Université de Strasbourg

UNIVERSITÉ DE STRASBOURG

École doctorale
Sciences de la **vie**
et de la **santé** | ED 414
Université de Strasbourg

ÉCOLE DOCTORALE – Sciences de la vie et de la santé – ED 414

Labo TRAVE - UMR 7104-U964

THÈSE **présentée par / DISSERTATION presented by :**

Kathleen WEIMER

soutenue le /defended on : **16 December 2024**

pour obtenir le grade de **/** to obtain the grade of **:**
Docteur de l'Université de Strasbourg / Strasbourg University Doctor

Discipline/Spécialité **/** Discipline/Specialty : Cellular and Molecular Biology

Mapping the HPV16 E6-host protein interactome using quantitative interactomic methods to investigate biophysical properties of binding and model cellular complex formation

THÈSE dirigée par / DISSERTATION supervisor :

Dr. TRAVÉ Gilles	Directeur de recherches, Université de Strasbourg/ IGBMC CNRS UMR 7104 (Illkirch-Graffenstaden France)

RAPPORTEURS :

Dr. BACHELERIE Francoise	Directrice de recherches, Université Paris-Saclay (Paris, France)
Dr. WOLFF Nicolas	Directeur de recherches, Institut Pasteur (Paris, France)

AUTRES MEMBRES DU JURY / OTHER MEMBERS OF THE JURY :

Dr. COYAUD Étienne	Chargée de recherches, University of Lille (Lille, France)
Dr. SIMON Claudia	Chargée de recherches, Boehringer Ingelheim Pharma GmbH (Biberach an der Riß, Germany)
Dr. MASSON Murielle	Chargée de recherches, École supérieure de biotechnologie Strasbourg ESBS, CNRS UMR7242 BSC (Illkirch-Graffenstaden, France)

Avertissement au lecteur / Warning to the reader

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition des membres de la communauté universitaire. Il est soumis à la propriété intellectuelle de l'auteur. Cela implique une obligation de citation et de référencement lors de l'utilisation de ce document. D'autre part, toute contrefaçon, plagiat, reproduction ou représentation illicite encourt une poursuite pénale.
This document is the result of a long process approved by the jury and made available to members of the university community. It is subject to the intellectual property rights of its author. This implies an obligation to quote and reference when using this document. Furthermore, any infringement, plagiarism, unlawful reproduction or representation will be prosecuted.

Code de la Propriété Intellectuelle

Article L122-4 :

Toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite. Il en est de même pour la traduction, l'adaptation ou la transformation, l'arrangement ou la reproduction par un art ou un procédé quelconque.
Any representation or reproduction in whole or in part without the consent of the author or his successors in title or assigns is unlawful. The same applies to translation, adaptation or transformation, arrangement or reproduction by any art or process whatsoever.

Articles L335-1 à L335-9. : Dispositions pénales / Penal provisions.

Licence attribuée par l'auteur / Licence attributed by the author

[*Veillez à ne conserver que la licence qui vous convient* / Make sure you only keep the licence that suits you best]

https://creativecommons.org/licenses/?lang=fr-FR

Kathleen WEIMER

Cartographie de l'interactome de la protéine hôte HPV16 E6 à l'aide de méthodes d'interactomique quantitative pour étudier les propriétés biophysiques de la liaison et modéliser la formation de complexes cellulaires.

Résumé

Les papillomavirus humains (HPV) constituent l'infection sexuellement transmissible la plus répandue. Alors que les présentations cliniques des infections à HPV peuvent aller de l'asymptomatique subclinique aux cancers anogénitaux, un sous-ensemble de HPV connus sous le nom de HPV à haut risque (HR-HPV) sont spécifiquement associés aux cancers, causant jusqu'à 99% de tous les cas de cancer du col de l'utérus. Parmi ces cas, plus de 50 % sont signalés comme étant liés à une infection par le HPV16. Il est bien établi que l'oncoprotéine E6 du HPV joue un rôle central dans la pathogenèse du cancer du col de l'utérus et l'E6 du HPV16 est considérée comme l'E6 prototypique. Étant donné la grande diversité génotypique et phénotypique des HPV, ceux-ci sont souvent considérés comme des systèmes modèles pour les études interactomiques. En outre, la capacité de l'E6 à reconnecter les réseaux cellulaires facilitant la progression de la maladie, en conjonction avec son importance clinique, a suscité plusieurs études interactomiques. Ici, nous visons à caractériser l'interactome quantitatif de l'E6 prototypique du HPV16 en utilisant le Holdup natif (nHU).

Mots clés : HPV, interactomique, IPP, E6, oncoprotéine, carcinogenèse, pathogénèse

Abstract

Human Papilloma Viruses (HPVs) are the most widespread sexually transmitted infection. While clinical presentations of HPV infections can range from subclinical asymptomatic to anogenital cancers, a sub-set of HPVs known as High-Risk HPVs (HR-HPVs) are specifically associated with cancers causing up to 99% of all cervical cancer cases. Of these cases >50% are reported as related to HPV16 infection. It’s been well established that the HPV oncoprotein E6 has a central role in cervical cancer pathogenesis and HPV16 E6 is regarded as the prototypical E6. Given the high genotypic and phenotypic diversity of HPVs, they are often considered model systems for interactomic studies. Furthermore, the ability of E6 to rewire cellular networks facilitating disease progression in conjunction with its clinical significance has prompted several interactomic investigations. Here we aim to characterize quantitative interactome of the prototypical HPV16 E6 using native Holdup (nHU).

Keywords : HPV, interactomics, PPIs, E6, oncoprotein, carcinogenesis, pathogenesis

Table of Contents

Presentation of Thesis Manuscript ..... **i**
Table of Contents **1**
Preface **4**
Introduction **6**
Interactomics 6
Introduction to interactomics 6
Standard methods in interactomics 7
Two-hybrid systems 7
Protein pull-downs 9
Protein chip technology 10
MS based approaches 11
Limitations in standard method in interactomics 13
Towards quantitative interactomics 15
Review: Molecules interact: but how strong and how much? 16
Methods in quantitative interactomics 28
Chromatographic retention assays 28
Interactomics and Disease 30
HPV as a model for interactomics 31
Review: HPV driven carcinogenesis 32
Introduction to research project **42**
Thesis project and objectives 42
Measuring the HPV16 E6 Affinity Interactome: Materials and Methods **43**
Materials and Methods (Main Project) 44
Measuring the HPV16 E6 Affinity Interactome: Data and Results **52**
Data and Results: Modelling HPV 16E6-host protein complex formation with quantitative interactomics (manuscript *in Preparation*) 53
Supplemental Materials and Methods, Data and Results **94**
Measuring 16E6-UBE3A interactions in different oligomerization states 95
Materials and Methods 95
HU of 16E6 with UBE3A 96
Data and Resulsts 96
HU of 16E6 with UBE3A 96
Mapping the E6* interactome 97
Materials and Methods 97

nHU with E6* 97
Data and Results 97
nHU with E6* 97
Evaluating the 16E6-PTPN14 interaction and its cellular effects 98
Materials and Methods 98
nHU spike-in with PTPN14 PY motif peptide 98
Monitoring PTPN14 in HEK293T-16E6 cell lines 98
Transfection of HEK293T-16E6 cell line with E7 98
Data and Results 97
nHU spike-in with PTPN14 PY motif peptide 100
Evaluating cellular effects of 16E6 on PTPN14 101
Evaluating cellular effects of 16E6 and E7 102
Investigating 16E6 interactions with the ubiquitin proteosome system 104
Materials and Methods 104
Generating knock-downs in HeLa and HEK293T cells 104
Data and Results 105
Investigating 16E6 interactions with the UPS 105
Evaluation of 16E6 inhibitors via nHU 106
Materials and Methods 106
16E6 treatment with inhibitors and nHU 106
Data and Results 107
Adapting nHU for therapeutic evaluation of E6 inhibitors 107
Systematic AlphaFold modelling of 100 16E6-binders 112
Materials and Methods 112
AlphaFold3 predictions 112
Data and Results 113
Evaluating AlphaFold models using Predictomes 113
Digital data provided in shared drive 122
Discussion and Future Perspectives 123
Discussion 124
Studying host-pathogen interactions with quantitative interactomics 124
Key findings from mapping the 16E6 host-affinity interactome 125

Intricately interwoven interactomes 126
Future Perspectives 127
Functional analysis of novel 16E6-binder 127
Continued study of HPV E6 with quantitative interactomics 127
Study of HPV E6 oncoproteins 128
Evaluating therapeutics with nHU 128
Developing a single affinity step quantitative interactomic method 128
References **130**
Contributions **139**
List of publications **141**
List of oral presentations **142**
Poster: The 16E6 affinity interactome 143
Poster: Investigation 16E6 interactions with native holdup 144
Thesis Summary in French **145**
Thesis Summary in English **149**
Acknowledgements **153**

Preface

Upon commencing my doctoral studies in the Trave team, Gilles asked me a question that he regarded as the only important question: “Why do you want to obtain a PhD.” I was honest in that I regarded it as a means to an end. I knew that I wanted a career in research and to have the autonomy I desired in my career I also needed the appropriate credentials. I have had longstanding goals to work with Médecins Sans Frontières, studying infectious disease from a perspective of translational medicine working to develop molecular diagnostics and therapeutics. To effectively achieve this, I decided that I would need a PhD in biochemistry. Now that I near the end of my doctoral studies and have experienced the process, I can fully appreciate what a PhD really means. It is a toolbox that one carries with them for the rest of one’s career and life. It equips one with the knowledge, resources, and liberty to ask challenging questions and construct solutions to problems. It is a commitment to constant growth and acquisition of knowledge. My time in Gilles’ lab has especially pushed me outside my comfort zone and limits in knowledge. I was challenged to adapt to new ways of thinking and consider new perspectives, and I believe this way of learning how to think critically is one of the greatest tools one can have. I have tried to reflect this value throughout my dissertation in hopes that it will insight interesting dialogues and prompt new ideas.

The journey of a thesis is often regarded to as one that is non-linear—full of ups, downs, twists, and curves. I think this is also a quality reflected in the format of my thesis manuscript. Throughout my PhD I was afforded opportunities to contribute to two literature reviews. The first, was a shared effort with colleagues from the team. The intention was that I could gain a sense of orientation in the field of quantitative interactomics, the prevailing methodology of the lab and a topic entirely new to me, while also integrating into the team. The collaboration and discussion that stemmed from this venture laid the foundation for my thesis project. The second opportunity presented itself nearing the completion of my thesis. Following a presentation of my research at a conference, I was asked to contribute to a special edition of the journal which was sponsoring the section I participated in. This exemplified a milestone in establishing my autonomy as a soon to be independent researcher. With encouragement and support from Gilles, I took this professional leap and wrote an individual piece providing a perspective on a topic that I have now studied at great lengths, the HPV oncoprotein E6. Not only do these two works embody much of the growth in my knowledge and skills over the course of my thesis, but they have generated compositions containing the most pertinent background information applicable to my thesis project. In an effort to present this information in a manner that is clear and concise, these two reviews have been embedded within the introduction section of my thesis manuscript and linked together by sections of supporting information.

Following the introduction I have provided an overview of the thesis project and the objectives that guided this research. Next, I provide the full experimental methodologies employed in this body of work. The results are then presented in two parts: (1) data and results in the form of a prepared manuscript for publication including the repetition of previously provided methods to retain the structure of the article. These results encompass the main results from my thesis project and the intention to publish and disseminate them throughout the scientific community as part of completing my thesis. (2) data and results from experiments and side projects conducted throughout my thesis that represent work that will not be included in this publication but were important contributions to the overall project and the future of this research. The project in its entirety is then appraised and evaluated, suggesting possible branch points for further study.

I am aware that this is a deviation from the traditional thesis manuscript format. The presented body of work was synthesized in the most organized and efficient way to accurately represent the cumulative efforts of my thesis project. While this has created a mélange of completed manuscripts interspersed by accessory sections to bridge any gaps in information and an overall condensed format, an extensive amount of research and effort has gone into the preparation of each respective manuscript so that they can be published. Furthermore, a reoccurring theme that I have experienced throughout my PhD is that there is no one-size fits all approach to anything in research. While of course we try to uphold universal standards to ensure good quality research that is rigorous, robust, and reproducible, it is also the individual styles, unique perspectives, and often bold ideas that advance science. With that said, I greatly appreciate those taking the time and effort to read and evaluate my work. I sincerely hope that you enjoy reading this manuscript as I enjoyed preparing it.

Kathleen Marie Weimer

December 2024

Strasbourg, France

Introduction

Interactomics:

Introduction to interactomics:

Interactomes are complex cellular networks that encompass the entirety of molecular interactions within a cell[1]. This creates an intricate framework of interactions that are temporal, often transient, and contextual by nature and ultimately underpin cellular function and regulation. Proteins rarely operate in isolation, with over 80% of proteins requiring interaction with other proteins to realize their function[2]. Furthermore, proteins possess a central role in cellular processes as they are considered the primary functional unit of the cell[3] and given the vast complexity and highly dynamic state of interactomes many studies, including this one, focus on interactomes from the perspective of protein-protein interactions (PPIs)[4]. In fact, the dynamic nature of interactomes is unique compared to that of other levels of cellular organization. The interactome does not demand the synthesis of new molecules thus affording more flexibility and plasticity as it swiftly responds to cellular stimuli[5]. In this way interactomes are also easily influenced by extrinsic factors, such as the transcriptome and proteome, leading to a high level of crosstalk between molecular strata with all communications converging at the interactome.

These key features have identified interactomes as interesting prospects for investigation, however it is these same features that contribute to their challenging nature for study. In general, proteins tend to be more biochemically complex than other molecules and therefore technically challenging to manipulate and study[6]. When we consider the amount of optimization that is sometimes necessary to produce a singular protein in the lab and realize that studying interactomes seeks to characterize the cell-wide network of protein interactions, often without prior knowledge as to which interactions might occur, the feat may seem impossible. Further still, there are several prerequisites to studying intact PPIs including that the proteins involved must be present, in sufficient stoichiometry and structural conformation, all within the appropriate cellular conditions and proximity[7,8]. As mentioned, interactomes are also highly dynamic as the interactions that constitute them are constantly changing in response to external factors. Proteins involved in multi-step cellular processes, such as signaling cascades or the catalytic activities of the spliceosome, undergo extensive interactomic shifts. For example, over the course of a splicing reaction, the spliceosome is constantly recruiting and releasing partners, altering complex composition to drive different steps of the splicing process[9]. This reflects the high degree of functional diversity orchestrated on the interactome level and the importance of acquiring these difficult to capture transient moments[5].

Standard methods in interactomics

Despite these challenges, several experimental methods have been developed to systematically study interactomes and attempts have even been made to measure organism wide reference interactomes[10-16]. Methodologies in studying interactomes include both proteomic based and interactomic based approaches that fall over a range of throughput. Standard techniques for studying interactomes at high throughput employed today include two-hybrid systems, protein pull-downs, protein chip technology, and various mass spectrometry (MS) based approaches (Fig. T1&T2). Among these, various computational methods are also often used, but as these are dependent on the analysis of data obtained through experimental techniques, we will not discuss them here.

Two-hybrid systems

The original pioneering method in interactomics was the complementation assay, yeast two-hybrid (Y2H). Like other complementation based approaches, such as fluorescence resonance energy transfer (FRET)[17] and bimolecular fluorescence complementation (BiFC)[18], Y2H started as a relatively low throughput method focused on only a small subset of interactions. However, it later became adapted for high throughput applications and, until recent years, was considered the gold standard in interactomic methods[12].

The principle of the assay comes from the modular construction of the Gal4 transcription activator in yeast[19]. Containing two separate functional domains— the N-terminal DNA binding domain (DBD) and the C-terminal transcriptional activation domain (AD)— it was determined that the two regions retained their respective activity regardless of the presence of the other, but both components were required for the overall activity of the transcription factor[20]. Exploiting these features, the Y2H was developed by fusing a gene encoding a protein of interest, or bait, to the DBD thus producing a DBD-bait vector while the AD gene was fused to a gene encoding a protein to be tested for interaction with the bait, creating an AD-prey vector[21]. Several AD-prey vectors created from cDNA libraries can then be used to create interaction screens with the DBD-bait. The two plasmids are then co-transformed and expressed in yeast where interactions between the bait and prey would also lead to the spatial proximity required of DBD and AD to activate transcription. The use of reporter genes, such as *LacZ*, then allows for the screening of positive interactions since the transformants will show blue in the presence of X-β-Gal[22] (Fig. T1).

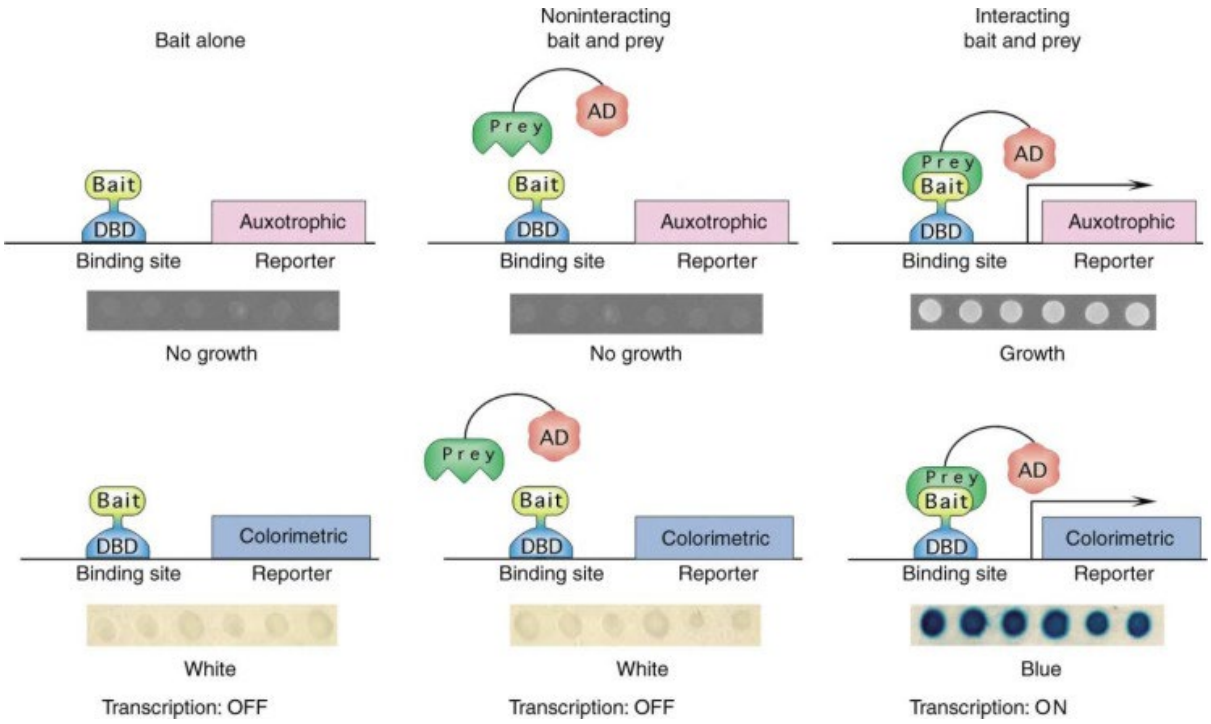


Figure T1 Schematic of yeast two-hybrid components and reporter gene activation. Three situations are depicted from left to right: (1) Transformants possessing bait alone – yeast containing only a bait and two reporters. While the DBD fused to the bait binds the upstream activation sequence (UAS) there is not transcriptional activation due to absence of AD. (2) Transformants possessing bait and prey proteins that do not interact – yeast containing bait, prey, and two reports but no transcriptional activation. The DBD fused bait binds the UAS, but the AD is not in close enough proximity to induce transactivation. (3) Transformants possessing bait and prey that do interact – yeast containing bait, prey, and to reporter display transcription of the reporter gene. Transformants are screened using selective media. For auxotrophic reporters positive colonies will grow, and negative colonies will not grow whereas for colorimetric reporters positive colonies will show in blue and negative colonies will show in white. *(Taken from Serebriiskii, 2013)*

Protein pull-downs

The protein pull-down is adjacent to co-immunoprecipitation in that they use the same principle for capturing interacting proteins, but instead of relying on antigen-antibody interactions the pull-down makes use of affinity tags (Fig. T2)[23]. To perform a pull-down the affinity tagged bait is immobilized on the corresponding affinity matrix and interacting preys are captured or “pulled down” with the fixed bait. Both bait and prey proteins can be obtained from a variety of sources including cell lysates, purified proteins, and in vitro transcription/translation systems. This also contributes to the versatility in pull-down experimental set ups. For example, a recombinantly purified protein can be used as bait to fish preys from cellular extracts. Following incubation of the bait with preys the intact interactions are eluted from the affinity matrix. For low throughput readout of specific interactions the pull-down elutions can be analyzed with Western Blot, but for high throughput analysis and interaction identification pull-down experiments can be coupled to MS[24]. Another frequent adaptation of pull-downs in interactomics is the use of proximity labelling technology, such as BioID[25] and APEX[26], which entails capturing endogenous interactions in vivo through fusing baits to promiscuous enzymes that will label preys as they come into proximity[27].

