

Séminaire international invité

1. Simultaneous peptide mapping, postranslational modifications and major N-glycosylation characterization of therapeutics mAbs by sheathless CESI-MS/MS - Comparison to nanoLC-MS/MS
R. Gahoual, **Y. N. Francois**, E. Leize-Wagner
Netherland Belgium CE-User meeting, 26 avril 2013, Breda, Holland

2. Simultaneous peptide mapping, postranslational modifications and major N-glycosylation characterization of therapeutics mAbs by sheathless CESI-MS/MS - Comparison to nanoLC-MS/MS
R. Gahoual, **Y. N. Francois**, E. Leize-Wagner
Swiss German CE-User meeting, 16-17 septembre 2013, Bâle, Suisse

3. Simultaneous peptide mapping, posttranslational modifications and major N-glycosylation characterization of trastuzumab by sheathless capillary electrophoresis - tandem mass spectrometry
Rabah Gahoual, Alicia Burr, Jean-Marc Busnel, Lauriane Kuhn, Philippe Hamman, Alain Beck, **Yannis-Nicolas François**, Emmanuelle Leize-Wagner
CE User meeting, 16-17 septembre, Basel, Suisse

4. Monoclonal antibodies complete primary structure and biosimilarity assessment in a single analysis using CESI-MS-MS
   Yannis Nicolas Francois
   Live Webcast LCGC, 4 juin 2014

5. Full antibody primary structure and microvariant characterization in a single injection using sheathless capillary electrophoresis–tandem mass spectrometry (CESI-MS)
   Rabah Gahoual, Jean-Marc Busnel, Alain Beck, **Yannis Nicolas François**, Emmanuelle Leize-Wagner
   CE Seminar, 23 octobre 2014, Namur, Belgique

6. Full antibody primary structure and microvariant characterization in a single injection using sheathless capillary electrophoresis–tandem mass spectrometry (CESI-MS)
   Yannis Nicolas François
   Live Webcast, 16 septembre 2014

7. Full antibody primary structure and microvariant characterization in a single injection using sheathless capillary electrophoresis–tandem mass spectrometry (CESI-MS)
   Rabah Gahoual, Jean-Marc Busnel, Alain Beck, **Yannis Nicolas François**, Emmanuelle Leize-Wagner
   Novartis Seminar, 15 octobre 2014, Basel, Suisse

8. Rapid and multi-level characterization of monoclonal antibody through CESI-MS workflow
   Y.N. Francois
   61ème International ASMS Conference, 9-13 juin 2013, Minneapolis, USA

9. Characterization of yeast mitochondrial proteome through CESI-MS workflow
   Y.N. Francois
   61ème International ASMS Conference, 9-13 juin 2013, Minneapolis, USA

10. Full antibody primary structure and microvariant characterization in a single injection using sheathless capillary electrophoresis–tandem mass spectrometry (CESI-MS)
    Rabah Gahoual, Jean-Marc Busnel, Alain Beck, **Yannis Nicolas François**, Emmanuelle Leize-Wagner
    1st CESI Seminar, 20 septembre 2014, Strasbourg, France

11. Sheathless capillary electrophoresis: a novel infusion platform for nanoESI study from intact proteins to high molecular mass noncovalent complexes
    R. Gahoual, Y. N. Francois, E. Leize-Wagner
    Workshop of the German-French PhD College "Membranes and Membrane Proteins" UFA 0407, 30 novembre 2014, Strasbourg

12. Characterization of Biotherapeutic Proteins using CE-MS
    Yannis Nicolas Francois
    Live Webcast Sepscience.com, 2 septembre 2015

13. Multilevel Characterization of Biotherapeutics using CESI-MS: from Intact Protein to Peptide Mapping Approach
    Yannis Nicolas Francois
    2nd CESI Seminar, 2 octobre 2015, Strasbourg, France

14. Biosimilarity assessment of monoclonal antibodies using sheathless CESI-MS
    Yannis Nicolas Francois
    Live Webcast LCGC, 29 février 2016

15. Structural Characterization of Antibody Drug Conjugate using sheathless CESI-MS
    Yannis Nicolas Francois
    Live Webcast LCGC, 16 mars 2016