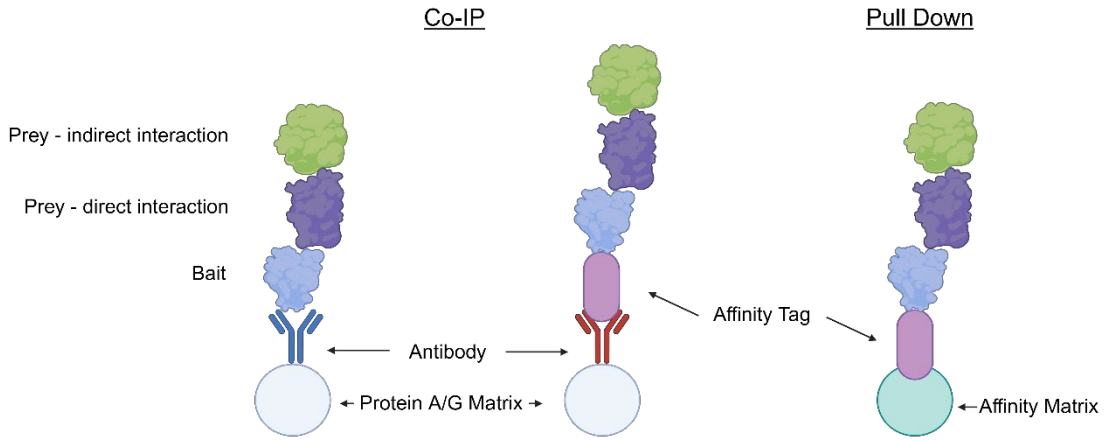


Figure T2 Coimmunoprecipitation and pull-down. Depicted from left to right: (1) The standard Co-IP format. When a specific antibody recognizing a protein of interest (bait) is available, it can be used to capture the protein. The antibody-protein complex is then captured by capturing the antibody with an immobilized protein A/G matrix, thus co-immunoprecipitating the bait with it. Following several washing steps to remove nonspecifically bound proteins, remaining proteins that interact with the bait (preys) can be recovered representing either direct or indirect interactions. (2) In the event that a specific antibody is not available, a tagged version of the protein can be used for co-IP with a corresponding antibody that recognizes the tag (i.e. HA tag). The experimental procedure then follows that of a traditional Co-IP. (3) When using an affinity tagged protein it is also possible to an affinity matrix that corresponds to the tag in place of the Protein A/G matrix coupled to an antibody, for example a GST tag with Glutathione Sepharose, through a process known as a protein “pull-down”. (Adapted from Isono & Schwechheimer, 2010; Created in BioRender.com)

Protein chip technology

The microarray allows fast and easy detection of thousands of parameters in parallel in a single experiment making it the ideal setup for large scale studies[28]. The procedure involves the capture and immobilization of molecules onto a solid support arranged into rows and columns over a very small area. The prepared microarrays are then exposed to samples containing corresponding molecules expected to interact[29]. Detection of complex formation can be deduced from readouts based on fluorescence, chemiluminescence, mass spectrometry, radioactivity, or electrochemistry making them very versatile. These assays are also highly sensitive and can be utilized across basic and applied biology (Fig. T3)[30].

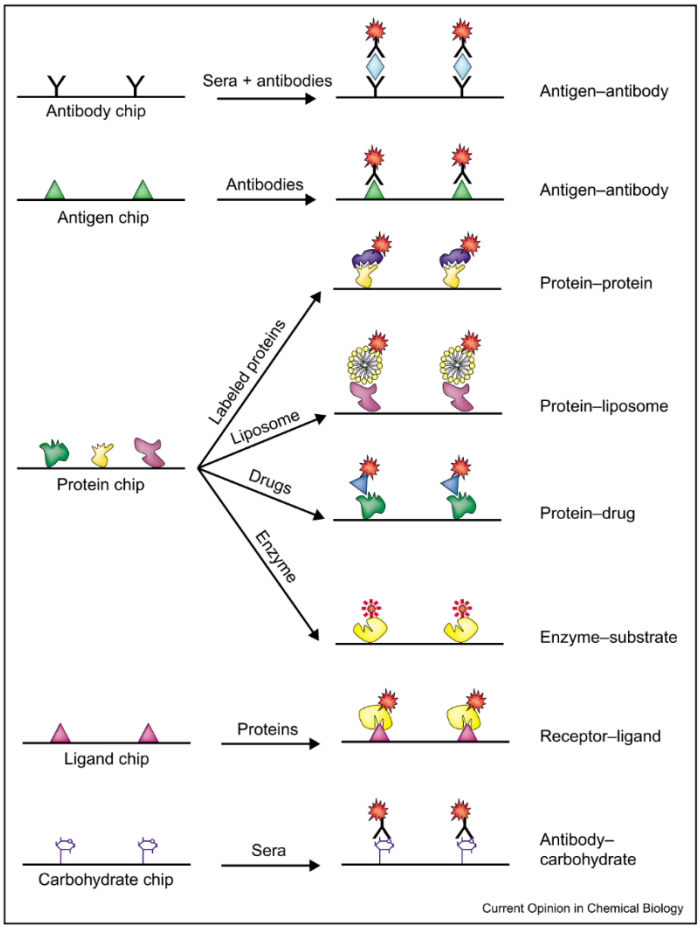


Figure T3 Applications of protein microarrays. There are two general types of protein microarrays: analytical and functional. Analytical microarrays (shown in rows 1-2, 4-5) involve a high-density array of fixed affinity reagents, often antibodies or antigens, that are used to detect interactions with proteins from complex mixtures. Small molecule and carbohydrate microarrays have also been developed as analytical microarrays that can be used to study protein interactions with ligands and carbohydrates. Functional microarrays (row 3) involve immobilizing a large number of purified proteins on a solid support. Functional microarrays are considered broader in terms of applicability in that they can study a wider range of biological activities than analytical microarrays (i.e. PPIs, protein-lipid, protein-nucleic acid, and enzyme substrate interactions) as well as assist in drug and drug target identification. (*Taken from Zhu & Snyder, 2003*)

MS based approaches

With growing technological advancements, the use of proteomics based MS methods in interactomics has greatly expanded (Fig. T4)[31]. The ever-increasing sensitivity of MS and the application of downstream bioinformatic tools for data processing and analysis has often led to coupling this technique with other experimental methods for interaction identification. As such, the protein pull-down assay coupled with MS has emerged as the most sensitive method to investigate PPIs with affinity purification MS (AP-MS) dominating the field as one of the most employed techniques[32]. This method involves the expression of an affinity tagged bait within a cell line that reflects the cellular environment of interest and thus supports biologically relevant interactions. Using gentle lysis protocols, the bait in complex with preys is purified with an affinity step which corresponds to the affinity tag of the bait. While AP-MS has proven especially useful in the identification of interactions, methods such as cross-linking MS (XL-MS) have developed in attempts to provide more detailed insight into the topologies of complex assemblies[33,34]. This involves the introduction of a chemical crosslinker into the cells which will create covalent bonds between certain functional groups as they become proximal, thus indicating an interaction (Fig. T5). Not only has this method allowed for the determination of complex interfaces, but it is also capable of capturing transient interactions that may be lost during the washing steps of AP-MS[33,35].

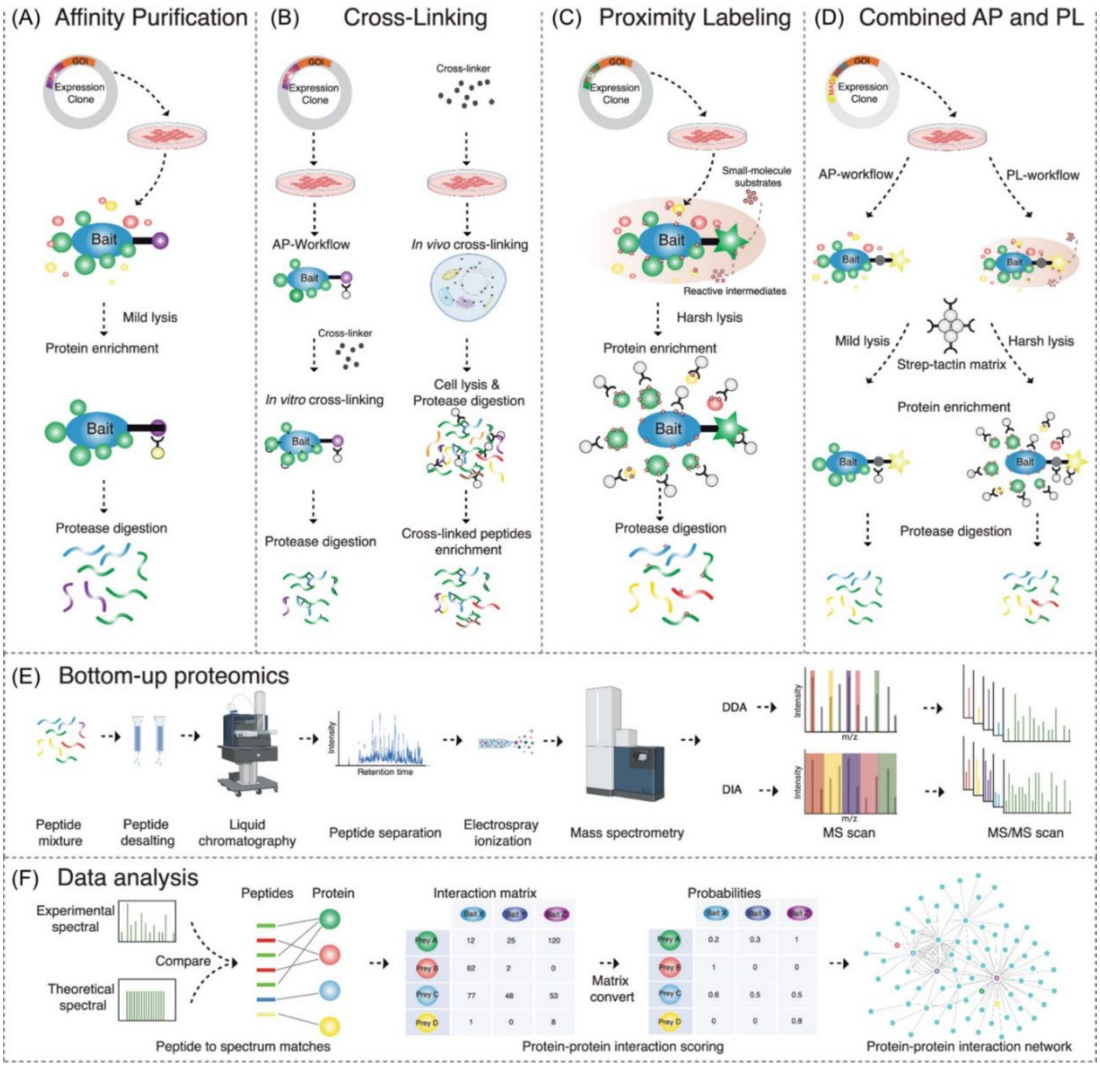


Figure T4 MS-based approaches to studying interactomes. (A) Affinity purification (AP) utilizes an affinity tagged protein of interest (POI) that is produced transiently or stably in selected cell lines. Following mild cell lysis, matrix conjugated to anti-affinity tag antibodies so that the tagged POI and its interacting partners can be selectively enriched. (B) Cross-linking can be performed in vitro following purification of protein complexes or in vivo with intact cells. The cross-linked proteins are then digested to produce cross-linked peptides, which are then enriched prior to MS analysis. (C) Proximity labeling (PL) involves fusion of a promiscuous enzyme to the POI and expression of said fusion in the cell line of choice. To induced in vivo labeling by the enzyme, the substrate is added to the cells allowing it to act as reactive intermediates for PL. The labelled proteins are then enriched. (D) Combination of AP and PL. A POI tagged with a Multiple Approaches Combined tag (MAC-tag) can be utilized for both AP and PL if the culture state and lysis buffer combination is appropriate. The same matrix is then used to enrich protein complexes. (E) The aforementioned techniques utilize top-down proteomics. In bottom-up proteomics peptides derived from proteolytic digestion are first desalted to remove salts that can interfere with analysis. The treated peptide mix is then applied to a liquid chromatography (LC) column, followed by electrospray ionization (EIS). The ionized peptides are then analyzed by MS using two primary strategies: (1) Data-dependent acquisition (DDA) in which ions are scanned, and the most abundant peptides are selected for MS/MS scans and (2) Data-independent acquisition (DIA) which involves the fragmentation of all peptides within a selected mass range. These approaches yield tandem mass spectra for peptide identification and inference of associated proteins. (F) Data analysis begins by comparing experimental spectra against a theoretical database establishing peptide-protein matches from which an interaction matrix is generated. The probabilities of interactions are statistically scored and those interactions identified as high-confidence interactions (HCIs) are extracted via a filtering process and used to construct PPI networks. (*Taken from Liu et al., 2024)*

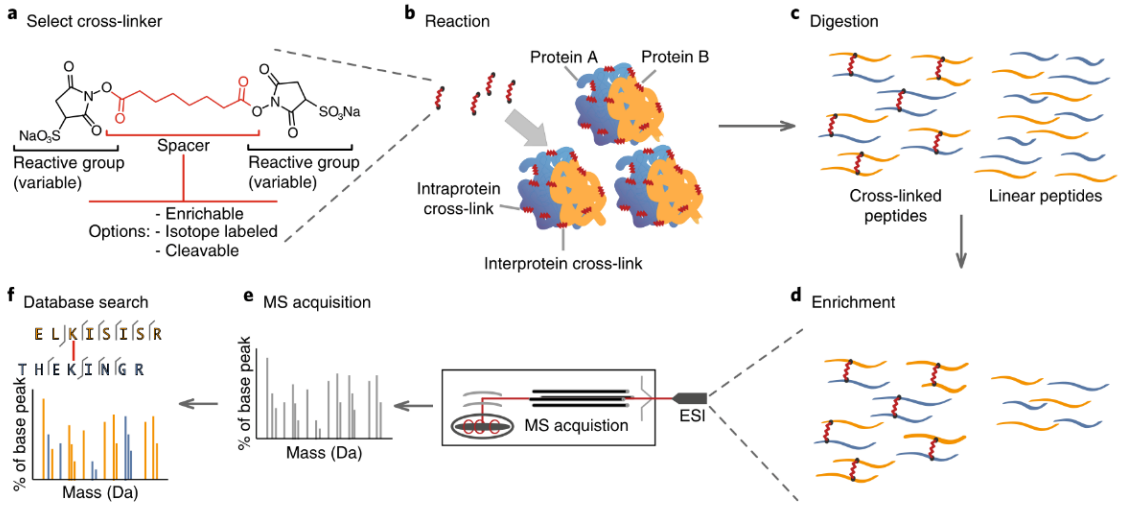


Figure T5 General XL-MS workflow. (A) Select an appropriate cross-linker for XL-MS experiments. Cross-linkers can be variable and depend on experimental set up. They include various chemistries and constructions including spacer lengths, cleavability, and labeling with specific moieties for biochemical enrichment. Reactive groups are also variable. (B) Reaction optimization. Concentrations and reaction times must be determined empirically for each application to encourage the optimal amounts of cross-linking. (C) Protein digestion can be performed in solution or gel producing a mix of cross-linked and linear peptides. (D) Following digestion, cross-linked peptides are typically enriched through chromatographic methods such as size-exclusion (SEC) or ion exchange (IEX). (E) Cross-linked peptide precursors are selected for fragmentation by MS/MS methods. (F) Identification of linked peptides obtained from spectra using computational methods. (*Taken from O'Reilly & Rappsilber, 2018*)

Limitations in standard interactomic methods

The evolution of such a diverse toolbox of high throughput interactomic techniques has made it possible to conduct proteome wide interactomic studies. However, as is true of all experimental methods, each respective method carries its own advantages and disadvantages, and new experimental procedures often emerge to cover the gaps left by another technique (Fig. T6). For example, the results of AP-MS experiments tend to be enriched in strong interactions formed by stable complexes and cannot distinguish between direct or indirect interactions. As a result, proximity labelling emerged as a way to capture transient interactions, but this method still can't differentiate indirect interactions[36]. Cross-linking MS has the benefit of capturing both stable and transient interactions and providing structural details that can determine if an interaction is direct, but the relative scarcity of cross-links formed during experiments represents a barrier to effectively employing this method[37]. While combining interactomic techniques can be an effective strategy to circumvent limitations of individual methods, this can create new challenges in the verification of interactions across methodologies. In fact, this represents a key challenge faced by the field of interactomics, as a whole, regardless of technical strategy: reproducibility.

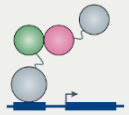
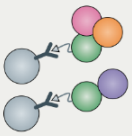
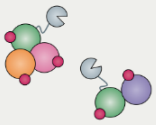
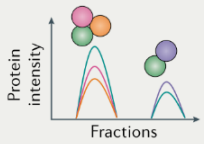
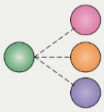
a	Yeast two-hybrid screens	Affinity purification coupled to mass spectrometry	Proximity labelling (BioID, APEX)	Cross-linking coupled to mass spectrometry	Co-fractionation coupled to mass spectrometry
Screening approach					
Measurement	Yeast phenotype	MS/MS	MS/MS	MS/MS	MS/MS
Requirements	Genetic engineering of fusion genes	Genetic tagging or available antibody	Genetic tagging with enzyme	None	None
Type of interactions	Binary PPIs	Binary PPIs and complexes (only via reciprocal pull-downs)	Proximal proteins	Binary PPIs, complexes and structural information	Binary PPIs and complexes
	Stable interactions	Very stable interactions	Transient interactions (down to ~1 min labelling speed)	Short-term interactions (~1 h cross-linking time)	Very stable interactions
Information from single experiment					
Proteome coverage	Only two proteins per assay; high-throughput multiplexing is possible to achieve high proteome coverage	Proteome coverage depends on selected bait proteins	Proteome coverage depends on selected bait proteins	Good proteome coverage in a single assay	Good proteome coverage in a single assay

Figure T6 Strategies for generating proteome-wide interactome maps. Medium through-put Y2H methods were the original standard for investigating PPIs (represented by solid lines between proteins). Advances in MS based proteomics pushed AP-MS to the forefront of interactomics as a state-of-the-art method for large-scale interactome mapping. Development of PL methods soon followed to identify more transient and proximal interactions with methods such as BioID or APEX which identify proteins in close spatiotemporal proximity (represented by dashed lines between proteins). XL-MS also provides the opportunity to capture transient interactions as well as provide more insight into the structural topologies of protein complexes (represented by proteins touching each other). More recently protein co-fractionation coupled to MS (CF-MS) has emerged as a promising new strategy. The process involves biochemical fractionation of cellular lysates to isolate macromolecular complexes without the need for affinity tagging or capture (*Taken from Bladau & Aebersold, 2020*)

Towards quantitative interactomic approaches

A longstanding challenge in interactomics is low overlap between independent studies and limited correlation among orthogonal methods. An analysis comparing large-scale Y2H screens reported a <15% overlap in identified interactions[38]. In general, much of this was attributed to inherent challenges of the Y2H system creating noisy datasets and newly developed techniques, such as AP-MS, were regarded as more dependable. Yet, the reliability of MS-based proteomics has also been called into question due to lack of reproducibility across labs which effects many interactomic methods as they rely on MS readout[39,40]. In the context of AP-MS, studies have shown that variation can be minimized through exacting standardized experimental procedures[41]. However, this doesn't remedy artifacts found within AP-MS methods[42] nor does it propose a way to minimize discrepancies across orthogonal methods. Investigations of published interaction databases have identified limited overlap across datasets[4,43] and have demonstrated that as low as 25% of reported interactions from AP-MS are reproducible[41]. Despite continuing development and refinement of interactomic methods many of these advances don't address the underlying cause of inconsistencies in reproducibility.

Difficulty in experimentally reproducing interactions stems from the fact that interactions are inherently difficult to detect. Unlike direct detection of molecules, interactions fall over a wide range of strengths and specificities[6]. As such, the interactions that can be detected by an assay will be subject to limitation by its sensitivity and detection threshold. Without clear criteria to define what constitutes an interaction their identification becomes problematic, especially at low signal where the distinction between interactors and contaminants becomes unclear[44]. Further complicating the issue, different methods employ the detection of different signals from different origins creating incongruence when attempting to make direct comparisons of the readout from different experiments. In order to increase continuity among datasets and alleviate issues related to reproducibility, a measurable standard metric should be applied across interactomic techniques. In doing so, this will also help establish clear parameters for defining interactions, increasing confidence in their identification. In this next section I present the manuscript for a review, prepared by myself and colleagues, which delves deeper into this topic proposing and discussing in detail the advantages of quantitative interactomic approaches for measuring protein binding affinities as a standard metric in interactomics.

Kathleen Weimer, Boglarka Zambo, & Gergo Gogl (2023). Molecules interact. But how strong and how much?. BioEssays.

Check for updates

Received: 12 January 2023 | Revised: 20 March 2023 | Accepted: 21 March 2023

DOI: 10.1002/bies.202300007

BioEssays

THINK AGAIN

Insights & Perspectives

Molecules interact. But how strong and how much?

Kathleen Weimer | **Boglarka Zambo** | **Gergo Gogl**

Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), INSERM U1258/CNRS UMR 7104/Université de Strasbourg, Illkirch, France

Correspondence
Gergo Gogl, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), INSERM U1258/CNRS UMR 7104/Université de Strasbourg, 1 rue Laurent Fries, BP 10142, F-67404 Illkirch, France.
Email: goglg@igbmc.fr

Abstract

Interactomics aims to characterize all interactions formed between molecules that comprise our body. Although it emerged from quantitative biophysics, it has devolved into a predominantly qualitative field of science over the past decades. Due to technical limitations at its onset, almost all tools in interactomics are qualitative, which persists in defining the discipline. Here, we argue that interactomics needs to return to a quantitative direction because the technical achievements of the last decade have overcome the original limitations that forced its current path. In contrast to qualitative interactomics which is constrained to charting lists of observed interactions, quantitative interactomics can also uncover answers to key questions such as the strength of interactions or how many of certain complexes can form in cells, thus providing researchers with more immediate proxies for understanding and predicting biological processes.

KEYWORDS
interactomics, protein–protein interactions, quantitative biology

INTRODUCTION

Biological processes are governed by noncovalent complexes at all levels of cellular life. Thanks to high throughput interaction mapping, we can now view the human protein–protein interactome with remarkable breadth and depth; yet, our view remains incomplete because current practices fail to incorporate the quantitative properties of these interactions. Due to recent advances, we can now also begin to integrate a quantitative perspective and understand the interactome on a much more nuanced and complete level than ever before.

Fundamental principles underpinning the formation of complexes have been formulated in the past centuries by pioneers in physical-chemistry and biophysics.[1] Modern biochemists carry on their legacy by developing refined theoretical and experimental approaches to investigate the thermodynamic, kinetic, or equilibrium binding properties of noncovalent complexe.[2–5] Although these methods are both precise and accurate, they require specialized expertise. Consequently, only a handful of interactions could be studied using biophysical approaches. This changed rapidly when biology entered the genomic era, allowing scientists to develop alternative, more accessible approaches for investigating noncovalent interactions in order to map entire interactomes proteome-wide.[6,7]

The rift that has emerged between this new field of interactomics and conventional biophysics is profound (Figure 1). Biophysics seeks to characterize complexes through their binding properties, whereas in contrast interactomics lists all interactions that form detectably between molecules of our cells. As a result, interactomics became a qualitative field of science where molecules either "bind" or "do not bind" and interactomes are routinely displayed, analyzed, and even predicted as simple "social networks of biomolecules".[8–10] In this work we present key properties and evaluate limitations of interactomics from the perspective of biochemistry, thus demonstrating the advantages of quantitative interactomics over current qualitative practices and reasoning why the field needs to find a way back to its biophysical roots.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2023 The Authors. *BioEssays* published by Wiley Periodicals LLC.

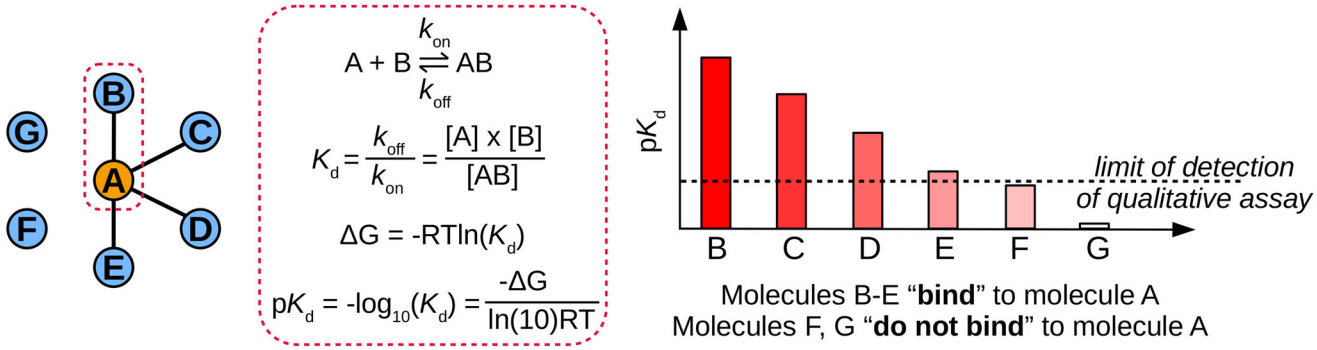


FIGURE 1 Qualitative interactomic, biochemical, and quantitative interactomic description of molecular interactions. On the left panel, an imaginary interaction network of "molecule A" is shown indicating six molecules, of which molecules B–E detectably bind based on qualitative interactomic measurements. All of these interactions, even the ones that do not appear as a binder in the first panel, can be described with a set of intrinsic binding constants, as exemplified in the middle panel. Such intrinsic biophysical binding constants constitute the affinity interactome, as exemplified on the right panel for the same interaction network shown in the left panel. Quantitative affinity-interatomic profiles immediately enable the ranking of the observed interactions which is not possible with simple qualitative interactomic network graphs. In addition, the detection thresholds of quantitative assays are better defined than "binding thresholds" of qualitative interactomic assays, increasing the reproducibility and reliability of interactomes.

QUALITATIVE INTERACTOMICS AND REMNANTS OF QUANTITATIVE FEATURES

A key milestone in the transition of biophysics into interactomics was the first description of a two-hybrid experiment in 1989, which allowed the discovery of complexes at a previously unprecedented pace.[11,12] Two-hybrid measures the possibility of interactions between pairs of molecules expressed in engineered environments. The Human Reference Interactome project applied this method between cca. 17,500 protein molecules and measured 53,000 interactions formed between human proteins.[13] Although these interactions were detected by the two-hybrid assay, the reacting molecules may never actually be present in the same environment in concentrations sufficient to form complexes in large quantities. For this reason, many alternative approaches were invented to measure the in situ amount of complexes in cells, or entire complexomes.

In the OpenCell project, affinity purification coupled to mass spectrometry (AP-MS) was used to find interaction partners of proteins labeled at endogenous levels, identifying 30,000 potential complexes using cca. 1000 protein baits.[14] In contrast to complexes identified by two-hybrid, partners of these interactions are certainly present at an adequate concentration in normal cellular context and are therefore able to form detectable amounts of complexes. However, with these methods, interactions are subject to being overlooked due to the detection threshold of mass spectrometry. To circumvent this, the BioPlex project adopted a similar assay, but with overexpressed baits, managing to identify more than 120,000 complexes with approximately 10,000 protein baits.[15,16] More recently, proximity labeling-based approaches, such as APEX or BioID, have become attractive for interactomic studies because these method often succeed in identifying even those highly transient interactions that are typically not detected by standard AP-MS or 2-hybrid experiments.[17,18] In these assays, bait proteins are produced tagged with biotin ligase in situ and the biotinylated proteins are captured using standard pulldown methods.[19] Finally, while two-hybrid, AP-MS and BioID are directed approaches (focused on a specific pair of molecules or a specific bait), co-fractionation based methods attempt to isolate and identify all complexes directly from cellular extracts that are stable enough to stay together during fractionation.[20,21]

All methods discussed above generate qualitative interactomic data, that is, lists of molecules that interact detectably. As in all experimental methods, interactomic approaches have a detection threshold. It is based on this limiting factor and the measured signal that we classify pairs of molecules as either "interacting" or "not interacting." However, even if we use the same word for the observed interactions of different assays, the measured signals can be of different origins. The underlying units of measurement of observed "interactions" are either equilibrium binding constants or concentrations of complexes. While the binding constants are intrinsic parameters of a system (i.e., they are constant between different states of the system), the concentration of reactants and ultimately the complex formation are extrinsic parameters (i.e., different in different states of the system).

Some of these methods study intrinsic binding preferences (e.g., 2-hybrid), while other approaches shed light on extrinsic features of cellular complexomes (e.g., AP-MS or co-fractionation). For methods that study extrinsic cellular complexomes or, in other words, relative amounts of complexes, the remnants of quantitative features are often still present within qualitative data. For example, the BioPlex project could show that the measured interaction networks change with cell type.[16] One may also measure the apparent molar ratio of bait and captured prey in AP-MS experiments in order to calculate an apparent "stoichiometry" for every observed interaction.[14,22] Such sophisticated AP-MS based studies can even reveal dynamic changes in the complexome.[23] Co-fractionation can also give an intricate insight into complexomic rearrangements of stable complexes under different proteomic conditions.[24,25] Finally, we recently showed that relative

15211878, 2023, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/bies.202300007 by Cochrane France, Wiley Online Library on [17/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

enrichment values of a given prey bound to multiple baits, immobilized in equal amounts during AP-MS, correlate with their relative intrinsic affinity differences.[26,27] Taken together, this indicates that although it is highly convenient to use qualitative interactomic methods, we should not forget the biophysical basis of complex formation when it comes to interpreting our data.

At last, we need to discuss how these remaining quantitative features in qualitative complexomics can be used to study perturbations in cells or differences between cell types. Although we can certainly use these observations to pinpoint key hotspots of complexomic variations, one has to consider that it is currently impossible to measure cellular complexomes without perturbing them. For example, in any assay where complexes are captured on solid supports and separated from their matrix through washing steps, the binding equilibrium is disrupted, resulting in partial or complete dissociation of obtained complexes in accordance to the applied washing protocols and the kinetic constants of the interactions.[28] We will later discuss how novel methods in quantitative interactomics can help overcome this obstacle to give a more complete view of the complexome.

THE ISSUE OF REPRODUCIBILITY IN QUALITATIVE INTERACTOMICS

From the beginning, qualitative interactomics faces a reproducibility crisis. This was initially realized when the first proteome-scale protein–protein interactome maps were charted in independent studies, revealing surprisingly large discrepancies between their results.[29] Since then, many interactomic studies have encountered the same reproducibility issue and attempt to find approaches to minimize this discrepancy, for example, by standardizing experimental procedures.[7] In part, the apparently poor reproducibility has a clear biological background.[30] For example, as mentioned above, 2-hybrid qualitatively studies intrinsic interactomes, but AP-MS qualitatively studies cellular complexomes. Consequently, 2-hybrid identifies pairs of molecules that could directly interact, while AP-MS identifies complexes that may contain indirectly bound components. Apart from these, much of the inconsistency between results comes from the thresholding of the measured signals.

As aforementioned, although rarely discussed otherwise in biological studies, all experimental methods have various kinds of detection thresholds and other bottlenecks. When these differ and, for example, one compares qualitative interactomes measured with low and high affinity detection thresholds, the overlap is expected to be smaller than when comparing results from two different methods with similar thresholds. Related to this issue is the concern of the negatome – pairs of molecules for which no interaction is observed. The interactions belonging to the negatome were found to have affinities or concentrations below the detection thresholds in experiments.[31,32] One should only concern analyzing the negatome when they can be paired with explicit detection thresholds, otherwise they cannot be relativized. Partners that appear to be "not interacting" could also appear as "interacting" in an assay with a lower detection threshold.[4] As previously established, one way to circumvent such limitations is to use highly standardized methods with similar binding thresholds.[7] Another solution is to completely change experimental strategies, measuring affinity interactomes as opposed to charting qualitative interactomes.

At last, validation metrics of qualitative interactomes are done by qualitatively comparing the data with reference interactomes, such as the calculation of Jaccard indexes or precision/recall values. It was recently shown that the most commonly used reference interactome for human protein–protein interactions (PPIs), CORUM, mostly consists of the highest affinity interactions of our interactome and was simultaneously found to be markedly underrepresented in weaker interactions that are inevitably more abundant in nature.[14,33] If comprehensive affinity interactomic databases would exist, the results of any measurement could be compared with published affinity values by calculating explicit affinity differences directly resulting in standardized, accurate and reproducible $\Delta\Delta G$ values.

ALTERNATIVES TO QUALITATIVE INTERACTOMICS

Many biochemists and biophysicists have already recognized the above limitations of interactomics. For example, the founders of AxCell Biosciences Corp. in the early 2000s were early to recognize the advantages of quantitative interactomics, and attempted to map affinities of every promiscuous domain-motif mediated interaction network by the mid-2000s.[34] Although their goals went unfinished, others proceeded with similar goals and measured (relative or absolute) affinities of large interaction networks of domain-motif interactions.[27,35–39] Scientists today have access to dozens of orthogonal experimental approaches that can quantify affinity interactomes at proteomic scales of either minimal binding fragments or even full-length proteins, but mainstream interactomics still remains mostly qualitative.

Measuring the estimated billions of interactions comprising the full human proteome and their variations can be a daunting task.[27] Taking into consideration that many PPIs are mediated through specific domain-motif interactions, efforts to measure the hundreds of millions of interactions constituting the full fragmental interactome appear to be much more feasible.[40] A common feature of these interactions is that they can be easily studied through their minimal binding fragments, such as using synthetic peptides and small globular domains. Among the numerous fragmentomic tools that researchers can use to study fragmental interactomes, phage display provides the highest throughput, as well as great sensitivity, yet it is only capable of identifying motif-mediated interactions and not quantifying their properties.[41]

Many low throughput techniques are available that can easily determine a handful of fragmentomic interactions with great precision and high accuracy. For example, surface plasmon resonance (SPR), biolayer interferometry (BLI), and isothermal calorimetry (ITC). However, these typically cannot be scaled up, preventing them from being suitable for large scale interactomic studies. Many optical assays can be employed at equally high precision and accuracy, as

15211878, 2023, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/bies.202300007 by Cochrane France, Wiley Online Library on [17/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

well as at higher throughput, such as fluorescence polarization (FP), microscale thermophoresis (MST), or MRBLE-pep.[37,42,43] Although these methods are typically more easily scalable and can be combined with peptide or protein arrays, they are resource intensive and can often be used only with special constraints due to technical limitations. Our recently developed fragmentomic holdup assay does not require any special instrumentation and can be used for any type of interaction determination.[27] We applied this method to measure all possible equilibrium binding constants between 50% of all human PDZ domains and approximately 10% of all predicted human PDZ-binding motifs in a combinatorial way, characterizing more than 65,000 interactions and quantifying more than 18,000 dissociation constants.

Regardless of their obvious benefits, fragmentomic approaches only shed light on localized, site-specific properties of interactions. While a well-defined minimal interacting fragment pair may result in the measurement of a singular high-quality affinity constant, it does not consider the global impact of multiple binding sites (e.g., cooperativity or avidity) or explain the distinct binding properties of proteoforms (resulting in allostery, autoinhibition, etc.).[2,44] To this end, investigating the properties of full-length PPIs are equally necessary. Methods for studying full-length PPIs share many challenges with fragmentomic approaches, but one of the most important is reagent preparation. Virtually all low-, medium-, and even high-throughput methods can be used to study interactions of full-length proteins, but purifying even a handful of active full-length proteins can take years of extensive work. Therefore, although feasible for a few cases, the same methods are not practical for studying interactions of full-length proteins at the proteomic scale. As mentioned above, efforts have been undertaken to map the Human Reference Interactome via the 2-hybrid system.[13] Although this was done with a qualitative version of the assay, a tri-fluorescent 2-hybrid assay was also developed that can study interactions quantitatively.[45] Other fluorescence- or luminescence-based assays have also been adapted to quantitatively study PPIs, such as Forster Resonance Energy Transfer (FRET), AlphaScreen, Homogeneous Time Resolved Fluorescence (HTFR), and Bioluminescence Resonance Energy Transfer (BRET), or using supported membrane sheets.[46–50] Although these methods are capable of characterizing interactions of full-length molecules, they are limited to the study of the properties of a single pair of molecules at a time and, therefore, still difficult to scale to the proteomic level. We have recently developed an approach called native holdup (nHU), which is capable of estimating equilibrium binding constants between a single recombinant bait and all proteins that are detectable from cell extracts.[51] With only a single nHU experiment coupled to a multiplex mass spectrometry readout, we showed that it is possible to get estimations about several thousands of equilibrium binding constants, all measured with full-length endogenous proteins and even large multicomponent complexes, exceeding the scalability of most alternative approaches. Although these approaches allow the measurement of the biophysical properties of intact macromolecular interactions often in their semi-native environment, they always lead to a single apparent affinity constant. However, as we have introduced, when discussing fragmental interactomics, proteins exist in a plethora of states referred to as proteoforms .[52] Regardless, these apparent affinities could readily provide important biological insight even if proteoforms are initially ignored and later could be refined, for example, with targeted fragmentomic characterizations.

In addition to the lack of standardized and scalable techniques that the community has agreed to widely adopt, one of the key barriers preventing the qualitative-to-quantitative transition of interactomics is the lack of universal quantitative affinity interactomic databases.[3] Compendiums of both low- and high-throughput interactomic studies provide extensive lists of interactions that allow scientists today to be able to find dozens to hundreds of putative interaction partners for any given protein of interest without doing any experiments. Public databases compiled of these results, such as IntAct, BioGrid, DIP, or STRING, are of unquestionable value.[53–56] Although some of these databases, like Intact, already record biophysical information when available, they were originally created to handle qualitative data and were only adopted later to capture quantitative information. Consequently, they do not capture all possible information for quantitative interactomics and they display interactomes in the form of qualitative interaction graphs, further solidifying the qualitative notion of the field. Previously developed databases fully dedicated for quantitative interactomics were highly specialized and could only store narrow types of data, usually measured with a particular technique.[27,57–61] Hopefully, as quantitative interactomic approaches become more widespread, new general biophysics-based databases will emerge, and current ones will adopt the quantitative view of the field.

COMPUTATIONAL QUANTITATIVE INTERACTOMICS

The recent machine learning boom has impacted nearly all domains and the field of interactomics is no exception. For over 50 years, predictive modeling of affinities has been an area of research interest, as the determination of drug-target affinities is critical to drug development and computational approaches require less resource investment than experimental ones.[62] Binding affinity prediction most often relies on the analysis of experimental or computationally docked structures of complexes using various algorithms.[63,64] With the recent advent of machine learning-based protein structure prediction tools, for example, AlphaFold, this approach is no longer limited by the availability of an experimentally solved structure or by the accuracy of molecular docking, as affinity predictions can be performed using accurately predicted models generated by advanced machine learning methods.[65] However important it may be to keep in mind that these are simply predictions, they serve as a good basis to evaluate candidates for further validation through experimental interactomic studies. For example, a recent approach used AlphaFold prediction with competing interactions to predict their relative binding affinities, however, their evaluation did not consider that binding affinities are determined not only by bound states, but also their free and all intermediate states as well.[66]

Despite the many applications for machine learning in interactomics, the field is currently self-limited by the mentioned lack of

15211878, 2023, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/bies.202300007 by Cochrane France, Wiley Online Library on [17/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

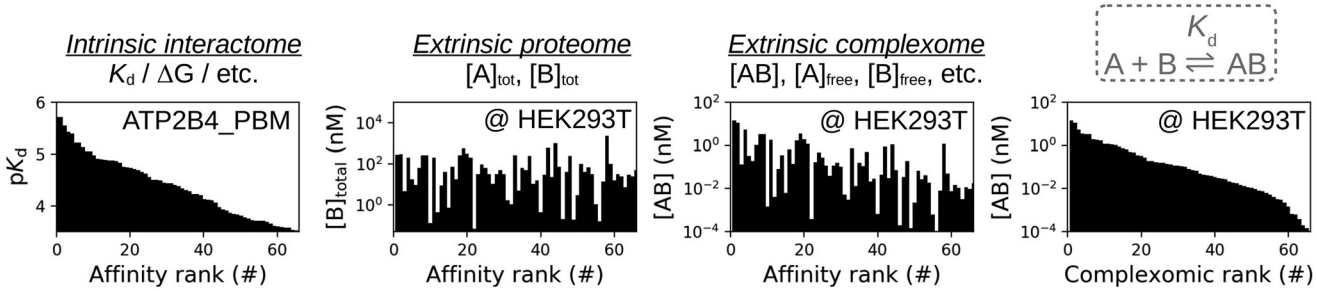


FIGURE 2 From interactomes to complexomes through the proteome. Intrinsic affinity interactomes (far left) can be combined with absolute proteomes of any cell type or state (middle left) in order to calculate the cellular complexome (middle right). This complexome can be used to re-rank partners according to their estimated cellular abundancies (far right). While the units of measurement of intrinsic affinity interactomes are some sort of physical quantities, like binding energies, the units of complexomes are concentrations. These can be also converted to a number of complexes per cell, as approximately 10^{-4} nM molar concentration corresponds to 1 molecule/complex per cell in eukaryotes.[71] Note that for an entire organism, the interactome remains constant, while the proteome and the complexome vary between cells. The intrinsic equilibrium binding affinity profile of the PDZ-binding motif of ATP2B4 was obtained from the ProfAff database and the absolute HEK293T proteome was obtained from the OpenCell database.[14,27] The simplest bimolecular binding model was used to estimate the amounts of complexes.

available affinity data and repositories to store them. Machine learning is a data driven method and it is known that "a dumb algorithm with lots of data beats a clever one with modest amounts of it."[67] The accuracy of this method is highly dependent on the availability of a large, high-quality training data set, where data quality can impact the model's predictive performance and insufficient data can prevent generalizability to new data.[68] That said, quantitative interactomic data naturally lends itself to machine learning and it is expected that experimental and computational quantitative interactomics will grow hand in hand.