16. Cutting-edge capillary electrophoresis characterization of therapeutical monoclonal antibodies and related products
    Yannis Nicolas Francois
    Live Webcast The Medecine Maker, 29 Juin 2016

Séminaire national invité

1. Intérêt du couplage « Off line » de l'électrophorèse capillaire (CE) avec la spectrométrie de masse (MS) à source d'ionisation laser assistée par matrice (MALDI). Application à la protéomique.
   M. Biacchi, N Pourhassan, **Y. Francois**, E. Leize
   17ème journée d'électrophorèse capillaire, 14 juin 2011, Paris

2. État de l'art de la CESI-MS dans le cadre de l'analyse de mélanges peptidiques complexes
   Y. Francois, E.Leize-Wagner, 29 novembre 2011, Paris

3. Enrichment by fraction collection developed for CE/MALDI-MS to analyze proteins by Top Down strategy
   M. Biacchi, R. Bhajun, Y.N. François, E. Leize-Wagner
   18ème journée d'électrophorèse capillaire, 9 octobre 2012, Paris

4. CESI-MS as a flexible platform: from nanoESIemitter for noncovalentcomplexes characterization to peptide mapping of monoclonal antibodies
   R. Gahoual, Y. N. Francois, E. Leize-Wagner
   18ème journée d'électrophorèse capillaire, 9 octobre 2012, Paris

5. CESI-MS as a flexible platform: from nanoESIemitter for non-covalent complexes characterization to peptide mapping of monoclonal antibodies
   R. Gahoual, Y. N. Francois, E. Leize-Wagner
   Journée de spectrométrie de masse Bruker, 27 novembre 2012, Paris

6. Characterization of monoclonal antibodies by mass spectrometry: Contribution of CE-MS
   R. Gahoual, **Y. N. Francois**, E. Leize-Wagner
   Séminaire Pierre Fabre, 19 mars 2013, Saint-Julien en Genevois, France

7. Sheathless Capillary Electrophoresis – Mass Spectrometry coupling as a flexible platform: From nanoESI infusion device for noncovalent complexes characterization to peptide mapping of monoclonal antibodies
   R. Gahoual, L. Kuhn, P. Hammann, **Y.N. Francois**, E.Leize-Wagner
   Séminaire IPHC, 23 mai 2013, Strasbourg, France

8. Overview of CE-MS coupling for proteomic and intact protein applications
   Y.N. Francois
   CJ SFEAP, 29-31 mai 2013, Strasbourg, France

9. CE-MS: de l'analyse protéomique à la protéine entière
   Y.N. Francois
   Journée de spectrométrie de masse Bruker, 25 novembre 2014, Paris

Conférence internationale orale

1. Multi-level characterization of trastuzumabusing sheathless capillary electrophoresis-tandem mass spectrometry
   Rabah Gahoual, Alicia Burr, Jean-Marc Busnel, Lauriane Kuhn, Philippe Hamman, Alain Beck, Yannis-Nicolas François, Emmanuelle Leize-Wagner
   MSB 2013, 10-14 mars 2013, Charlottesville, Virginie, USA

2. Simultaneous Peptide Mapping, Posttranslational Modifications and Major N-Glycosylation Characterization of Trastuzumab by Sheathless Capillary Electrophoresis-Tandem Mass Spectrometry
   Rabah Gahoual, Alicia Burr, Jean-Marc Busnel, Lauriane Kuhn, Philippe Hamman, Alain Beck, Yannis-Nicolas François, Emmanuelle Leize-Wagner
   HPLC 2013, 16-20 juin 2013, Amsterdam, Holland

3. Complete sequence coverage in one injection followed by posttranslational modifications and major N-glycosylations characterization of monoclonal antibodies by sheathless capillary electrophoresis - tandem mass spectrometry
   Rabah Gahoual, Alicia Burr, Jean-Marc Busnel, Lauriane Kuhn, Phillipe Hammann, Alain Beck, Yannis Nicolas François, Emmanuelle Leize-Wagner
   International EUPA 2013, 14-17 octobre 2013, Saint-Malo, France

4. Novel sheathless CE-MS interface as an original and powerful infusion platform for nanoESI study: from intact proteins to high molecular mass noncovalent complexes
   Rabah Gahoual, Jean-Marc Busnel, Philippe Wolff, **Yannis Nicolas François**, Emmanuelle Leize-Wagner
   Strasbourg workshop on native MS and Ion Mobility, 8-9 avril 2013, Strasbourg France