ESTIMATING STEADY-STATE OF COMPLEXOMES AND THEIR DYNAMIC CHANGES

Interactions can be studied by their intrinsic (e.g., binding constants) or extrinsic (e.g., cellular concentrations) parameters. Thanks to modern proteomics, we have easy ways to estimate the total bulk concentrations of protein molecules present in cellular extracts.[69] One of the main benefits of quantitative interactomics is that it could be used to convert experimental proteome information into estimated complexomes using computational modeling.

Using first-principle modeling, we can approximate the amounts of formed complexes under binding equilibrium, by combining these absolute concentrations with intrinsic affinities for every molecule pair.[26,27,51,70] Simple estimations of steady states can be very coarse, as they only consider the presence of interacting molecules in isotropic cells at complete binding equilibrium (Figure 2). Still, even an estimated complexome could provide a useful picture of biological mechanisms.[71–74] While quantitative proteomics showed us a reference scale about the total amounts of protein molecules in cells, we still lack such a scale for molecular complexes. Certainly, all molecular complexes are somewhat less abundant as their ingredients, but by how much? Are they roughly equal since the complexome is dominated by extremely stable complexes? Or are only a small fraction of molecules in complex at a time? These simplistic calculations are also sufficient to show how weaker affinity complexes can arise in larger quantities than stronger ones due to higher abundances of their reactants. Since these calculations are simple and fast, they can also be used to rapidly translate observed proteomic-perturbations into complexomic perturbations, revealing how the complexomic landscape is rearranged by various factors, such as by mutations or by different environmental signals.

A more complete complexomic calculation can be performed by extending the number of reactants in the calculations from binary complexes (pairs of molecules) to whole signaling pathways or even the entire complexome, taking into account series of competing-cooperating interactions. For example, when members of the same domain family compete to bind a group of proteins containing a family of cognate recognition motifs, the overall steady-state of the complexomic network is determined by the affinities of all possible interactions, as well as the total concentration of every participating molecule. Using total cellular concentrations and either directly measured kinetic rate constants or arbitrary kinetic constants calculated from equilibrium binding affinities, rule-based binding simulations can be performed to estimate the overall steady-state of the complete system.[75]

Complexomic modeling can also be used to study dynamic cellular processes and not only steady-state binding equilibriums. For example, cellular signal transduction is a dynamic process, where the cellular content, for example, phosphorylation states, continuously changes over time. Quantitative interactomics can also be extended to study the properties of modification-dependent interactions, as we have shown in the past on phosphorylation and acetylation-dependent interactions.[76,77] These, in combination with phosphoproteomic data, can help us predict dynamic changes in complexomes during complex signaling events.[78] However, for future large-scale dynamic modeling, we will also need other parameters in addition to equilibrium binding constants, such as reaction rate constants and other kinetic parameters, whose measurement will likely be the next big challenge of quantitative interactomics.

15211878, 2023, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/bies.202300007 by Cochrane France, Wiley Online Library on [17/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The spatial resolution of any complexomic modeling is bound to the spatial resolution of the proteomic measurement, but not to the intrinsic biophysical properties, which do not have spatial dimensions. Cellular compartmentalization, a property inherent to eukaryotic cells, certainly leads to differential complexomes across the cell. Fortunately, there are alternative approaches to bulk proteomics to quantify amounts of molecules in space.[79,80] Spatial proteome-complexome relationships were already implied based on past qualitative data, and future quantitative progress in the field may reveal the spatial self-organization principles of highly interconnected molecular networks in cells.[14,81] Although we may still be far from reaching the coverage of quantitative affinity interactomes sufficient for this type of complex first-principle modeling, establishing the clear conceptual difference between interactomes and complexomes is definitely the first step towards this goal.

COMPARATIVE QUANTITATIVE INTERACTOMICS AND COMPLEXOMICS

In many instances, interactomics is used to measure similarities or differences in binding properties between molecules. For example, the qualitative interactomes of many near-complete protein families were studied in depth to connect the particular roles of family-members with certain biological processes.[18,82,83] One of the main advantages of quantitative biophysical measurements is that instead of making qualitative deductions based on a simple list of interaction partners, we can use standardized and accurate metrics of similarities, such as ΔΔG values. This is often achieved by studying the affinity interactomes of a series of proteins, or even entire protein families, against a concise set of known interaction partners, such as peptide motifs, using conventional biophysical methods.[37,84] When quantitative interactomic approaches are performed on the proteomic scale, instead of analyzing a few selected individual interaction partners, we can consider entire binding profiles and quantify interactomic similarities through multidimensional affinity spaces by calculating Euclidean distances (e.g., cumulative ΔΔG values). As a result, quantitative interactomics is capable of determining a comprehensible, unbiased and reproducible picture of interactomic similarity, a critical asset to understanding differences between close paralogous proteins or differences that occur in turn due to post-translational modifications, or genetic mutations.

Any change in affinity interactomes, for example, induced by genetic variations, is expected to lead to altered cellular physiology. Nevertheless, much is still not yet understood regarding the interactome-phenotype relationship, because many layers of biological networks, in increasing levels of complexity, span between intrinsic properties and macroscopic phenotypes. Qualitative studies have shown that mutations drastically altering the interactome are often associated with diseases and that different mutations in the same gene causing different interactomic profiles are associated with distinguishable phenotypes.[85] Since these comparative interactomes were studied in a qualitative manner, experimentally surveyed interactomic impacts were categorized as either quasi-WT resulting in no detectable changes, edgetic resulting in partial loss or gain of interactions, or quasi-null resulting in total abolishment of interactions. One clear disadvantage of qualitative assays is that they only consider "interacting" and "not interacting" molecules; therefore, mutations causing relative perturbations in biophysical traits of interactions go undetected. Interactions unaffected by mutations according to qualitative interactomic assays may still display enhanced or weakened affinities, and these unaccounted for effects could also contribute to phenotypes through slightly altered complexomes. The accumulation of small relative variations in affinities or concentrations of several interactions can contribute to the development of disease just as much as large perturbations.

As briefly discussed above, much attention has been paid in the past not only to study the interactomic differences of different proteins or their variants, but also to study differences in the interaction networks formed in different cellular contexts.[9] Since the affinity interactome is an intrinsic property of an organism, biophysical constants remain the same in every cell of our body, unless somatic mutations occur. However, each cell has a unique proteome, which can vary markedly between different tissues; their unique cellular complexomes can be used to measure changes in amounts of formed complexes between different cellular types and states. Related to this problem, pathogenic hijacking is of particular importance. Many host-pathogen interactions occur through PPIs where newly introduced pathogenic components interact and perturb an already established host interaction network.[86,87] As a constant intrinsic physical property, affinities of the host network remain unchanged, albeit the affinity interactomes of certain host proteins are extended with new interactions formed with the hijacking proteins. However, depending on the type of the infected host cell, large variations are expected in hijacked complexomes.[26] In addition, infection can often trigger extensive proteomic changes in the host cell.[27,88] Shifts in the abundance of available partners are reflected during complex formation, leading to different complexomic landscapes at various stages of infection, and studying these changes may not only give us insight into the pathophysiology of infectious disease, but can help identify key targets at critical time points during infection to block pathogenic reprogramming in cells.

UNDERSTANDING SPECIFICITY AND PROMISCUITY IN THE LIGHT OF UNBIASED INTERACTOMES

Specific interactions can be described according to the law of mass action, while nonspecific binding refers to different kinds of interactions, such as those resulting from the Debye-Hückel theory.[1] For simplicity, if we assume specific binding between any pairs of protein molecules and if we only consider a single proteoform per protein, the entire unbiased human affinity PPI interactome consists of 300,000,000 "combinatorial" interactions. Albeit the numerical values of these affinities may slightly change with changing conditions in theory, in practice they can be considered constants because most conditions, such as temperature, pressure, or ionic strength, only vary

15211878, 2023, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/bies.202300007 by Cochrane France, Wiley Online Library on [17/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

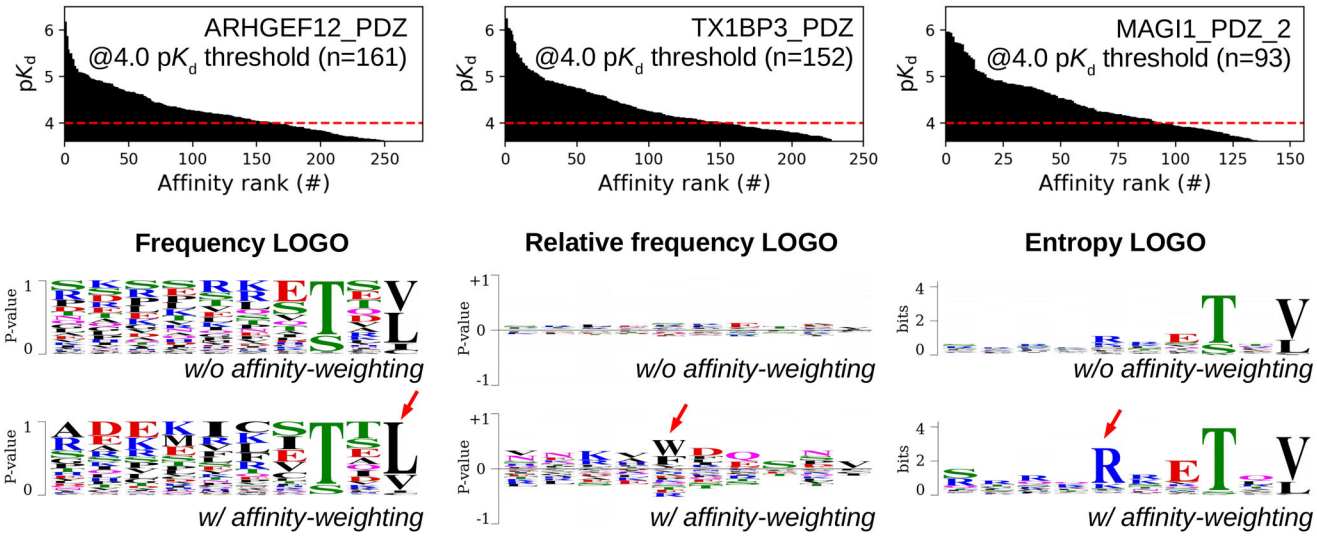


FIGURE 3 Affinity weighting can reveal key determinants of optimal binding. For interactions mediated by short linear motifs, sequence LOGO representation is often used to reveal specificity determinants. We show three examples of different LOGO representations of PDZ interactions with or without affinity weighting using data taken from the ProfAff database.[27] Sequences of interaction partners were taken above a 4.0 pK_d affinity threshold (red dashed line) and the number of partner sequences with affinities above the threshold used for calculations is indicated on the affinity profiles. In affinity-weighted LOGO calculations, the same set of interacting peptide sequences was used but their corresponding affinities were also considered as described in detail before.[27] In frequency LOGO calculations, the frequencies of each amino acid are counted at every motif position. In relative frequency LOGO calculations, the amino acid frequencies of interacting peptides are normalized to the amino acid frequencies of all assayed peptides (including the "non-binder" sequences). In entropy LOGO calculations, the frequencies of each amino acid are converted to Shannon entropies. Red arrows indicate some marked differences that only appear after affinity weighting; preference for Leu over Val at the C-terminal position of ARHGEF12 interacting peptides; preference for aromatic residues of TX1BP3 interacting peptides and the preference of an Arg residue for high-affinity MAGI1 interaction motifs.

within a very narrow range under physiological conditions, with the notable exception of concentrations of certain key components, such as Mg^{2+} or H_3O^+ ions. While most of these combinatorial interactions are likely to display extremely weak affinities, they are unlikely to lead to substantial complex formation under physiological cellular conditions. However, even these interactions are likely to contribute to the formation of cellular complexomes and therefore should be considered when discussing specificity or promiscuity.[89]

Based on more than 65,000 measured interactions including more than 18,000 quantified steady-state fragmentomic affinities of PDZ domains, we have found that the binding free energies follow an exponential distribution that may relate to the Boltzmann distribution, similarly to the distribution of other energy-related properties of molecules and interactions.[27] Although this observation was made for highly similar domain-motif interactions, the affinities of more diverse types of interactions show similar trends.[51] Consequently, affinity interactomes are dominated by weak interactions, and only a small fraction of all detectable partners mediate the strongest affinities. Based on the overall shape of such affinity profile or, in other words, based on the affinity distribution of the interactome, the promiscuity of the selected molecule can be determined[77] (Figure 3). A highly promiscuous molecule has many interaction partners with similar affinities, none of which display outstanding affinity. In contrast, a less promiscuous molecule will have at least one interaction partner, which will display much greater affinity than the others. While it is tempting to call such a molecule highly specific or to say that its interactome is specific, it should be kept in mind that even the weaker affinity interactions are specific by definition. Therefore, from a pragmatic biochemical perspective, it is highly improbable to find molecules with exclusively specific interactomic profiles, and it is plausible that even the most specific molecules will display weak affinities against all other molecules present in our body.

Since quantitative interactomics is most advanced for domain-motif interactions, we could not avoid mentioning the specificity determinants of such interactions in the light of affinity profiles. Conventionally, binding preferences of these interactions are often studied with bioinformatic approaches, such as position-specific scoring matrices (PSSM), and specificity determinants are often visualized in the form of sequence LOGOs.[90–92] However, most of these approaches do not discriminate between interaction partners that display strong or weak affinities and only define "binders" and "nonbinders." We previously proposed that such calculations could be modified in a way to take affinity information into account and this way, affinity-weighted PSSMs or LOGO representation could be an important asset to visualize determinants of high affinity interactions.[27]

CONCLUSION

Quantitative interactomics has already shown in many cases how powerful can it be in answering key biological questions. In this essay, instead of providing a comprehensive review about these major

15211878, 2023, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/bies.202300007 by Cochrane France, Wiley Online Library on [17/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

achievements, we tried to focus on important limitations of current trends and provide an alternative perspective that could go beyond the boundaries of current interactomics. In the light of recent advances, many of the original technical limitations of interactomics, which can stand in the way of quantitative measurements, have almost been overcome. We believe that the time has come for interactomics to return to its quantitative roots. Qualitative interactomics has already given us a nearly comprehensive map of the molecular wiring of our cells.[93] Although there will be much to come until proteome-wide comprehensive affinity maps will be available, on a case-by-case basis, researchers can already consider alternative quantitative approaches over conventional ones. With the help of these, there are deeper questions that can be addressed, such as how strong are the intrinsic affinities of complexes of cellular molecules or how much can be found from them under given cellular conditions. Ultimately, turning the field into a more quantitative direction will help us to describe the cellular distribution of complexes and their alterations in disease.

ACKNOWLEDGMENTS

We are grateful to all our colleagues and friends for their enlightening discussions and support. In particular, we would like to thank Gilles Trave for our stimulating discussions, which have inspired most of the ideas in this article.

CONFLICT OF INTEREST STATEMENT

The authors have no conflict of interest to declare.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

ORCID

Gergo Gogl https://orcid.org/0000-0002-8597-3711

REFERENCES

1. Wyman, J., & Gill, S. J. (1990). *Binding and linkage*. University Science Books, Mill Valley, CA.
2. Ivarsson, Y., & Jemth, P. (2019). Affinity and specificity of motif-based protein–protein interactions. *Current Opinion in Structural Biology*, *54*, 26–33. https://doi.org/10.1016/j.sbi.2018.09.009
3. Jarmoskaite, I., Alsadhan, I., Vaidyanathan, P. P., & Herschlag, D. (2020). How to measure and evaluate binding affinities. *ELife*, *9*, 1–34. https://doi.org/10.7554/ELIFE.57264
4. Pollard, T. D. (2010). A guide to simple and informative binding assays. *Molecular Biology of the Cell*, *21*(23), 4061–4067. https://doi.org/10.1091/mbc.E10-08-0683
5. Corzo, J., & Santamaria, M. (2006). Time, the forgotten dimension of ligand binding teaching. *Biochemistry and Molecular Biology Education*, *34*(6), 413–416. https://doi.org/10.1002/bmb.2006.494034062678
6. Bludau, I., & Aebersold, R. (2020). Proteomic and interactomic insights into the molecular basis of cell functional diversity. *Nature Reviews Molecular Cell Biology*, *21*(6), 327–340. https://doi.org/10.1038/s41580-020-0231-2
7. Varjosalo, M., Sacco, R., Stukalov, A., Van Drogen, A., Planyavsky, M., Hauri, S., Aebersold, R., Bennett, K. L., Colinge, J., Gstaiger, M., & Superti-Furga, G. (2013). Interlaboratory reproducibility of large-scale human protein-complex analysis by standardized AP-MS. *Nature Methods*, *10*(4), 307–314. https://doi.org/10.1038/nmeth.2400
8. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: A software environment for integrated models. *Genome Research*, *13*(11), 2498–2504. https://doi.org/10.1101/gr.1239303
9. Vidal, M., Cusick, M. E., & Barabási, A. L. (2011). Interactome networks and human disease. *Cell*, *144*(6), 986–998. https://doi.org/10.1016/j.cell.2011.02.016
10. Kovács, I. A., Luck, K., Spirohn, K., Wang, Y., Pollis, C., Schlabach, S., Bian, W., Kim, D. K., Kishore, N., Hao, T., Calderwood, M. A., Vidal, M., & Barabási, A. L. (2019). Network-based prediction of protein interactions. *Nature Communications*, *10*(1), 1–8. https://doi.org/10.1038/s41467-019-09177-y
11. Fields, S., & Song, O. K. (1989). A novel genetic system to detect protein-protein interactions. *Nature*, *340*(6230), 245–246. https://doi.org/10.1038/340245a0
12. Young, K. H. (1998). Yeast two-hybrid: So many interactions, (in) so little time. *Biology of Reproduction*, *58*(2), 302–311. https://doi.org/10.1095/biolreprod58.2.302
13. Luck, K., Kim, D. K., Lambourne, L., Spirohn, K., Begg, B. E., Bian, W., Brignall, R., Cafarelli, T., Campos-Laborie, F. J., Charloteaux, B., Choi, D., Coté, A. G., Daley, M., Deimling, S., Desbuleux, A., Dricot, A., Gebbia, M., Hardy, M. F., Kishore, N., … Calderwood, M. A. (2020). A reference map of the human binary protein interactome. *Nature*, *580*(7803), 402–408. https://doi.org/10.1038/s41586-020-2188-x
14. Cho, N. H., Cheveralls, K. C., Brunner, A.-D., Kim, K., Michaelis, A. C., Raghavan, P., Kobayashi, H., Savy, L., Li, J. Y., Canaj, H., Kim, J. Y. S., Stewart, E. M., Gnann, C., McCarthy, F., Cabrera, J. P., Brunetti, R. M., Chhun, B. B., Dingle, G., Hein, M. Y., … Leonetti, M. D. (2022). OpenCell: Endogenous tagging for the cartography of human cellular organization. *Science*, *375*, eabi6983. https://doi.org/10.1126/science.abi6983
15. Huttlin, E. L., Bruckner, R. J., Paulo, J. A., Cannon, J. R., Ting, L., Baltier, K., Colby, G., Gebreab, F., Gygi, M. P., Parzen, H., Szpyt, J., Tam, S., Zarraga, G., Pontano-Vaites, L., Swarup, S., White, A. E., Schweppe, D. K., Rad, R., Erickson, B. K., … Harper, J. W. (2017). Architecture of the human interactome defines protein communities and disease networks. *Nature*, *545*(7655), 505–509. https://doi.org/10.1038/nature22366
16. Huttlin, E. L., Bruckner, R. J., Navarrete-Perea, J., Cannon, J. R., Baltier, K., Gebreab, F., Gygi, M. P., Thornock, A., Zarraga, G., Tam, S., Szpyt, J., Gassaway, B. M., Panov, A., Parzen, H., Fu, S., Golbazi, A., Maenpaa, E., Stricker, K., Guha Thakurta, S., … Gygi, S. P. (2021). Dual proteome-scale networks reveal cell-specific remodeling of the human interactome. *Cell*, *184*(11), 3022–3040.e28. https://doi.org/10.1016/j.cell.2021.04.011
17. Rhee, H., Zou, P., Udeshi, N. D., Martell, J. D., Mootha, V. K., Carr, S. A., & Ting, A. Y. (2013). Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. *Science*, *339*(3), 1328.
18. Göös, H., Kinnunen, M., Salokas, K., Tan, Z., Liu, X., Yadav, L., Zhang, Q., Wei, G. H., & Varjosalo, M. (2022). Human transcription factor protein interaction networks. *Nature Communications*, *13*(1), 1–16. https://doi.org/10.1038/s41467-022-28341-5
19. Roux, K. J., Kim, D. I., Burke, B., & May, D. G. (2018). BioID: A screen for protein-protein interactions. *Current Protocols in Protein ScienceCurrent Protocols in Protein Science*, *91*(1), 19.23.1–19.23.15. https://doi.org/10.1002/cpps.51
20. Skinnider, M. A., & Foster, L. J. (2021). Meta-analysis defines principles for the design and analysis of co-fractionation mass spectrometry experiments. *Nature Methods*, *18*(7), 806–815. https://doi.org/10.1038/s41592-021-01194-4
21. Havugimana, P. C., Goel, R. K., Phanse, S., Youssef, A., Padhorny, D., Kotelnikov, S., Kozakov, D., & Emili, A. (2022). Scalable multiplex co-fractionation/mass spectrometry platform for accelerated protein interactome discovery. *Nature Communications*, *13*(1), 1–8. https://doi.org/10.1038/s41467-022-31809-z