5. Full antibody primary structure and micro-variant characterization in a single injection using transient isotachophoresis with ultra-low flow capillary electrospray ionization – tandem mass spectrometry (CESI- MS)
   Rabah Gahoual, Jean-Marc Busnel, Alain Beck, **Yannis Nicolas François**, Emmanuelle Leize-Wagner
   NVMS, 13-15 avril 2014, Kerkrade, Pays-Bas

6. Monoclonal antibodies complete primary structure and biosimilarity assessment in a single analysis by sheathless capillary electrophoresis-mass spectrometry
   Rabah Gahoual, Jean-Marc Busnel, Alain Beck, **Yannis Nicolas François**, Emmanuelle Leize-Wagner
   IMSC, 24-29 aout 2014, Geneva, Suisse

7. Measuring Variability of Glycosylation Analysis Methods for Monoclonal Antibodies: A NIST Interlaboratory Study
   M. De Leoz, YN. Francois among many contributors
   64th ASMS Conference, 5-9 Juin, San-Antonio, USA

Conférence nationale orale

1. Enrichment by fraction collection developed for CE/MALDI-MS to analyze proteins by Top Down strategy
   M. Biacchi, R. Bhajun, Y.N. François, E. Leize-Wagner
   29ème Journées Francaises de Spectrométrie de Masse, 17-20 septembre 2012, Orléans

2. Evaluation of Sheathless Capillary Electrophoresis as an Infusion Platform for the NanoESI Study of Intact Proteins and Non-Covalent Complexes
   Rabah Gahoual ; Yannis FRANCOIS ; Emmanuelle LEIZE-WAGNER
   RCJSM, 2-6 avril 2012, La Bresse
3. Simultaneous Peptide Mapping, Posttranslational Modifications and Major N-Glycosylation Characterization of Trastuzumab by Sheathless Capillary Electrophoresis-Tandem Mass Spectrometry
   Rabah Gahoual, Alicia Burr, Jean-Marc Busnel, Lauriane Kuhn, Philippe Hamman, Alain Beck, Yannis-Nicolas François, Emmanuelle Leize-Wagner
   SEP13, 4-7 juin 2013, Paris, France

4. Glycoform Separation and Characterization of Cetuximab Variants by Middle-up Off-line CE-UV/ESI-MS
   N. Said, M. Biacchi, R. Gahoual, A. Beck, Y.N. François, E. Leize-Wagner
   SMAP 2015, 15-18 septembre 2015, Ajaccio

5. Overview of CE-MS methodologies for protein characterization and proteomic application
   Y.N. François
   CJSFEAP 2016, 24-26 mai 2016, Lille, France

Affiche internationale

Sheathless Capillary Electrophoresis as an Infusion Platform for the NanoESI Study of Intact Proteins and Non-Covalent Complexes
Rabah Gahoual ; **Yannis Francois** ; Emmanuelle LEIZE-WAGNER
60ème International ASMS Conference, 20-24 mai 2012, Vancouver

Enrichment by fraction collection developed for CE/MALDI-MS to analyze proteins by Top Down strategy
M. Biacchi, R. Bhajun, Y.N. François, **E. Leize-Wagner**
60ème International ASMS Conference, 20-24 mai 2012, Vancouver

Evaluation of Sheathless Capillary Electrophoresis as an Infusion Platform for the NanoESI Study of Intact Proteins and Non-Covalent Complexes
Rabah Gahoual ; **Yannis FRANCOIS** ; Emmanuelle LEIZE-WAGNER
19ème International Symposium ITP2012, 30 septembre - 3 octobre 2012, Baltimore

Enrichment by fraction collection developed for CE/MALDI-MS to analyze proteins by Top Down strategy
M. Biacchi, R. Bhajun, Y.N. François, E. Leize-Wagner

19ème International Symposium ITP2012, 30 septembre - 3 octobre 2012, Baltimore

CE-UV/MALDI-MS: a new platform for Enrichment by fraction collection to analyze intact proteins by Top Down approach
M. Biacchi, R. Bhajun, Y.N. François, E. Leize-Wagner
MSB 2013, 10-14 mars 2013, Charlottesville, Virginie, USA

CE-UV/MALDI-MS: a new platform using an integrative matrix delivery system for proteomics and intact proteins applications
M. Biacchi, J. Hertzog, C. Adolf, **Y.N. François**, E. Leize-Wagner
HPLC 2013, 16-20 juin 2013, Amsterdam, Holland

Simultaneous peptide mapping, posttranslational modifications and major N-glycosylation characterization of trastuzumab by sheathless CE-ESI-MS/MS - Comparison to nanoLC-MS/MS
Rabah Gahoual, Alicia Burr, Jean-Marc Busnel, Lauriane Kuhn, Phillipe Hammann, Alain Beck, **Yannis Nicolas François**, Emmanuelle Leize-Wagner
61ème International ASMS Conference, 9-13 juin 2013, Minneapolis, USA

CE-UV/MALDI-MS: a new platform for proteomic and intact protein characterization
M. Biacchi, Anja Resemann, Pierre-Olivier Schmit, Alain Beck, Y.N. François, E. Leize-Wagner
61ème International ASMS Conference, 9-13 juin 2013, Minneapolis, USA

CEToolbox: Specialized calculator for Capillary Electrophoresis users as an Android application
J. Pansanel, M. Biacchi, **R. Gahoual**, E. Leize-Wagner, Y.N. François
HPLC 2013, 16-20 juin 2013, Amsterdam, Holland

CE-UV/MALDI-MS: a new platform for proteomic and intact protein characterization
M. Biacchi, **Anja Resemann**, Pierre-Olivier Schmit, Alain Beck, Y.N. François, E. Leize-Wagner
12ème International HUPO Conference, 14-18 septembre 2013, Yokohama, Japan

CEToolbox: Specialized calculator for Capillary Electrophoresis users as an Android application
J. Pansanel, M. Biacchi, R. Gahoual, E. Leize-Wagner, **Y.N. François**
CE Pharm, 6-9 octobre 2013, Arlington, USA

CE-UV/MALDI-MS: a new platform for proteomic and intact protein characterization
M. Biacchi, Anja Resemann, Pierre-Olivier Schmit, **Alain Beck**, Y.N. François, E. Leize-Wagner
Mass Spec 2013, 23-26 septembre 2013, Boston, USA

Simultaneous peptide mapping, posttranslational modifications and major N-glycosylation characterization of trastuzumab by sheathless CE-ESI-MS/MS - Comparison to nanoLC-MS/MS
Rabah Gahoual, Alicia Burr, Jean-Marc Busnel, Lauriane Kuhn, Phillipe Hammann, **Alain Beck**, Yannis Nicolas François, Emmanuelle Leize-Wagner
Mass Spec 2013, 23-26 septembre 2013, Boston, USA

CEToolbox: Specialized calculator for Capillary Electrophoresis users as an Android application
J. Pansanel, M. Biacchi, **R. Gahoual**, E. Leize-Wagner, Y.N. François
Mass Spec 2013, 23-26 septembre 2013, Boston, USA

CE-UV/MALDI-MS: a new platform for proteomic and intact protein characterization
M. Biacchi, Anja Resemann, Pierre-Olivier Schmit, Alain Beck, **Y.N. François**, E. Leize-Wagner
CE Pharm, 6-9 octobre 2013, Arlington, USA

Intact protein separation by CE-UV/MALDI-MS for top-down proteomics
Michael Biacchi, Alain Beck, **Yannis Francois**, Emmanuelle Leize-Wagner
62ème International ASMS Conference, 15-19 juin 2014, Baltimore, USA

Monoclonal antibodies complete primary structure and biosimilarity assessment in a single analysis using transient isotachophoresis sheathless capillary electrophoresis-tandem mass spectrometry
Rabah Gahoual, Jean-Marc Busnel, Johana Chicher, Lauriane Kuhn, Philippe Hammann, Alain Beck, Yannis Francois, Emmanuelle Leize-Wagner
62ème International ASMS Conference, 15-19 juin 2014, Baltimore, USA