15211878, 2023, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/bies.202300007 by Cochrane France, Wiley Online Library on [17/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

22. Hein, M. Y., Hubner, N. C., Poser, I., Cox, J., Nagaraj, N., Toyoda, Y., Gak, I. A., Weisswange, I., Mansfeld, J., Buchholz, F., Hyman, A. A., & Mann, M. (2015). A Human Interactome in three quantitative dimensions organized by stoichiometries and abundances. *Cell*, *163*(3), 712–723. https://doi.org/10.1016/j.cell.2015.09.053
23. Ciuffa, R., Uliana, F., Mannion, J., Mehnert, M., Tenev, T., Marulli, C., Satanowski, A., Keller, L. M. L., Ramírez, P. N. R., Ori, A., Gstaiger, M., Meier, P., & Aebersold, R. (2022). Novel biochemical, structural, and systems insights into inflammatory signaling revealed by contextual interaction proteomics. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(40), 1–12. https://doi.org/10.1073/pnas.2117175119
24. Salas, D., Stacey, R. G., Akinlaja, M., & Foster, L. J. (2020). Next-generation interactomics: Considerations for the use of co-elution to measure protein interaction networks. *Molecular and Cellular Proteomics*, *19*(1), 1–10. https://doi.org/10.1074/mcp.R119.001803
25. Bludau, I., Nicod, C., Martelli, C., Xue, P., Heusel, M., Fossati, A., Uliana, F., Frommelt, F., Aebersold, R., & Collins, B. C. (2022). Rapid profiling of protein complex re-organization in perturbed systems. *BioRxiv*, 1–33. https://www.biorxiv.org/content/10.1101/2021.12.17.473177v2
26. Gogl, G., Tugaeva, K., Eberling, P., Kostmann, C., Trave, G., & Sluchanko, N. N. (2021). Hierarchized phosphotarget binding by the seven human 14-3-3 isoforms. *Nature Communications*, *12*(1677), 1–12. https://doi.org/10.1038/s41467-021-21908-8
27. Gogl, G., Zambo, B., Kostmann, C., Cousido-Siah, A., Morlet, B., Durbesson, F., Negroni, L., Eberling, P., Jane, P., Nomine, Y., Zeke, A., Ostergaard, S., Monsellier, E., Vincentelli, R., & Trave, G. (2022). Quantitative fragmentomics allow affinity mapping of interactomes. *Nature Communications*, *13*(1), 1–18.
28. Luck, K., & Travé, G. (2011). Phage display can select over-hydrophobic sequences that may impair prediction of natural domain-peptide interactions. *Bioinformatics*, *27*(7), 899–902. https://doi.org/10.1093/bioinformatics/btr060
29. Goll, J., & Uetz, P. (2006). The elusive yeast interactome. *Genome Biology*, *7*(6), 223. https://doi.org/10.1186/gb-2006-7-6-223
30. Sharifi Tabar, M., Parsania, C., Chen, H., Su, X. D., Bailey, C. G., & Rasko, J. E. J. (2022). Illuminating the dark protein-protein interactome. *Cell Reports Methods*, *2*(8), 100275. https://doi.org/10.1016/j.crmeth.2022.100275
31. Mei, S., & Zhang, K. (2019). Neglog: Homology-based negative data sampling method for genome-scale reconstruction of human protein–protein interaction networks. *International Journal of Molecular Sciences*, *20*(20), 5075. https://doi.org/10.3390/ijms20205075
32. Blohm, P., Frishman, G., Smialowski, P., Goebels, F., Wachinger, B., Ruepp, A., & Frishman, D. (2014). Negatome 2.0: A database of non-interacting proteins derived by literature mining, manual annotation and protein structure analysis. *Nucleic Acids Research*, *42*(D1), D396–D400. https://doi.org/10.1093/nar/gkt1079
33. Giurgiu, M., Reinhard, J., Brauner, B., Dunger-Kaltenbach, I., Fobo, G., Frishman, G., Montrone, C., & Ruepp, A. (2019). CORUM: The comprehensive resource of mammalian protein complexes –2019. *Nucleic Acids Research*, *47*(D1), D559–D563. https://doi.org/10.1093/nar/gky973
34. Hu, H., Columbus, J., Zhang, Y., Wu, D., Lian, L., Yang, S., Goodwin, J., Luczak, C., Carter, M., Chen, L., James, M., Davis, R., Sudol, M., Rodwell, J., & Herrero, J. J. (2004). A map of WW domain family interactions. *Proteomics*, *4*(3), 643–655. https://doi.org/10.1002/pmic.200300632
35. Smith, M. J., Hardy, W. R., Murphy, J. M., Jones, N., & Pawson, T. (2006). Screening for PTB domain binding partners and ligand specificity using proteome-derived NPXY peptide arrays. *Molecular and Cellular Biology*, *26*(22), 8461–8474. https://doi.org/10.1128/mcb.01491-06
36. Stiffler, M. A., Chen, J. R., Grantcharova, V. P., Lei, Y., Fuchs, D., Allen, J. E., Zaslavskaia, L. A., & MacBeath, G. (2007). PDZ domain binding selectivity is optimized across the mouse proteome. *Science*, *317*(10), 364–370.
37. Simon, M., Gógl, G., Ecsédi, P., Póti, Á., Kardos, J., & Nyitray, L. (2020). High throughput competitive fluorescence polarization assay reveals functional redundancy in the S100 protein family. *FEBS Journal*, *287*(13), 1–13. https://febs.onlinelibrary.wiley.com/doi/10.1111/febs.15175
38. Delalande, F., Gogl, G., Rohrbacher, A., Kostmann, C., Eberling, P., Carapito, C., Travé, G., & Monsellier, E. (2022). The Holdup Multiplex, an assay for high-throughput measurement of protein- ligand affinity constants using a mass-spectrometry readout. *BioRxiv*, 1–36.
39. Hause, R. J., Leung, K. K., Barkinge, J. L., Ciaccio, M. F., Chuu, C., & Jones, R. B. (2012). Comprehensive binary interaction mapping of SH2 domains via fluorescence polarization reveals novel functional diversification of ErbB receptors. *PLoS ONE*, *7*(9), e44471. https://doi.org/10.1371/journal.pone.0044471
40. Tompa, P., Davey, N. E., Gibson, T. J., & Babu, M. M. (2014). A million peptide motifs for the molecular biologist. *Molecular Cell*, *55*(2), 161–169. https://doi.org/10.1016/j.molcel.2014.05.032
41. Benz, C., Ali, M., Krystkowiak, I., Simonetti, L., Sayadi, A., Mihalic, F., Kliche, J., Andersson, E., Jemth, P., Davey, N. E., & Ivarsson, Y. (2022). Proteome-scale mapping of binding sites in the unstructured regions of the human proteome. *Molecular Systems Biology*, *18*(1), 1–27. https://doi.org/10.15252/msb.202110584
42. Jerabek-Willemsen, M., André, T., Wanner, R., Roth, H. M., Duhr, S., Baaske, P., & Breitsprecher, D. (2014). MicroScale Thermophoresis: Interaction analysis and beyond. *Journal of Molecular Structure*, *1077*, 101–113. https://doi.org/10.1016/j.molstruc.2014.03.009
43. Nguyen, H. Q., Roy, J., Harink, B., Damle, N. P., Latorraca, N. R., Baxter, B. C., Brower, K., Longwell, S. A., Kortemme, T., Thorn, K. S., Cyert, M. S., & Fordyce, P. M. (2019). Quantitative mapping of protein-peptide affinity landscapes using spectrally encoded beads. *ELife*, *8*(e40499), 1–28. https://doi.org/10.7554/eLife.40499
44. Stefan, M. I., & Le Novère, N. (2013). Cooperative binding. *PLoS Computational Biology*, *9*(6), 2–7. https://doi.org/10.1371/journal.pcbi.1003106
45. Cluet, D., Amri, I., Vergier, B., Léault, J., Audibert, A., Grosjean, C., Calabrési, D., & Spichty, M. (2020). A quantitative tri-fluorescent yeast two-hybrid system: From flow cytometry to in cellula affinities. *Molecular and Cellular Proteomics*, *19*(4), 701–715. https://doi.org/10.1074/mcp.TIR119.001692
46. Song, Y., Madahar, V., & Liao, J. (2011). Development of FRET assay into quantitative and high-throughput screening technology platforms for protein-protein interactions. *Annals of Biomedical Engineering*, *39*(4), 1224–1234. https://doi.org/10.1007/s10439-010-0225-x
47. Yasgar, A., Jadhav, A., Simeonov, A., & Coussens, N. P. (2016). AlphaScreen-based assays: Ultra-high-throughput screening for small-molecule inhibitors of challenging enzymes and protein-protein interactions. *High Throughput Screening*, *1*, 77–98. https://doi.org/10.1007/978-1-4939-3673-1
48. Degorce, F., Card, A., Soh, S., Trinquet, E., Knapik, G. P., & Xie, B. (2009). HTRF: A technology tailored for drug discovery – A review of theoretical aspects and recent applications. *Current Chemical Genomics*, *3*(1), 915–933. https://doi.org/10.1007/s10236-017-1064-1
49. Trepte, P., Kruse, S., Kostova, S., Hoffmann, S., Buntru, A., Tempelmeier, A., Secker, C., Diez, L., Schulz, A., Klockmeier, K., Zenkner, M., Golusik, S., Rau, K., Schnoegl, S., Garner, C. C., & Wanker, E. E. (2018). LuTHy: A double-readout bioluminescence-based two-hybrid technology for quantitative mapping of protein–protein interactions in mammalian cells. *Molecular Systems Biology*, *14*(7), 1–20. https://doi.org/10.15252/msb.20178071
50. Erlendsson, S., Thorsen, T. S., Vauquelin, G., Ammendrup-Johnsen, I., Wirth, V., Martinez, K. L., Teilum, K., Gether, U., & Madsen, K. L. (2019). Mechanisms of pdz domain scaffold assembly illuminated by use of supported cell membrane sheets. *ELife*, *8*, 1–27. https://doi.org/10.7554/eLife.39180

15211878, 2023, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/bies.202300007 by Cochrane France, Wiley Online Library on [17/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

51. Zambo, B., Morlet, B., Negroni, L., Trave, G., & Gogl, G. (2022). Native holdup (nHU) to measure binding affinities from cell extracts. *Science Advances*, *8*(51), eade3828.
52. Aebersold, R., Agar, J. N., Amster, I. J., Baker, M. S., Bertozzi, C. R., Boja, E. S., Costello, C. E., Cravatt, B. F., Fenselau, C., Garcia, B. A., Ge, Y., Gunawardena, J., Hendrickson, R. C., Hergenrother, P. J., Huber, C. G., Ivanov, A. R., Jensen, O. N., Jewett, M. C., Kelleher, N. L., … Zhang, B. (2018). How many human proteoforms are there? *Nature Chemical Biology*, *14*(3), 206–214. https://doi.org/10.1038/nchembio.2576
53. Orchard, S., Ammari, M., Aranda, B., Breuza, L., Briganti, L., Broackes-Carter, F., Campbell, N. H., Chavali, G., Chen, C., Del-Toro, N., Duesbury, M., Dumousseau, M., Galeota, E., Hinz, U., Iannuccelli, M., Jagannathan, S., Jimenez, R., Khadake, J., Lagreid, A., … Hermjakob, H. (2014). The MIntAct project – IntAct as a common curation platform for 11 molecular interaction databases. *Nucleic Acids Research*, *42*(D1), D358–D363. https://doi.org/10.1093/nar/gkt1115
54. Oughtred, R., Rust, J., Chang, C., Breitkreutz, B. J., Stark, C., Willems, A., Boucher, L., Leung, G., Kolas, N., Zhang, F., Dolma, S., Coulombe-Huntington, J., Chatr-aryamontri, A., Dolinski, K., & Tyers, M. (2021). The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Science*, *30*(1), 187–200. https://doi.org/10.1002/pro.3978
55. Salwinski, L., Miller, C. S., Smith, A. J., Pettit, F. K., Bowie, J. U., & Eisenberg, D. (2004). The database of interacting proteins: 2004 update. *Nucleic Acids Research*, *32*(DATABASE ISS.), 449D–451. https://doi.org/10.1093/nar/gkh086
56. Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., Doncheva, N. T., Legeay, M., Fang, T., Bork, P., Jensen, L. J., & von Mering, C. (2021). The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Research*, *49*(Database issue), 10800–10800. https://doi.org/10.1093/nar/gkab835
57. Wassermann, A. M., & Bajorath, J. (2011). BindingDB and ChEMBL: Online compound databases for drug discovery. *Expert Opinion on Drug Discovery*, *6*(7), 683–687. https://doi.org/10.1517/17460441.2011.579100
58. Wang, R., Fang, X., Lu, Y., Yang, C. Y., & Wang, S. (2005). The PDBbind database: Methodologies and updates. *Journal of Medicinal Chemistry*, *48*(12), 4111–4119. https://doi.org/10.1021/jm048957q
59. Nikam, R., Kulandaisamy, A., Harini, K., Sharma, D., & Michael Gromiha, M. (2021). ProThermDB: Thermodynamic database for proteins and mutants revisited after 15 years. *Nucleic Acids Research*, *49*(D1), D420–D424. https://doi.org/10.1093/nar/gkaa1035
60. Siva Shanmugam, N. R., Jino Blessy, J., Veluraja, K., & Michael Gromiha, M. (2020). ProCaff: Protein-carbohydrate complex binding affinity database. *Bioinformatics*, *36*(11), 3615–3617. https://doi.org/10.1093/bioinformatics/btaa141
61. Harini, K., Srivastava, A., Kulandaisamy, A., & Gromiha, M. M. (2022). ProNAB: Database for binding affinities of protein-nucleic acid complexes and their mutants. *Nucleic Acids Research*, *50*(D1), D1528–D1534. https://doi.org/10.1093/nar/gkab848
62. Kastritis, P. L., & Bonvin, A. M. J. J. (2013). On the binding affinity of macromolecular interactions: Daring to ask why proteins interact. *Journal of the Royal Society Interface*, *10*(79), 20120835. https://doi.org/10.1098/rsif.2012.0835
63. Yan, Z., Guo, L., Hu, L., & Wang, J. (2013). Specificity and affinity quantification of protein-protein interactions. *Bioinformatics (Oxford, England)*, *29*(9), 1127–1133. https://doi.org/10.1093/bioinformatics/btt121
64. Wang, D. D., Chan, M. T., & Yan, H. (2021). Structure-based protein–ligand interaction fingerprints for binding affinity prediction. *Computational and Structural Biotechnology Journal*, *19*, 6291–6300. https://doi.org/10.1016/j.csbj.2021.11.018
65. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*, 583–589. https://doi.org/10.1038/s41586-021-03819-2
66. Chang, L., & Perez, A. (2022). Ranking peptide binders by affinity with AlphaFold. *Angewandte Chemie International Edition*, *2022*, e202213362. https://doi.org/10.1002/anie.202213362
67. Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, *55*(10), 78–87. https://dl.acm.org/citation.cfm?id=2347755
68. Dias, R., & Kolaczkowski, B. (2017). Improving the accuracy of high-throughput protein-protein affinity prediction may require better training data. *BMC Bioinformatics*, *18*(suppl 5), 7–18. https://doi.org/10.1186/s12859-017-1533-z
69. Wiśniewski, J. R., Hein, M. Y., Cox, J., & Mann, M. (2014). A "proteomic ruler" for protein copy number and concentration estimation without spike-in standards. *Molecular and Cellular Proteomics*, *13*(12), 3497–3506. https://doi.org/10.1074/mcp.M113.037309
70. Catozzi, S., Halasz, M., & Kiel, C. (2021). Predicted 'wiring landscape' of Ras-effector interactions in 29 human tissues. *Npj Systems Biology and Applications*, *7*(1), 1–15. https://doi.org/10.1038/s41540-021-00170-0
71. Milo, R., Jorgensen, P., Moran, U., Weber, G., & Springer, M. (2009). BioNumbers—The database of key numbers in molecular and cell biology. *Nucleic Acids Research*, *38*(1), D750–D753. https://doi.org/10.1093/nar/gkp889
72. Flamholz, A., Phillips, R., & Milo, R. (2014). The quantified cell. *Molecular Biology of the Cell*, *25*(22), 3497–3500. https://doi.org/10.1091/mbc.E14-09-1347
73. Phillips, R., & Milo, R. (2009). A feeling for the numbers in biology. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(51), 21465–21471. https://doi.org/10.1073/pnas.0907732106
74. Liberali, P., & Pelkmans, L. (2012). Towards quantitative cell biology. *Nature Cell Biology*, *14*(12), 1233. https://doi.org/10.1038/ncb2648
75. Harris, L. A., Hogg, J. S., Tapia, J. J., Sekar, J. A. P., Gupta, S., Korsunsky, I., Arora, A., Barua, D., Sheehan, R. P., & Faeder, J. R. (2016). BioNetGen 2.2: Advances in rule-based modeling. *Bioinformatics*, *32*(21), 3366–3368. https://doi.org/10.1093/bioinformatics/btw469
76. Gogl, G., Jane, P., Caillet-Saguy, C., Kostmann, C., Bich, G., Cousido-Siah, A., Nyitray, L., Vincentelli, R., Wolff, N., Nomine, Y., Sluchanko, N. N., & Trave, G. (2020). Dual specificity PDZ- and 14-3-3-binding motifs: A structural and interactomics study. *Structure*, *28*(7), 747–759.e3. https://doi.org/10.1016/j.str.2020.03.010
77. Jané, P., Gógl, G., Kostmann, C., Bich, G., Girault, V., Caillet-Saguy, C., Eberling, P., Vincentelli, R., Wolff, N., Travé, G., & Nominé, Y. (2020). Interactomic affinity profiling by holdup assay: acetylation and distal residues impact the PDZome-binding specificity of PTEN phosphatase. *PloS One*, *15*(12), 1–24. https://doi.org/10.1371/journal.pone.0244613
78. Gógl, G., Biri-Kovács, B., Durbesson, F., Jane, P., Nomine, Y., Kostmann, C., Bilics, V., Simon, M., Reményi, A., Vincentelli, R., Trave, G., & Nyitray, L. (2019). Rewiring of RSK–PDZ interactome by linear motif phosphorylation. *Journal of Molecular Biology*, *431*(6), 1234–1249. https://doi.org/10.1016/j.jmb.2019.01.038
79. Politi, A. Z., Cai, Y., Walther, N., Hossain, M. J., Koch, B., Wachsmuth, M., & Ellenberg, J. (2018). Quantitative mapping of fluorescently tagged cellular proteins using FCS-calibrated four-dimensional imaging. *Nature Protocols*, *13*(6), 1445–1464. https://doi.org/10.1038/nprot.2018.040
80. Lundberg, E., & Borner, G. H. H. (2019). Spatial proteomics: A powerful discovery tool for cell biology. *Nature Reviews Molecular Cell Biology*, *20*(5), 285–302. https://doi.org/10.1038/s41580-018-0094-y
81. Qin, Y., Huttlin, E. L., Winsnes, C. F., Gosztyla, M. L., Wacheul, L., Kelly, M. R., Blue, S. M., Zheng, F., Chen, M., Schaffer, L. V., Licon, K.,