Monoclonal Antibodies Complete Primary Structure and Biosimilarity Assessment in a Single Analysis Using Transient Isotachophoresis Sheathless Capillary Electrophoresis-tandem Mass Spectrometry
Rabah Gahoual, Jean-Marc Busnel, Johana Chicher, Lauriane Kuhn, Philippe Hammann, **Alain Beck**, Yannis Francois, Emmanuelle Leize-Wagner
11ème CASSS MS, 9-12 septembre 2014, Napa, USA

Monoclonal antibodies complete primary structure and biosimilarity assessment in a single analysis using transient isotachophoresis sheathless capillary electrophoresis-tandem mass spectrometry
Rabah Gahoual, Jean-Marc Busnel, Johana Chicher, **Lauriane Kuhn**, Philippe Hammann, Alain Beck, Yannis Francois, Emmanuelle Leize-Wagner
63ème International ASMS Conference, 10-14 juin 2015, Saint-Louis, USA

High-order structural interrogation of antibody-drug conjugates by a combination of intact, middle-up and bottom-up techniques using sheathless capillary electrophoresis-mass spectrometry
Nassur Said, Rabah Gahoual, Lauriane Kuhn, Alain Beck, **Yannis Nicolas François**, Emmanuelle Leize-Wagner
64ème International ASMS Conference, 5-9 juin 2016, San Antonio, USA

Glycoform Separation and Characterization of Cetuximab Variants by Middle-up Off-line CE-UV/ESI-MS
Nassur Said, Michael Biacchi, Rabah Gahoual, Charly Renard, Alain Beck, Yannis Nicolas François, Emmanuelle Leize-Wagner
64ème International ASMS Conference, 5-9 juin 2016, San Antonio, USA

High-order structural interrogation of antibody-drug conjugates by a combination of intact, middle-up and bottom-up techniques using sheathless capillary electrophoresis-mass spectrometry
Nassur Said, Rabah Gahoual, Lauriane Kuhn, Alain Beck, **Yannis Nicolas François**, Emmanuelle Leize-Wagner
26ème International IMSC, 20-26 aout 2016, Toronto, Canada

Glycoform Separation and Characterization of Cetuximab Variants by Middle-up Off-line CE-UV/ESI-MS
Nassur Said, Michael Biacchi, Rabah Gahoual, Charly Renard, Alain Beck, Yannis Nicolas François, Emmanuelle Leize-Wagner
26ème International IMSC, 20-26 aout 2016, Toronto, Canada

Affiche nationale

CESI-MS. Principe et application pour l'analyse sensible de mélanges peptidiques complexes
M. Biacchi, N. Pourhassan, A. A. Heemskerk, B. Schoenmaker, **Y. Francois**, E. Leize, J. M. Busnel, O.A. Mayboroda
28ème Journées Francaises de Spectrométrie de Masse, 19-22 septembre 2011, Avignon

Caractérisation de protéines entières par couplage indirect CIEF/MALDI-MS
Michaël Biacchi, Nina Pourhassan, **Yannis Francois**, Emmanuelle Leize-Wagner
28ème Journées Francaises de Spectrométrie de Masse, 19-22 septembre 2011, Avignon

Proteinscape, un outil convivial et performant au service de la protéomique : de la gestion des projets à l'exploitation des données de spectrométrie de masse
L. Kuhn, P-O. Schmit, Y. François, E. Leize-Wagner, P. Hammann
28ème Journées Francaises de Spectrométrie de Masse, 19-22 septembre 2011, Avignon

Evaluation of Sheathless Capillary Electrophoresis as an Infusion Platform for the NanoESI Study of Intact Proteins and Non-Covalent Complexes
Rabah Gahoual ; Yannis FRANCOIS ; Emmanuelle LEIZE-WAGNER
29ème Journées Francaises de Spectrométrie de Masse, 17-20 septembre 2012, Orléans

Evaluation of Sheathless Capillary Electrophoresis as an Infusion Platform for the NanoESI Study of Intact Proteins and Non-Covalent Complexes
Rabah Gahoual ; Yannis FRANCOIS ; Emmanuelle LEIZE-WAGNER
18ème journée d'électrophorèse capillaire, 9 octobre 2012, Paris

CE-UV/MALDI-MS: a new platform for proteomic and intact protein characterization
M. Biacchi, Anja Resemann, Pierre-Olivier Schmit, Alain Beck, Y.N. François, E. Leize-Wagner
SEP13, 4-7 juin 2013, Paris, France