15211878, 2023, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/bies.202300007 by Cochrane France, Wiley Online Library on [17/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Bäckström, A., Vaites, L. P., Lee, J. J., Ouyang, W., Liu, S. N., Zhang, T., Silva, E., Park, J., … Ideker, T. (2021). A multi-scale map of cell structure fusing protein images and interactions. *Nature*, *600*(7889), 536–542. https://doi.org/10.1038/s41586-021-04115-9

82. Neveu, G., Cassonnet, P., Vidalain, P. O., Rolloy, C., Mendoza, J., Jones, L., Tangy, F., Muller, M., Demeret, C., Tafforeau, L., Lotteau, V., Rabourdin-Combe, C., Travé, G., Dricot, A., Hill, D. E., Vidal, M., Favre, M., & Jacob, Y. (2012). Comparative analysis of virus-host interactomes with a mammalian high-throughput protein complementation assay based on Gaussia princeps luciferase. *Methods*, *58*(4), 349–359. https://doi.org/10.1016/j.ymeth.2012.07.029
83. Buljan, M., Ciuffa, R., van Drogen, A., Vichalkovski, A., Mehnert, M., Rosenberger, G., Lee, S., Varjosalo, M., Pernas, L. E., Spegg, V., Snijder, B., Aebersold, R., & Gstaiger, M. (2020). Kinase interaction network expands functional and disease roles of human kinases. *Molecular Cell*, *79*(3), 504–520.e9. https://doi.org/10.1016/j.molcel.2020.07.001
84. Garai, Á., Zeke, A., Gógl, G., Törő, I., Ferenc, F., Blankenburg, H., Bárkai, T., Varga, J., Alexa, A., Emig, D., Albrecht, M., & Reményi, A. (2012). Specificity of linear motifs that bind to a common mitogen-activated protein kinase docking groove. *Science Signaling*, *5*(245), ra74. https://doi.org/10.1126/scisignal.2003004.Specificity
85. Sahni, N., Yi, S., Taipale, M., Fuxman Bass, J. I., Coulombe-Huntington, J., Yang, F., Peng, J., Weile, J., Karras, G. I., Wang, Y., Kovács, I. A., Kamburov, A., Krykbaeva, I., Lam, M. H., Tucker, G., Khurana, V., Sharma, A., Liu, Y. Y., Yachie, N., … Vidal, M. (2015). Widespread macromolecular interaction perturbations in human genetic disorders. *Cell*, *161*(3), 647–660. https://doi.org/10.1016/j.cell.2015.04.013
86. Davey, N. E., Travé, G., & Gibson, T. J. (2011). How viruses hijack cell regulation. *Trends in Biochemical Sciences*, *36*(3), 159–169. https://doi.org/10.1016/j.tibs.2010.10.002
87. Sámano-Sánchez, H., & Gibson, T. J. (2020). Mimicry of short linear motifs by bacterial pathogens: A drugging opportunity. *Trends in Biochemical Sciences*, *45*(6), 526–544. https://doi.org/10.1016/j.tibs.2020.03.003
88. Paget-Bailly, P., Meznad, K., Bruyère, D., Perrard, J., Herfs, M., Jung, A. C., Mougin, C., Prétet, J. L., & Baguet, A. (2019). Comparative RNA sequencing reveals that HPV16 E6 abrogates the effect of E6*I on ROS metabolism. *Scientific Reports*, *9*(1), 1–15. https://doi.org/10.1038/s41598-019-42393-6
89. Hofmann, J. L., Maheshwari, A. J., Sunol, A. M., Endy, D., & Zia, R. N. (2022). Ultra-weak protein-protein interactions can modulate proteome-wide searching and binding. *BioRxiv*, 1-36. https://www.biorxiv.org/content/10.1101/2022.09.30.510365v1
90. Krystkowiak, I., Manguy, J., & Davey, N. E. (2018). PSSMSearch: A server for modeling, visualization, proteome-wide discovery and annotation of protein motif specificity determinants. *Nucleic Acids Research*, *46*(W1), W235–W241. https://doi.org/10.1093/nar/gky426
91. Schneider, T. D., & Stephens, R. M. (1990). Sequence logos: A new way to display consensus sequences. *Nucleic Acids Research*, *18*(20), 6097–6100. https://doi.org/10.1093/nar/18.20.6097
92. Crooks, G. E., Hon, G., Chandonia, J. M., & Brenner, S. E. (2004). WebLogo: A sequence logo generator. *Genome Research*, *14*(6), 1188–1190. https://doi.org/10.1101/gr.849004
93. Lambourne, L., Yadav, A., Wang, Y., Desbuleux, A., Kim, D.-K., Cafarelli, T., Pons, C., Kovács, I. A., Jailkhani, N., Schlabach, S., Ridder, D. D., Luck, K., Bian, W., Shen, Y., Yang, Z., Mee, M. W., Helmy, M., Jacob, Y., Lemmens, I., … Vidal, M. (2022). Binary interactome models of inner- versus outer-complexome organisation. *BioRxiv*,1-89. http://biorxiv.org/content/early/2022/07/22/2021.03.16.435663.abstract

How to cite this article: Weimer, K., Zambo, B., & Gogl, G. (2023). Molecules interact. But how strong and how much? *BioEssays*, *45*, e2300007. https://doi.org/10.1002/bies.202300007

15211878, 2023, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/bies.202300007 by Cochrane France, Wiley Online Library on [17/11/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Methods in quantitative interactomics

While the concept of quantitative interactomics is not novel, historically, approaches within this domain implore the use of low throughput methods. In pursuit of large-scale studies, interactomics turned away from quantitative methods in favor of scalable ones that sacrificed molecular details for higher throughput. Even so, the desire to obtain quantitative details from interactomics is prevalent in current practices across the field[10,45-47]. However, one could argue that many of the methods currently being implored for quantitative study do not objectively measure the interactome. Many methodologies focus on the quantification of some signal, such as luminescence or fluorescence, as an indicator of the presence of an interaction. As previously mentioned, without stringent parameters to classify what qualifies as an interaction this measurement does not provide much detail beyond the signal strength which does not reflect any inherent properties about the interaction other than its presence. Other methods focus on measuring extrinsic properties of the interactome, such as quantified proteomics or quantification of complex formation and stoichiometries. While these are certainly interesting characteristics, they represent functions of the interactome at work and are variable. As neither of these approaches involve measuring inherent properties of the interactome we regard them in this context as “pseudo-quantitative”. That said, to date, several methods for quantitatively measuring protein affinities, an intrinsic property of interactomes, have been developed. While many of these are still limited to low (SPR, BLI, ITC) and medium throughput (FP and MST) methods and therefore unable to accommodate the large-scale demands of the field, high throughput approaches are available.

Chromatographic retention assays

A family of assays, known as the holdup (HU) assays, have been established for high-throughput protein affinity measurements (Fig. T7)[48-52]. These methods are all based on the principle of chromatographic retention[53]. The premise is similar to that of a pull-down in that it relies on the immobilization of an affinity tagged bait on an affinity matrix. However, unlike the pull-down, following exposure of the bait to preys there are no wash steps allowing for the direct measurement of the remaining prey. This is then compared to a non-binding control where the degree of depletion of certain preys after exposure to the bait can be calculated. Baits are supplied in great excess of the preys in solution and all protein concentrations are fixed. This allows for the accurate calculation of protein affinities (K_d), even for weak interactions, and because interactions are measured at equilibrium, transient interactions are also detected. The methods are highly adaptable and automatable, with assay readout accomplished through intrinsic fluorescence, which can be read via plate reader, or MS and the implementation of automated pipetting for large scale panels.

Initial applications of holdup assays were used to measure protein fragment interactions, or fragmentomics. In this format combinations of domain, peptide, and full-length interactions can be measured. Quantifying the estimated few hundreds of

millions of interactions that comprise the full fragmentomic interactome represents a more practically feasible feat than measuring the full-length human interactome which is likely represented by upwards of five hundred million interactions[50]. Furthermore, the use of HU protein-peptide panels is useful in the evaluation of therapeutics. Nevertheless, fragmentomic approaches provide a limited view of the binding events and thus necessitate the development of an approach capable of measuring the affinities of full-length interactions. To this end, adaptations of the HU Multiplex that involve coupling to methods in high throughput protein expression and purification in conjunction with automated pipetting are in development[54]. Additionally, the native holdup (nHU) has emerged as a way to measure the affinities of protein interactions in their pseudo-native state directly from whole cell lysates [51].

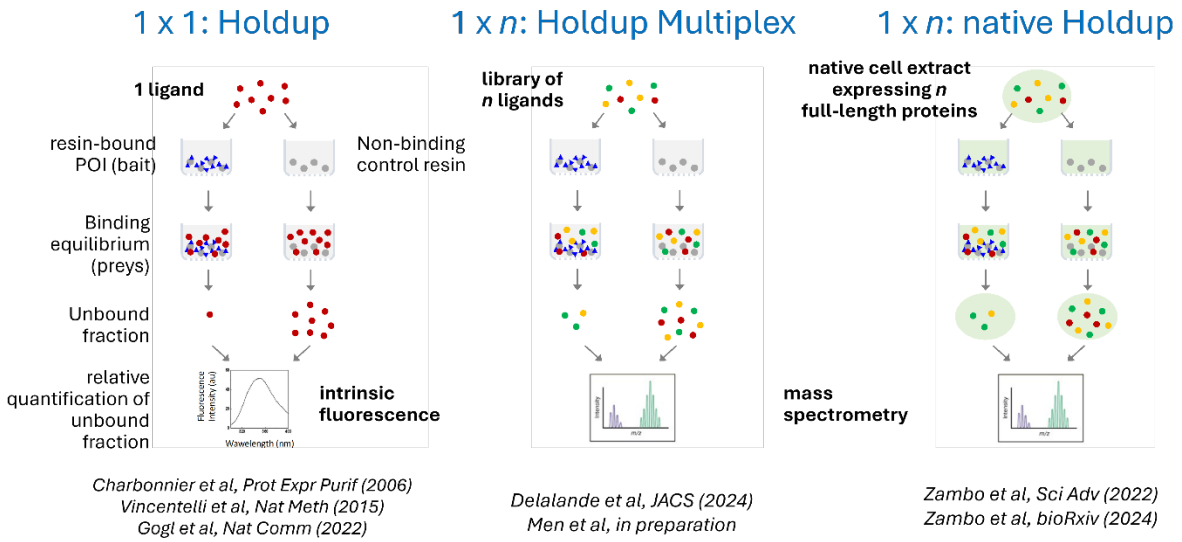


Figure T7 Holdup assays for quantitative interactomics. All Holdup (HU) family assays are based on the same principle of chromatographic retention. This includes the same steps of immobilizing bait on an affinity matrix at a known concentration, exposure to a solution containing preys, separation of bound and unbound fractions, and analysis and quantification of the depletion of preys from the unbound fractions. From left to right: (1) the traditional HU assay allows evaluation of binary interactions between singular baits and singular preys. (2) the HU Multiplex allows the assessment of a singular bait against complex mixes containing multiple preys. (3) the native HU measures interactions of a singular bait against cognate preys present in cell extracts. MS analysis can then be used to identify interactors. (Figure provided by Elodie Monsellier with some modifications)

Interactomics and disease:

We've progressed in strides from original postulations based on hypotheses such as Beadle and Tatum's notion of "one-gene-one-protein-one-function"[55] which was over simplified and vastly reductionist compared to the highly dynamic and complex architecture that we now understand underpins this relationship[5]. Leading into this paradigm shift was the surprising report by the International Human Genome Sequencing Consortium that the genome is composed of significantly less protein coding genes than what was originally predicted[56]. This introduced the idea of a certain functional flexibility and efficiency that allows the genome to dictate beyond what it directly encodes. As more researchers adopted this perspective[57-59] in parallel with the computational developments of recent decades it has become exceedingly apparent that the genotype-phenotype link spans several layers of molecular organization mounting in complexity and dynamics (Fig. T8). This has necessitated the use of systems biology approaches which seek to investigate the molecular relationships underscoring cellular processes within the context of integrated systems therefore producing a more complete view of phenotypes and pathologies[60].

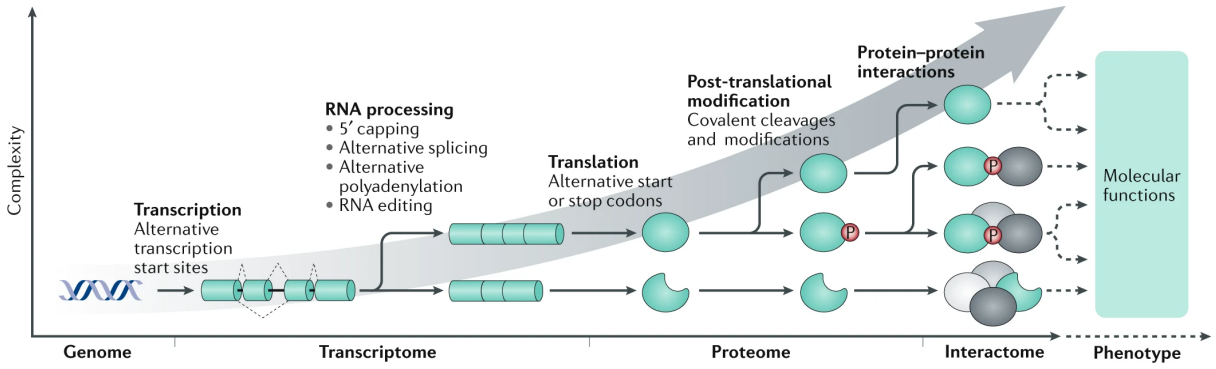


Figure T8 The generation of functional diversity at different molecular levels. Cellular complexity arises from many mechanisms that expand molecular diversity beyond that encoded by the protein-coding genome. These mechanisms include an increase in coding potential through various methods of transcriptional processing such as using alternative transcription start sites as well as 5' capping, alternative splicing, alternative polyadenylation, and RNA editing at the co-transcriptional of post-transcriptional level. Additionally, by using alternative start and stop codons during translation and introducing post-translational modifications a high degree of protein diversity is generated. These proteins can create interactions with one another to form multiple distinct functional units that can potentially perform a variety of downstream functions. The functional impact of these different layers of molecular organization on phenotypes (dashed lines) remains to be explored. *(Taken from Bladau & Aebersold, 2020)*

Sitting directly proximal to the phenotypic layer of molecular organization, the interactome represents a critical junction for understanding disease as any perturbances within the interactome will be directly reflected in the phenotype. Furthermore, studies investigating the role of genetic mutations within the context of the interactome have reported that mutations which drastically alter interactomes are often associated with genes and various mutations within the same gene can produce distinctive interactomic profiles which are associated with distinguishable phenotypes[61-64]. With this in mind, interactomic approaches have become ideal for furthering our understanding of disease. In particular, interactomics is advantageous for investigating infectious disease, adeptly addressing pathogenic hi-jacking in which a foreign entity infiltrates the host network and rewires it to suit its own survival[65-67]. This methodology has been applied across various pathogens to further our understanding of how viruses[68-70], bacterium[71-73], fungi[74,75], and parasites[76-78] illicit infection through this restructuring of host networks.

HPV as a model for interactomics

A longstanding model in host-pathogen interactomic studies is Human Papillomaviruses (HPVs) due to their high genotypic and phenotypic diversity[79,80]. Clinical presentation of HPV infection can range from transient and subclinical to persistent and cancer causing[81] and the nearly 200 identified HPV isotypes[82] have been subclassed based on this distinction of oncogenic potential[83,84]. This has resulted in the identification of 16 high-risk HPVs (hr-HPVs)[85]. At the crux of HPV oncogenesis is a triad of virally encoded oncoproteins: E5, E6, and E7[86]. This further validates HPV's value as a model system for interactomics, as any one of these oncoproteins represents a singular component capable of disrupting host interaction networks to drive pathogenesis.

To further explore the central role of these oncoproteins, particularly E6 and how its complex entanglement with host cellular networks underpins HPV-related pathogenesis I have provided in the subsequent pages a review I authored and submitted for publication:

Kathleen Weimer (2024) Too many cooks in the kitchen: HPV driven carcinogenesis – the result of collaboration or competition?. Tumour Virus Research. (During the time of defense this piece was *in review* and published shortly after on 27 December 2024. The original submitted draft has been replaced by this final published version below.)

Tumour Virus Research 19 (2025) 200311

Contents lists available at ScienceDirect

Tumour Virus Research

journal homepage: www.journals.elsevier.com/tumour-virus-research

Too many cooks in the kitchen: HPV driven carcinogenesis – The result of collaboration or competition?

Weimer Kathleen

IGBMC – CBI: Institut de génétique et de biologie moléculaire et cellulaire, Centre de biologie intégrative, 1 rue Laurent Fries, Illkirch-Graffenstaden, BP 10142, 67404, France

ARTICLE INFO

Keywords:
HPV
E6
Oncoproteins
Carcinogenesis
Papillomavirus
Oncogenesis

ABSTRACT

Infection by Human Papillomaviruses accounts for the most widespread sexually transmitted infection worldwide. Clinical presentation of these infections can range from subclinical and asymptomatic to anogenital cancers, with the latter associated with persistent infection over a significant period of time. Of the over 200 isotypes of the human virus identified, a subset of these has been characterized as high-risk due to their ability to induce oncogenesis. At the core of Papillomavirus pathogenesis sits three virally encoded oncoproteins: E5, E6, and E7. In this review we will discuss the respective roles of these proteins and how they contribute to carcinogenesis, evaluating key distinguishing features that separate them from their low-risk counterparts. Furthermore, we will consider the complex relationship between this trio and how their interwoven functional networks underpin the development of cancer.

1. An introduction to HPV life cycle and pathologies

Research efforts to understand Human Papillomaviruses (HPVs) often focus on high-risk HPVs (hr-HPVs), particularly the prototypical HPV16, due to their role in cervical and other cancers [1,2]. These investigations have uncovered a universal pattern of viral gene expression that can be extended, with some modification, across different HPV groups [3]. Infection commences when the virus enters cells of the basal layer, presumably gaining access to the epithelium through microabrasions [4] or, in the case of cervical infection, by traversing the squamocolumnar junction between the endo- and ectocervix [5]. Upon access, an effective infection is established in dividing basal epithelial cells (Fig. 1) [6,7]. This first stage involves viral DNA replication and genome maintenance [8]. During this step, the viral genome, in complex with L1 and L2, is trafficked to the nucleus [9,10] where early replication is initiated, producing between 50 and 100 episomal copies per nucleus [11]. In parallel, a period of cell proliferation is provoked, creating a layer of basal cells harboring replicated episomes [4]. Viral protein expression remains low during this phase due to E2-mediated repression of the early promoter in efforts to subvert immune detection [12]. This makes E1 and E2 the primary players, as they sequester host machinery for DNA replication [13,14]. Following ongoing division of infected cells, some progeny remain in the basal layer, acting as an episomal reserve, while others ascend, migrating through the epithelial strata, towards the productive phase of the viral life cycle [15]. It is here, in the suprabasal layer, that viral oncoproteins E5, E6, and E7 assume more prominent roles, postponing terminal differentiation and blocking departure from the cell cycle, thereby stabilizing an environment that promotes genome amplification at high copy number [16–18]. At the summit of the epithelium, after terminal differentiation and the expression of late genes L1, L2, and E1^4, the HPV life cycle concludes with virion assembly in the nucleus followed by subsequent release from the epithelial surface [19].

Often HPV infections are asymptomatic, cleared by the host immune system within a year or two devoid of any indication of infection [20]. In some cases, persistent infection may occur, and if these infections are not readily resolved by the immune system, they may result in cancer progression (Fig. 1) [21]. In this progression, cervical cancer is preceded by cervical disease, or precancerous lesions, a common occurrence that is the consequence of HPV infection [8,22]. These precancerous lesions, or "cervical intraepithelial neoplasia" (CIN), are scored into three grades: CINI, CINII, and CINIII [23,24]. In the CINI stage, HPV infection is still considered transient and it is not uncommon to detect infections by multiple HPV types [25]. The prognosis for cervical cancer development from CINI is low, with HPV infections often resolving. However, progression from CINI to CINIII is considered indicative of underlying cancer formation [26]. Of HPV-infected women, approximately 10 % show signs of one of these stages, representing oncogenic transformation

E-mail address: weimerk@igbmc.fr.

https://doi.org/10.1016/j.tvr.2024.200311
Received 11 November 2024; Received in revised form 21 December 2024; Accepted 22 December 2024
Available online 27 December 2024
2666-6790/© 2024 The Author. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

of the cervix [27]. However, it is important to note that hr-HPV induced oncogenesis is an uncommon occurrence [2,8]. Oncogenic transformation is predominantly confined to the site of the squamocolumnar junction within a small region of metaplastic squamous epithelium, deemed the "transformation zone", where cells appear to be especially susceptible to hr-HPVs [4,5,28,29]. Further still, carcinogenesis is not an ideal outcome for the virus as transformed cells are differentially defective, ultimately terminating viral replication, which is tightly tethered to this process [6,30,31]. With this in mind remission from CIN states is frequent, occurring spontaneously in approximately 80–90 % of CINI cases and up to 60 % of CINIII cases [26].