CEToolbox: Specialized calculator for Capillary Electrophoresis users as an Android application
J. Pansanel, M. Biacchi, **R. Gahoual**, E. Leize-Wagner, Y.N. François
SEP13, 4-7 juin 2013, Paris, France

Glycoform Separation and Characterization of Cetuximab Variants by Middle-up Off-line CE-UV/ESI-MS
Nassur Said, Michael Biacchi, Rabah Gahoual, Charly Renard, Alain Beck, Yannis Nicolas François, Emmanuelle Leize-Wagner
JFSM 2016, 26-30 Septembre 2016, Bordeaux, France

IV. Action d'enseignement et de diffusion

Encadrement d'étudiants

	Thèse	*Master 2*	*Master 1*	*Licence 3*	*Licence 2*
Etudiant	**4**	**9**	**10**	**13**	**6**

Thèse :

Dr. Michael BIACCHI
Thèse soutenue le 23 septembre 2014
(50%)

Dr. Rabah GAHOUAL
Thèse soutenue le 29 septembre 2014

M. Nassur SAID
Soutenance de thèse prévue en 2017

M. Jérémie GIORGETTI
Soutenance de thèse prévue en 2019

Master 2 recherche :

2010 : M. Bruno GUYONNET (6 mois) (50%)
2011 : Mme Nina POURHASSAN (6 mois) (100%)
2012 : Mme Alicia BURR (6 mois) (100%)
2014 : M. Nassur SAID (6 mois) (100%) – Mme Clotilde DEVAUX (6 mois) (50%)
2015 : M. Jérémie GIORGETTI (6 mois) (100%) – Camille LINKER (6 mois) (50%)
2016 : M. Jordan KIRMANN (6 mois) (100%) – Charly RENARD (6 mois) (50%)

Master 1 :

2010 : Mme Nina POURHASSAN (4 mois) (100%)
2011 : Mme Iliana TSONEVA (4 mois) (100%)
2012 : M. Ricky BHAJUN (3 mois) (100%) – M. Heidi SOUFFARY (3 mois) (100%) - M. Marc LIBERATORE (6 mois) (25%)
2013 : Mme Jasmine HERTZOG (2 mois) (100%)
2015 : M. Charly RENARD (3 mois) (100%) – M. Jordan KIRMANN (3 mois) (100%)
2016 : M. Anthony LECHNER (2 mois) (100%) - Julie CANONGE (2 mois) (100%)

Licence 3 :

2010 : M. T. DURAND (1 mois) – M. H. SIRINDIL (1mois),
2011 : Mme L. ATI (1 mois), Mme E. Bonfanti (1 mois)
2013 : Mme, T. MARLEUX (1 mois),
2012 : M. PAQUET (1 mois), Mme J. HERTZOG (2 mois), M. C. ADOLF (2 mois),
2014 : Mme C. MAZOHIE (1 mois), Mme M. METZELER (1 mois)
2015 : Mme EL MASOUSI (1 mois), M. J. HEISER (2 mois), M. S. PASHINA (2 mois)

Licence 2 et 2ème année d'école :

2012 : M. T. GUGGENBUHL (2 mois), Mme M. COURBET (1 mois), M. GUILLAUME (1 mois)
2013 : Mme M. DAVRIL (1 mois), Mme M. MONNIER (1 mois)
2016 : Mme C. BALHONZER (1 mois)

Activité d'enseignement

Niveau Licence

Licence 3 Chimie et Chimie Physique : cours et TD de Spectrométrie de masse (6 heures + 8 heures), cours d'électrophorèse capillaire (5 heures), TD de protocole analytique (4 heures)
Licence 3 Pro Synthèse : Cours de spectrométrie de masse (4 heures), cours de chromatographie (4 heures), TP sciences analytiques (32 heures)
Licence 3 Pro Science du médicament : cours de chromatographie gaz (10 heures), cours d'électrophorèse capillaire (4 heures)
Licence 3 Chimie-Biologie : TP sciences analytiques (50 heures)
Licence 3 STE : cours de spectrométrie de masse (4 heures)