In the event that tumorigenesis does occur, it is instigated by the increased expression of virally encoded oncoproteins, principally E6 and E7. Activities of these proteins involve targeting host factors to induce cell growth, chromosomal instability, and decreased differentiation. Upregulation of these proteins occurs in the wake of HPV genome integration into that of the host [32]. Many molecular models explaining integration have been proposed, but it remains a complex process, further compounded by host co-factors, making it difficult to reach a clear consensus [27]. Several "hot spots" within the host genome have been suggested as preferential sites for HPV integration [33–35], but ambiguity remains regarding when integration is initiated and its contribution to carcinogenesis, with evidence of integration found in all grades of CIN [36–39]. However, it is understood that integration events diverge from the normal HPV replication cycle, leading to an abortive infection where the production of virions arrests despite ongoing synthesis of viral proteins [15,40]. In the process of integration, expression of E2 is lost, alleviating repression of E6 and E7 [8,27]. This has catastrophic consequences for the cell, as the aberrant activity of these proteins exceeds their attempts to evade immune detection and sustain the viral lifecycle, instead driving immortalization and malignancy.

2. HPV oncoproteins – E5, E6, and E7

At the heart of HPV pathogenesis is a triad of oncoproteins: E5, E6, and E7 [41]. While these proteins are known for onsetting oncogenesis, their primary roles are necessary for the HPV life cycle [8]. These proteins play critical roles in regulating the host immune response, circumventing detection by disrupting gene expression, hijacking cellular networks via protein-protein interactions, inducing posttranslational modifications, and facilitating the cellular trafficking of key host immune modulators [42,43]. Studies monitoring E6 and E7 mRNA in middle and lower layers of the epithelium exhibited elevated levels [44–46] compared to biomarker studies measuring oncoprotein expression in the upper stratum which demonstrated diminished levels [47]. Taken together, considering the early roles of these proteins, such as the requirement of E6 for episomal genome maintenance [16,48,49] and E7 activation of the G1/S checkpoint supporting viral replication [50], it indicates that the activity of these proteins is likely most important during early infection. Despite this, their dysregulation is viewed as deterministic of cancer progression as their efforts to drive the viral life cycle coincidently prompts a persistent pro-proliferative cellular state [51]. While continuous oncoprotein expression is necessary to uphold the transformed phenotype, it alone is not sufficient for transformation, and given that the majority of infected cells do not progress to cancer, this reaffirms that other factors must influence carcinogenesis [52]. Upregulation of oncoprotein expression appears to follow integration and can confer a selective growth advantage to cells [53,54]. However, as it is not the predominant phenotype, this appears to be a fate stimulated by interactions with the host as opposed to the virus itself. Here we will review each oncoprotein, evaluating key interactions with host factors, with an emphasis on E6, to highlight how this interplay can lead to the establishment of cancer.

2.1. E5—

The three early genes—E5, E6, and E7 have been classified as oncoproteins due to their functions in driving carcinogenesis [55]. However, not all HPVs encode E5 with types from the Beta, Gamma, and Mu genera lacking the E5 ORF entirely [56]. As such, primary transformative and oncogenic functions have largely been attributed to E6 and E7 with E5 reduced to a supporting role. However, after demonstrating the protein's ability to induce anchorage-independent growth in murine fibroblasts and keratinocytes [57–59] E5 earned its class as a *bona fide* oncoprotein.

More recently, evolutionary studies of genital HPVs have given rise to the hypothesis that carcinogenic HPVs all stem from a shared lineage and encode E5 [60–62]. Further classification of the E5 ORF into four groups—alpha, beta, gamma, and delta—revealed a clustering based on clinical manifestations and identified a subset of 10 E5 encoding

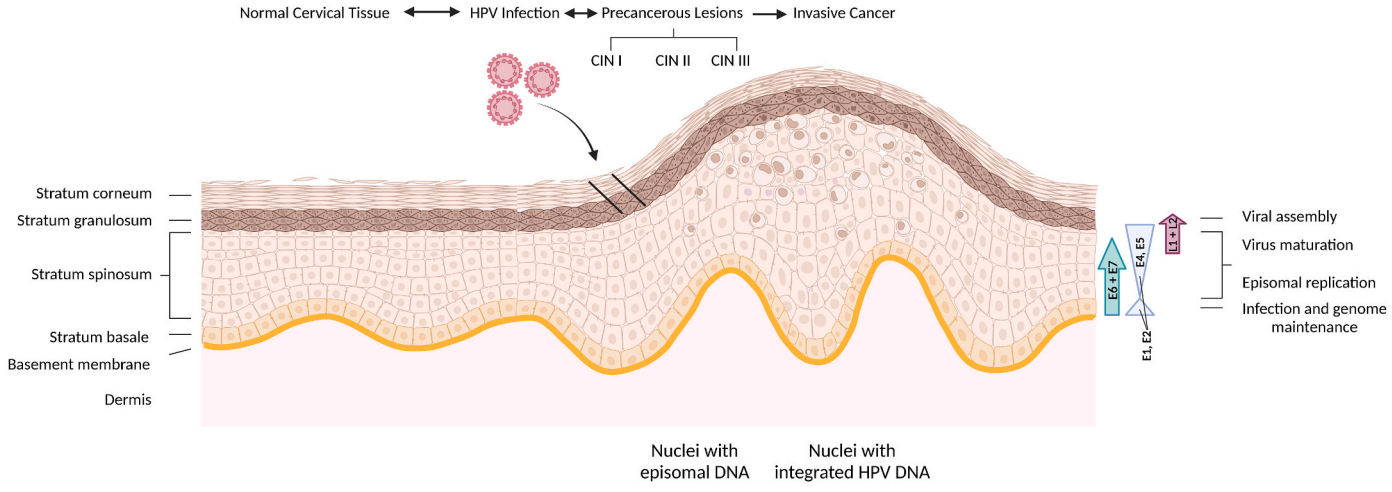


Fig. 1. The progression of HPV induced cancer in epithelial tissue.
Schematic representation of epithelial tissue reorganization in the development of HPV related cancers. On the left epithelial layers are defined. HPV enters through microabrasions or the junction between the endo- and ecto-cervix establishing low copy number episomal replication in basal cells. As infected cells climb through the stratum of the epithelial tissue, changes in viral protein levels can be viewed on the right side with their corresponding stages of the viral life cycle. Following infection precancerous lesions, graded CINI-III, may progress to invasive cancer as the result of persistent infection over the course of up to 20 years. This transition is marked by a decrease in episomal DNA and an increase of cells containing integrated viral DNA. Created in BioRender. Weimer, K. (2024) https://BioRender.com/z51i407.

low-risk (lr-HPVs) which elicit benign venereal warts [63]. Of these, E5β HPVs are the only group associated with cutaneous lesions. Additionally, analysis of differentially active 16E5 variants in vitro drew an association between variants displaying the greatest mitogenic activity and those most commonly detected in the population and in cervical lesions [64]. Taken together, this suggests that while E5 is not essential for the viral life cycle, it does confer certain advantages that favor infection and transformation [65]. Even so, the exact functions of E5 remain elusive, and its contributions have been limited to the early stages of tumorigenesis [51,65–67], leaving the protein primarily credited with augmenting the transformative effects of E7 and working in conjunction with E6 and E7 to drive malignancy [68–73]. Efforts to determine the role of E5 have been complicated by the challenging biochemical nature of the protein, making it difficult to produce and study in the lab, along with a lack of specific antibodies that impedes the study of HPV-infected cells expressing endogenous levels of E5.

The HPV16 E5 protein is the most well-characterized of the E5 proteins, forming an 83 amino-acid polypeptide [74] with other E5 proteins varying in size from 40 to 85 amino acids [75]. These proteins are abundant in hydrophobic amino acids gathered into a multi-pass transmembrane protein with a cytosolic C-terminus [76–78]. Given these characteristics, it has been suggested that E5 belongs to a family of proteins, known as viroporins [79]— a group of viral membrane proteins characterized by the presence of at least one amphipathic helix as part of a hydrophobic domain that undergoes membrane insertion [80]. As viroporins homo-oligomerize, the amphipathic helices arrange themselves as a membrane-spanning hydrophilic central pore with hydrophobic residues facing outwards, interacting with the lipid bilayer [81]. This assembly acts as a channel, facilitating the passage of small molecules and ions thus allowing them to regulate ion homeostasis. While further study is needed to determine the ion selectivity of the E5 channel [82], it has been demonstrated that E5 oligomerizes in cells and forms a functional pore structure in liposomes [80]. Comparative analysis of lr- and hr-HPV E5 sequences denotes a difference in the conservation of two amino acids located at the end of the sequence [83]. This region corresponds to a predicted transmembrane helix where hr-HPVs retain histidine and alanine in place of the tyrosine and isoleucine found in lr-HPVs. While small hydrophobic residues have high packing values and are common within transmembrane helices, charged residues are less frequent [84]. An investigation of other viroporins evaluated the structural impact of a HxxxW motif. While terminal tryptophans are often described as anchoring residues for integral membrane structures [85], this study established that the H residue creates a change in tilt angle of the helix [86]. This change in tilt not only accommodates the larger size of the residue, but also increases solvent accessibility which may stimulate changes in ion selectivity [87]. Furthermore, since this tilt shift occurs even in the absence of a tryptophan it is likely histidine has a similar impact on hr-HPV E5s. Finally meta-analysis of trends in the amino acid composition of transmembrane helices suggests that increased sequence complexity is correlated with functionality. As such, transmembrane helices primarily composed of simple hydrophobic amino acids will mainly act as anchors compared to those with more diversified sequences which are associated with functionality [88]. Taken together, this suggests structural and functional differences in the E5 proteins of high and low risk types, however, further study is necessary to solidify this theory.

Furthermore, overexpression studies show that E5 localizes primarily to the Endoplasmic Reticulum (ER) and Golgi Apparatus (GA) [89,90], potentially associating with the plasma membrane as well, hinting towards a role in the trafficking of cytoplasmic membrane proteins [67,91]. In addition, a myriad of other functions has been described for E5 [63], including but not limited to, stimulation of the epidermal growth factor receptor (EGFR) signaling cascade [59,92], altering apoptotic response [93], inducing changes in membrane lipid composition [94,95], and blocking trafficking of major histocompatibility complexes (MHC) classes -I and -II [91,96]. Due to our limited understanding of E5 function, it is difficult to pinpoint exactly where it intersects with its counterparts, E6 and E7. That said, the evidence that E5 augments effects of E6 and E7 likely indicates a regulatory network underpinning their functionality.

2.2. E6—

The discovery that, in cervical tumors and cervical cancer-derived cell lines, the E6 ORF was retained and expressed even years after initial transformation led to speculation about its role as an oncoprotein [97–99]. Its intrinsic ability to transform cells was verified in several models [100–105], however not all E6 proteins possess these capacities. While E6s of hr-HPVs have proven sufficient for transformation, lr-HPVs are not capable of transforming primary cells [106]. Additionally, in the case of hr-HPVs, E6 is considered to have weakly transformative capabilities, working in tandem with E7 to establish carcinogenesis. This has given rise to the paradigm that E7 operates as the founder of tumorigenesis while E6 perpetuates malignancy [107]. Supporting this theory is the evidence that immortalization of primary human keratinocytes requires the full viral genome intact, containing both E6 and E7 ORFs [108].

E6 proteins are small proteins approximately 150 amino acids in length [109], and fold into two domains, E6N and E6C, named according to their distal location at the N- or C-terminus respectively. Each domain assumes a structure consisting of a triple-stranded β-sheet and two short helices with another short helix bridging the two regions [110,111]. These proteins possess a total of four CXXC zinc binding motifs arranged into two pairs, forming two zinc finger domains that are essential for E6 activity [112–115]. At its core, E6 forms a charged, hydrophobic binding groove aptly termed the “LxxLL binding pocket” due to its ability to recognize and bind a short linear LxxLL motif, normally found in the disordered regions of proteins within acidic peptides that arrange as a helix upon contact [116–121]. Both lr- and hr-HPVs adopt this core fold, allowing them to attract LxxLL partners, but variations in binding profiles have been described, leading to their subcategorization. This dichotomy divides E6 proteins that utilize the LxxLL pocket to associate with the cellular E3 ubiquitin ligase, E6-Associated Protein (E6AP), a characteristic shared by Alpha, DyoDelta, Dyopi, Omega, and Omikron Papillomaviruses, from ones that interact with the Notch co-activator Mastermind Like 1 (MAML-1),as seen in the remaining genera [122–126]. This disrupts the traditional convention of using tropism to cluster HPV types as Alpha and Beta genera contain types responsible for both cutaneous and mucosal infections. Sequence and structural analysis of E6 contact residues within the LxxLL pocket revealed a lack of conservation, suggesting some E6 proteins may target multiple LxxLL partners, but it does not clarify the selectivity between E6AP and MAML-1. Additionally, functional assays demonstrated that E6 binding to E6AP alone is not sufficient to induce its degradation [122,127]. As such, there are low-risk α-HPVs which associate with E6AP but do not degrade it. Conversely, some β-HPVs have been linked to nonmelanoma skin cancers as E6 represses Notch trans-activation through its interaction with MAML-1 [126]. Given that the Notch signaling pathway is a critical determinant in keratinocyte differentiation and cell cycle arrest, it represents a key target for the HPV life cycle. Taken together, this demonstrates the importance of the nature of E6 interactions with their host factors in addition to the targets themselves. Finally, at the C-terminal of hr-HPVs a PDZ (PSD-95/DLG/ZO-1) binding motif (PBM) can be found [128]. This motif is considered a class I PBM formed by a X-S/T-X-V/L consensus sequence [129] which recognizes and binds PDZ containing partners [130]. Many PDZ proteins contain multiple copies of the domain in addition to other protein-protein interaction motifs, allowing them to serve as interaction hubs, erecting scaffolding for the assembly of multi-protein complexes involved in an array of cellular functions [131]. This structural distinction between lr-HPVs and hr-HPVs also distinguishes the group functionally. When trying to understand differences between viruses of the two groups, genome analysis

has demonstrated a high degree of conservation between virus types. This makes the PBM of hr-HPVs a defining feature which can be used as a molecular signature to indicate oncogenic potential [132–134].

Utilizing its various protein interfaces E6 is able to target an array of host factors involved in diverse cellular processes (Fig. 2) and disrupt cellular function despite total lack of any inherent enzymatic activity [15]. The profound effect of E6 on cellular function is owed, in part, to its association with E6AP, which can lead to the recruitment and degradation of targets such as p53 [135,136] and some PDZ proteins [137–139]. Meanwhile, other studies have implicated the involvement of a ubiquitin ligase other than E6AP in the degradation of some PDZs [140,141]. Successful sequestration of p53 occurs after E6 undergoes conformational changes induced by E6AP binding, thus revealing an additional interface for direct p53 interaction [142–144]. Based on this model of tertiary complex formation, brought through binding-induced conformational changes, and our own E6 binding studies (data not yet published), E6 may interact through other intermediaries in a similar manner, leading to other unexplored interfaces. This mechanism could also underpin the promiscuous binding profiles exhibited by many E6 proteins as well as their diversity despite structural conservation. To this end, many interatomic studies [145–154] have sought to characterize E6 interactions with aims of understanding their contributions to cancer development [155]. Comparison of E6 interactions from lr- and hr-HPVs revealed overlap in cellular targets, indicating some advantage in supporting the viral life cycle [156]. This overlap is also shared by oncoproteins of the same HPV type, suggesting that they are critical targets and that the accumulation of cellular effects from multiple successful viral attacks likely contributes to cancer development [157,158].

2.3. *Other forms of E6* –

2.3.1. *Phospho-E6*

We've discussed that the E6 C-terminal PBM acts as a critical determinant in distinguishing hr- and lr-HPVs, however, this region is not restricted to its interactions with PDZ partners, displaying dual functionality as a phospho-protein. Within the PBM, situated at T156, is a phospho-acceptor site that, upon modification, transitions E6 PBM targeting from interactions with PDZs in favor of associations with members of the 14-3-3 protein family [159,160]. Early studies asserted E6 phosphorylation was facilitated by either protein kinase A (PKA) or AKT, recognizing the phospho-site in a sequence dependent manner [161,162]. Further exploration revealed a correlation between DNA damage and increased levels of phospho-E6, linking the protein to kinases involved in regulating the cellular stress response, such as DNA PK [163].

Much remains to be understood regarding the role of the 14-3-3 protein family in HPV related pathogenesis. There are seven human isoforms of 14-3-3 proteins and to date, phospho-E6 has been shown to interact directly with 14-3-3ζ which stabilizes levels of E6 [161,164,165]. In general, this family is implicated in nearly every facet of cellular function and they are particularly known for their involvement in cellular signaling, meaning their potential contributions to cancer development are endless [166]. Enrichment of expressed 14-3-3ζ was identified in several cervical cancer cell lines via comparative proteomic analysis with a non-tumorigenic cell line [167]. Furthermore, this same study suggests that 14-3-3ζ causes dysregulation of the cell cycle, driving malignant transformation and determining cell fate. Other

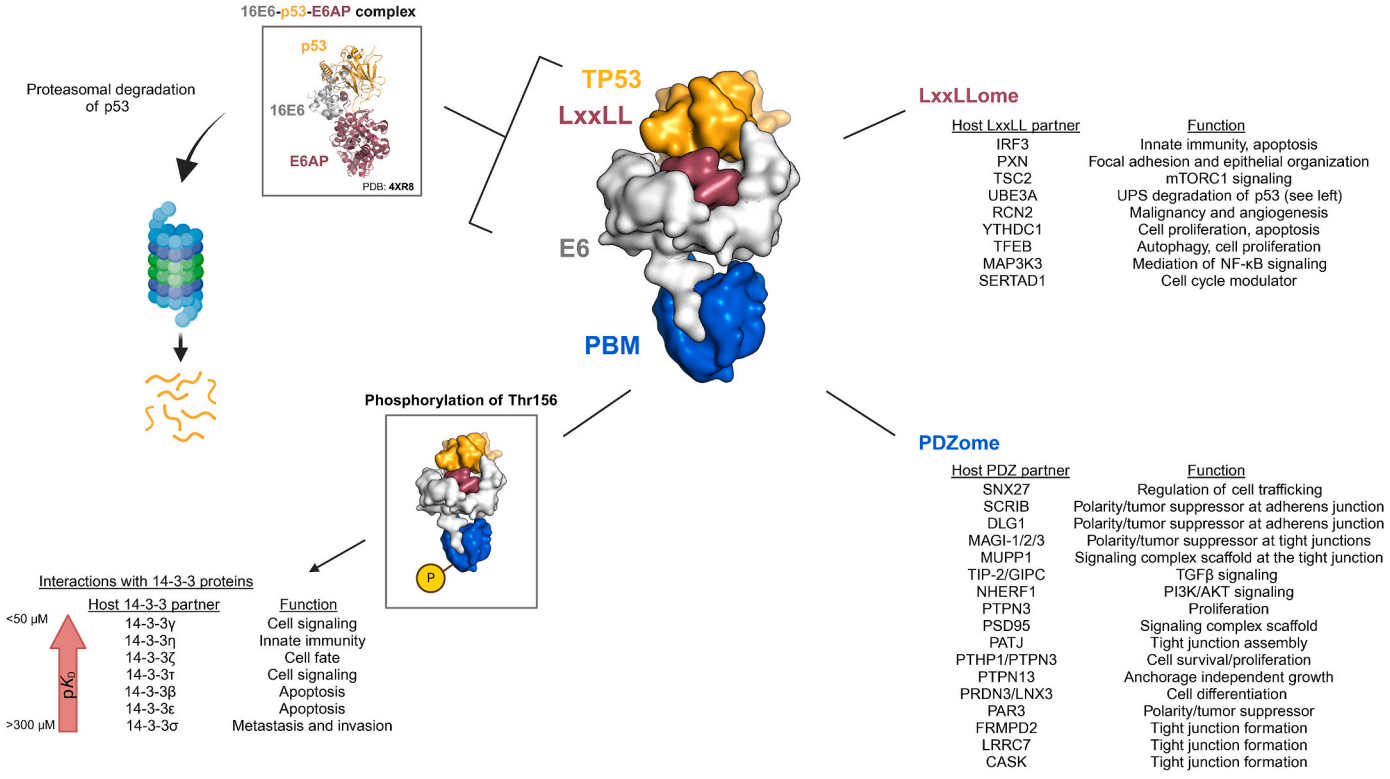


Fig. 2. Interactions of hr-HPV E6.
The canonical hr-HPV E6 structure is shown with interfaces labelled as follows: LxxLL binding pocket is highlighted in raspberry, the TP53 (p53) binding interface is highlighted in gold, the C-terminal PBM is highlighted in blue, and the remaining E6 surface is shown in grey. In the upper left-hand box, the structure of the tertiary 16E6-p53-E6AP complex is displayed (PDB: 4XR8). The proteasomal degradation of p53 is mediated by this complex. In the lower left-hand box, a diagram of E6 modified by phosphorylation of T156 is shown. This introduces interactions with members of the 14-3-3 family with their potential functions in cancer listed. An affinity scale is represented by a red arrow, indicating the strength of interactions which range from 300 μM–50 μM based on fragmentomic in vitro binding assays although only E6 binding to 14-3-3ζ has been functionally implicated *in vivo*. To the right interactions with LxxLL partners (top right) and PDZ partners (bottom right) are listed and referred to the LxxLLome and PDZome respectively.
Created in BioRender. Weimer, K. (2024) https://BioRender.com/m12c626.

models have attempted to explain 14-3-3 function in the context of HPV infection by determining the role of E6 phospho-regulation. In HeLa cells, loss of E6AP increased E6 phosphorylation making it the dominant species in these cells where it is also known that E6 levels drop due to its destabilization. The upregulation of modified E6 is dependent upon transcriptionally active p53 and DNA PK [168]. Although, previous studies showed a connection between the presence of an E6 PBM with a phospho-acceptor site and transcriptional inhibition of p53 by E6. Taken together, these results imply that these proteins are part of a negative feedback loop, where p53 activation prompts phosphorylation of E6, in turn enhancing E6's ability to transcriptionally regulate p53 [163,168]. While the full extent of phospho-E6's roles remain to be established, especially in the context of the other 14-3-3 family members, this is another example of the complex connection erected by E6 and its related proteoforms.