Niveau Master

Master SACEB : Cours d'électrophorèse capillaire (10,5 heures)
Master 1CPAM : TP Spectrométrie de masse (8 heures)
Master 1 VRV : TP Phytochimie (40 heures)
Master 2 VRV : TP Phytochimie (40 heures)

Formation continue

Formation du coupage CE-MS : cours + TD + TP (24 heures)

Service d'enseignements

Année	Heures équivalent TD
2009-2010	228
2010-2011	256
2011-2012	234
2012-2013	212
2013-2014	220
2014-2015	232
2015-2016	239

V. Rayonnement et diffusion

Activité d'expertise

2012 : Membre d'un comité de sélection pour le concours MCD 31 Chaire 4078 – Université de Strasbourg
2015 : Membre d'un comité de sélection pour le poste d'ATER
2015 : Expertise d'un projet région Ile de France (DIM Analytics)
2015-2016 : Expertise d'un projet anses
2016 : Membre du comité de sélection du concours externe BAP 242 769

Participation à des jurys de soutenance de thèse

23/09/2014 : Membre Invité de la thèse de Michael Biacchi – Université de Strasbourg
Directeur de thèse : Dr. E. Leize-Wagner
29/09/2014 : Membre Invité de la thèse de Rabah Gahoual – Université de Strasbourg

Directeur de thèse : Dr. E. Leize-Wagner

Referee de revues Internationales indexées

Analytical Chemistry, Journal of Chromatography A, Electrophoresis, Rapid Communication in Mass Spectrometry, Analytical Method, Proteomics

Organisation de congrès nationaux/Internationaux

2014 : Membre du comité d'organisation du Workshop international « Research and innovation of CESI-MS coupling » organisé à l'Université de Strasbourg du 30 septembre au 2 octobre.

Chairman de la session « metabolomic »

28 février 2014 Organisation d'une journée scientifique – Université de Strasbourg

VI. Reponsabilités scientifiques :

Obtention de financements (equipement, bourse de doctorat…)

2015 : Co-rédacteur de Projet IDEX enseignement « Initiation aux méthodes de la recherche en Nanobiophysicochimie ». Obtention de 165 k€

2015 : Porteur de projet « caractérisation de modifications post-traductionnelles dans le but de dosage immunologique ». Obtention de 10k€ par la société Beckman Coulter (2 stages M2 recherche)

Participant à hauteur de 15% du projet Idex « Initiation aux méthodes de la recherche en Nanobiophysicochimie ». Obtention de 150 k€

2014 : Porteur de projet pour l'obtention d'une bourse de thèse Région avec le conseil générale de Mayotte. Obtention d'un financement de 45 k€ (Thèse de Nassur Said)

Coordinateur d'un achat d'une électrophorèse capillaire pour l'enseignement de Travaux Pratique. Obtention de 25 k€

Porteur de projet « caractérisation de systèmes complexes par CESI-MS ». Obtention de 3k€ par la société Bruker (1 stages M2 recherche)

Obtention de financements industriels

2012-2015 : Porteur de projet « mise en place d'un laboratoire d'application CESI-MS » avec la société Sciex Separation. Obtention d'un budget de 800 k€ sous la forme de prêt renouvelable de machines de dernière génération (CESI8000, Triple TOF) et up-date.

2014-2016 : Participation à une collaboration avec la Société Bruker « application à la caractérisation de protéines entières par CESI-MS ». Obtention d'un budget de 500 k€ pour un équipement de spectrométrie de masse haute résolution.

2012-2016 : Coordinateur d'une collaboration industrielle avec la société Pierre Fabre « caractérisation fine à différents niveaux d'anticorps monoclonaux, de biosimilaires et de conjugués ». Obtention d'échantillons dont la valeur peut atteindre 10 k€. Collaboration ayant permis la publication de 6 articles.

Repsonsabilité/ fonctions dans instances nationales/locales

2011-2015 : Membre élu du Comité National des Universités en section 31. Membre suppléant actif (2 sessions)
2011-2015 : Membre élu du comité scientifique de la 31ème section de l'Université de Strasbourg
2013-2018 : Membre du conseil d'UMR 7140 Chimie de la Matière Complexe.
2015-2019 : Membre élu du Conseil de la Faculté de Chimie de Strasbourg.
2015-2019 : Membre nommé de la commission recherche de la Faculté de Chimie de Strasbourg