2.3.2. E6*

Many HPV genes are expressed as polycistronic pre-mRNAs that produce various transcripts through alternative splicing, generating differential patterns in mRNA expression throughout infection [169,170]. Processing of the E6-E7 ORF by alternative splicing is a common feature of hr-HPVs, not found in lr-HPVs [171], resulting in the production of multiple transcripts, containing truncated forms of E6. Of these forms, all hr-HPVs encode the E6*I, normally referred to as E6*, splice variant [169]. Originally, it was hypothesized that the sole role of E6* was to promote E7 translation. The gap between E6 termination and E7 initiation on the E6/E7 mRNA is not adequate to support efficient translation of E7; however, this is not an issue with the spliced transcripts [102,172,173]. Other studies, however, have shown that E7 is predominantly translated from non-spliced mRNA [174,175]. While existence of the splice variant at the protein level has been a point of contention, there have been several reports suggesting biological activity as a protein [176–178] with detection of protein expression in cervical cancer cell lines [179] and binding assays with synthesized E6* peptides (data not yet published). Although, it is worth noting the E6* protein has yet to be detected in infected cells *in vivo* [180].

E6* shares the first 44 amino acids with full-length E6 and possesses an additional 13 residues gained through the intron removal process [177]. This corresponds to the conservation of only a portion of the E6N domain. Additionally, most proteins contain a hydrophobic (L/M/I)XX(L/I/V)X(L/V/I) motif which is implicated in E6*'s association with E6-E6AP binding [176]. Other than these features, not much is known regarding the structure or biophysical properties of E6*. Much like its full-length counterpart, E6* exhibits a multitude of functions and, similar to the interplay between other oncoproteins, there is a dynamic relationship between E6 and E6* that determines the protein's functional capacities. Investigations have implicated E6* independently in the involvement of many activities including p53 and WNT/β-catenin signaling, apoptosis, oxidative stress and DNA damage, degradation of some PDZ substrates, and the inflammatory response [176,177,179,181–184]. In an interesting turn of events, some of these attributes are anti-tumor functions while others have contradictory roles in tumorigenesis. It's been proposed that the function of E6*, at least in part, is to counteract the oncogenic effects of E6. Depending on the HPV type, the roles of E6* differ, a trend observed for all HPV oncoproteins, reflecting the intricate functional networks they weave.

2.4. E7—

Upon understanding that the genome of hr-HPVs contain transformative properties [185], the HPV protein E7 was quickly implicated in this process [186–191]. Subsequent studies indicated a complementation between E7 and E6 activity leading to the conclusion that both parties were necessary for transformation [108,192,193]. Current understanding, however, suggests that each oncoprotein possesses independent transformation capabilities that are compounded when combined [194]. Furthermore, each oncoprotein has developed unique tactics to reprogram cells, yet these strategies often converge. Frequently, they involve preying on corresponding cellular processes through different targets, although, the overlap of specific cellular targets also occurs. For example, E6 and E7 both promote cell proliferation by affecting tumor suppressor proteins. E6 facilitates the degradation of p53 via the ubiquitin-proteasome system (UPS), whereas E7 drives cell cycle progression through targeting the tumor suppressor, pRB, either directly, through a shared mechanism of UPS mediated degradation of unphosphorylated pRB [195,196], or through indirect targeting of pocket proteins [197,198]. On the contrary, E7 expression has been shown to stabilize p53 [199–201], possibly through multiple mechanisms [202–204]. Furthermore, studies have demonstrated that E6AP has stabilizing effects on E7 in addition to E6 [205]. The interplay between E6 and E7 has led to the paradigm that the two oncoproteins work in concert to drive the viral life cycle, perhaps even intending to balance one another, as they inadvertently result in carcinogenesis as an unwanted by-product [206].

Similarities between the E6 and E7 oncoproteins extend beyond their shared host targets. E7 is a small protein of 100 amino acids that can be divided and defined by three conserved regions: CR1, CR2, and CR3 [50,207]. This division of primary sequence is analogous to that of adenovirus E1A to which the N-terminus of E7 proteins share sequence and functional homology [188]. These properties are also shared with the SV40 large T antigen and involve a fully conserved LXCXE motif which confers high affinity interaction with pRb [197,208–211]. Whereas in the CR3 region of E7 resides two CXXC motifs partitioned by approximately 30 residues creating a zinc binding domain [212–214], similar to the dual zinc binding domains located at the N- and C-terminals of E6 [113,215–217]. Comparison of E6N, E6C, and E7 zinc binding domain sequences suggests that E7 arose from a duplication event of a singular ancestral E6 domain [218,219]. Said ancestral sequence possibly belonged to a single-domain E6 protein present in the supposed proto-papillomavirus [62], with the discovery of single-domain avian and turtle E6 proteins further supporting this hypothesis [220]. Therefore, it is possible that E6N also stemmed from the duplication and divergence of this precursor [221]. Furthermore, despite low sequence identity between E6N and E6C sequences they share structural homology [222] in contrast to E7 which shares sequence identity with the E6 domain but adopts a significantly different fold [218]. This could indicate that while the proteins stem from a common ancestor, they adopt different functions leading to different protein conformations as they adapt to their niche. A theory reaffirmed by the low sequence conservation of E6 or E7 proteins from various viruses, indicating that the viral type necessitates the form and function of each set of proteins [218]. This is reflected in interactomic studies which have demonstrated that the different types of E6 or E7 proteins possess different interactomes [145,147,149,150,152–154,156,223–225] and supports the idea that some functions evolved are simply adaptive as opposed to perpetuating viral reproduction [61]. However, the question remains as to whether during this process E6 and E7 have co-evolved, working harmoniously with one another to drive the viral life cycle or, instead, are in competition with one another. Due to the bicistronic organization of the HPV genomes [226] the pair is always present together in infected cells. It stands to reason that this has driven co-evolution of the duo, however, it is possible that the two are not working in alliance and these adaptive functions could include attempts to adapt to selective pressure onset by direct competition between the pair. Overlap in targets could represent attempts to thwart one another and demonstrate functional advantage which instead backfires because the oncoproteins are continuously successful in outcompeting one another. For example, earlier in section 2.4 we described that E7 expression has a stabilizing effect on p53. This could have been a pressure exerted on E6 by E7 that was subverted by E6 through the emergence of the phospho-E6 proteoform and development of transcriptional mechanisms to regulate p53 [168,227–229]. Also mentioned is the stabilizing effect of E6AP on E7. In this study, increases

in the level of E6AP were also associated with increases of E7, but these were subject to regulation by E6 [205]. This indicates a potential regulatory loop between the oncoproteins, where E6 can exert control over E7 expression. Taken together, this could represent the remnants of a prior E7 interaction, which was lost after out competition by E6. Therefore as the selective pressure increases, the pair must become more inventive and ultimately the cumulative effects of their ingenuity results in cancer.

3. Conclusions

The process of HPV induced carcinogenesis is complex with many players involved. While much progress has been made, aiding our understanding of the respective roles of each oncoprotein, studies often focus on singular features of the singular components: structure, localization, expression levels, and individual targets or pathways. Interactomic studies have begun to open our eyes to the system-wide perturbances caused by these proteins, but much is left to be deciphered regarding how these proteins interact with the host and, potentially more importantly, how they interact with each other. This unholy trinity shares a motive to perpetuate the viral life cycle, and in their efforts to meet this shared objective, it could be as simple as having too many cooks in the kitchen. Their repeated efforts to target mutual cellular mechanisms are too successful, accumulating and leading to the cellular catastrophe that is cancer. Furthermore, it appears that certain host factors possess the power to convey a propensity for carcinogenesis meaning that the trio is not acting alone in determining this fate.

Funding sources

Work in the Trave lab is funded by institutional support from le Centre National de la Recherche Scientifique (CNRS) and l'Institut National de la Santé et de la Recherche Médicale (INSERM) with projects related to HPV receiving further funding from Ligue contre le Cancer and IMCBio.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

I would like to extend a warm thank you to Lawrence Banks for extending the opportunity to share this piece of work. I would like to the Gilles Trave for his endless support and the Trave team for always supplying input and interesting discussions.

Data availability

No data was used for the research described in the article.

References

[1] E.F. Dunne, I.U. Park, HPV and HPV-associated diseases, Infect. Dis. Clin. 27 (4) (2013) 765–778.
[2] M. Stanley, Pathology and epidemiology of HPV infection in females, Gynecol. Oncol. 117 (2 Suppl) (2010) S5–S10.
[3] J. Doorbar, The papillomavirus life cycle, J. Clin. Virol. 32 (2005) 7–15.
[4] J. Doorbar, Papillomavirus life cycle organization and biomarker selection, Dis. Markers 23 (4) (2007) 297–313.
[5] M. Herfs, et al., A discrete population of squamocolumnar junction cells implicated in the pathogenesis of cervical cancer, Proc. Natl. Acad. Sci. USA 109 (26) (2012) 10516–10521.
[6] D. Pyeon, et al., Establishment of human papillomavirus infection requires cell cycle progression, PLoS Pathog. 5 (2) (2009) e1000318.
[7] K. Egawa, Do human papillomaviruses target epidermal stem cells? Dermatology 207 (3) (2003) 251–254.
[8] Sheila V. Graham, The human papillomavirus replication cycle, and its links to cancer progression: a comprehensive review, Clin. Sci. 131 (17) (2017) 2201–2221.
[9] M. Bienkowska-Haba, et al., Cyclophilins facilitate dissociation of the human papillomavirus type 16 capsid protein L1 from the L2/DNA complex following virus entry, J. Virol. 86 (18) (2012) 9875–9887.
[10] S. DiGiuseppe, et al., The nuclear retention signal of HPV16 L2 protein is essential for incoming viral genome to transverse the trans-Golgi network, Virology 458–459 (2014) 93–105.
[11] G.A. Maglennon, P. McIntosh, J. Doorbar, Persistence of viral DNA in the epithelial basal layer suggests a model for papillomavirus latency following immune regression, Virology 414 (2) (2011) 153–163.
[12] H. Romanczuk, F. Thierry, P.M. Howley, Mutational analysis of cis elements involved in E2 modulation of human papillomavirus type 16 P97 and type 18 P105 promoters, J. Virol. 64 (6) (1990) 2849–2859.
[13] M. Bergvall, T. Melendy, J. Archambault, The E1 proteins, Virology 445 (1–2) (2013) 35–56.
[14] A.A. McBride, The papillomavirus E2 proteins, Virology 445 (1–2) (2013) 57–79.
[15] J.A. Scarth, et al., The human papillomavirus oncoproteins: a review of the host pathways targeted on the road to transformation, J. Gen. Virol. 102 (3) (2021).
[16] H.K. Wang, et al., Robust production and passaging of infectious HPV in squamous epithelium of primary human keratinocytes, Genes Dev. 23 (2) (2009) 181–194.
[17] N.S. Banerjee, et al., Human papillomavirus (HPV) E7 induces prolonged G2 following S phase reentry in differentiated human keratinocytes, J. Biol. Chem. 286 (17) (2011) 15473–15482.
[18] A. Pal, R. Kundu, Human papillomavirus E6 and E7: the cervical cancer hallmarks and targets for therapy, Front. Microbiol. 10 (2019) 3116.
[19] C. Cerqueira, J.T. Schiller, Papillomavirus assembly: an overview and perspectives, Virus Res. 231 (2017) 103–107.
[20] J. Doorbar, The human Papillomavirus twilight zone–Latency, immune control and subclinical infection, Tumour Virus Research (2023) 200268.
[21] N. Egawa, et al., Human papillomaviruses; epithelial tropisms, and the development of neoplasia, Viruses 7 (7) (2015) 3863–3890.
[22] T.A. Berman, J.T. Schiller, Human papillomavirus in cervical cancer and oropharyngeal cancer: one cause, two diseases, Cancer 123 (12) (2017) 2219–2229.
[23] J. Doorbar, et al., Human papillomavirus molecular biology and disease association, Rev. Med. Virol. 25 (Suppl 1) (2015) 2–23. Suppl Suppl 1.
[24] C. De Witte, et al., Imiquimod in cervical, vaginal and vulvar intraepithelial neoplasia: a review, Gynecol. Oncol. 139 (2) (2015) 377–384.
[25] K. Cuschieri, et al., Multiple high risk HPV infections are common in cervical neoplasia and young women in a cervical screening population, J. Clin. Pathol. 57 (1) (2004) 68–72.
[26] C.M. Martin, J.J. O'Leary, Histology of cervical intraepithelial neoplasia and the role of biomarkers, Best Pract. Res. Clin. Obstet. Gynaecol. 25 (5) (2011) 605–615.
[27] M.A. Molina, et al., HPV integration and cervical cancer: a failed evolutionary viral trait, Trends Mol. Med. 30 (9) (2024) 890–902.
[28] H. Deng, et al., HPV16-Immortalized cells from human transformation zone and endocervix are more dysplastic than ectocervical cells in organotypic culture, Sci. Rep. 8 (1) (2018) 15402.
[29] E. Burghardt, A. Ostör, Site and origin of squamous cervical cancer: a histomorphologic study, Obstet. Gynecol. 62 (1) (1983) 117–127.
[30] S. Hong, L.A. Laimins, Regulation of the life cycle of HPVs by differentiation and the DNA damage response, Future Microbiol. 8 (12) (2013) 1547–1557.
[31] A. Kirk, S.V. Graham, The human papillomavirus late life cycle and links to keratinocyte differentiation, J. Med. Virol. 96 (2) (2024) e29461.
[32] M. Pett, N. Coleman, Integration of high-risk human papillomavirus: a key event in cervical carcinogenesis? J. Pathol.: A Journal of the Pathological Society of Great Britain and Ireland 212 (4) (2007) 356–367.
[33] W. Li, et al., Characteristic of HPV integration in the genome and transcriptome of cervical cancer tissues, BioMed Res. Int. 2018 (2018) 6242173.
[34] M. Kamal, et al., Human papilloma virus (HPV) integration signature in Cervical Cancer: identification of MACROD2 gene as HPV hot spot integration site, Br. J. Cancer 124 (4) (2021) 777–785.
[35] Z. Hu, et al., Genome-wide profiling of HPV integration in cervical cancer identifies clustered genomic hot spots and a potential microhomology-mediated integration mechanism, Nat. Genet. 47 (2) (2015) 158–163.
[36] F.L. Sand, et al., Risk of CIN3 or worse with persistence of 13 individual oncogenic HPV types, Int. J. Cancer 144 (8) (2019) 1975–1982.
[37] J.E. Myers, et al., Detecting episomal or integrated human papillomavirus 16 DNA using an exonuclease V-qPCR-based assay, Virology 537 (2019) 149–156.
[38] J. Huang, et al., Comprehensive genomic variation profiling of cervical intraepithelial neoplasia and cervical cancer identifies potential targets for cervical cancer early warning, J. Med. Genet. 56 (3) (2019) 186–194.
[39] T. Hu, et al., Testing for viral DNA integration among HPV-positive women to detect cervical precancer: an observational cohort study, Bjog 131 (3) (2024) 309–318.
[40] S.R. Georgescu, et al., New insights in the pathogenesis of HPV infection and the associated carcinogenic processes: the role of chronic inflammation and oxidative stress, Journal of immunology research 2018 (1) (2018) 5315816.
[41] D. Estêvão, et al., Hallmarks of HPV carcinogenesis: the role of E6, E7 and E5 oncoproteins in cellular malignancy, Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms 1862 (2) (2019) 153–162.

[42] J.A. Westrich, C.J. Warren, D. Pyeon, Evasion of host immune defenses by human papillomavirus, Virus Res. 231 (2017) 21–33.
[43] N.E. Davey, G. Travé, T.J. Gibson, How viruses hijack cell regulation, Trends Biochem. Sci. 36 (3) (2011) 159–169.
[44] E. Beyer-Finkler, et al., Cell differentiation-related gene expression of human papillomavirus 33, Med. Microbiol. Immunol. 179 (1990) 185–192.
[45] M.H. Stoler, et al., Differentiation-linked human papillomavirus types 6 and 11 transcription in genital condylomata revealed by in situ hybridization with message-specific RNA probes, Virology 172 (1) (1989) 331–340.
[46] M.H. Stoler, et al., Human papillomavirus type 16 and 18 gene expression in cervical neoplasias, Hum. Pathol. 23 (2) (1992) 117–128.
[47] K. Middleton, et al., Organization of human papillomavirus productive cycle during neoplastic progression provides a basis for selection of diagnostic markers, J. Virol. 77 (19) (2003) 10186–10201.
[48] J.T. Thomas, et al., Human papillomavirus type 31 oncoproteins E6 and E7 are required for the maintenance of episomes during the viral life cycle in normal human keratinocytes, Proc. Natl. Acad. Sci. USA 96 (15) (1999) 8449–8454.
[49] R.B. Park, E.J. Androphy, Genetic analysis of high-risk e6 in episomal maintenance of human papillomavirus genomes in primary human keratinocytes, J. Virol. 76 (22) (2002) 11359–11364.
[50] A. Roman, K. Munger, The papillomavirus E7 proteins, Virology 445 (1) (2013) 138–168.
[51] C.A. Moody, L.A. Laimins, Human papillomavirus oncoproteins: pathways to transformation, Nat. Rev. Cancer 10 (8) (2010) 550–560.
[52] M. Stanley, Pathology and epidemiology of HPV infection in females, Gynecol. Oncol. 117 (2) (2010) S5–S10.
[53] S. Jeon, B.L. Allen-Hoffmann, P.F. Lambert, Integration of human papillomavirus type 16 into the human genome correlates with a selective growth advantage of cells, J. Virol. 69 (5) (1995) 2989–2997.
[54] A.A. McBride, A. Warburton, The role of integration in oncogenic progression of HPV-associated cancers, PLoS Pathog. 13 (4) (2017) e1006211.
[55] O. Basukala, L. Banks, The not-so-good, the bad and the ugly: HPV E5, E6 and E7 oncoproteins in the orchestration of carcinogenesis, Viruses 13 (10) (2021).
[56] T. Gheit, Mucosal and cutaneous human papillomavirus infections and cancer biology, Front. Oncol. 9 (2019).
[57] P. Leechanachai, et al., The E5 gene from human papillomavirus type 16 is an oncogene which enhances growth factor-mediated signal transduction to the nucleus, Oncogene 7 (1) (1992) 19–25.
[58] D. Pim, M. Collins, L. Banks, Human papillomavirus type 16 E5 gene stimulates the transforming activity of the epidermal growth factor receptor, Oncogene 7 (1) (1992) 27–32.
[59] S.W. Straight, et al., The E5 oncoprotein of human papillomavirus type 16 transforms fibroblasts and effects the downregulation of the epidermal growth factor receptor in keratinocytes, J. Virol. 67 (8) (1993) 4521–4532.
[60] M. Schiffman, et al., The carcinogenicity of human papillomavirus types reflects viral evolution, Virology 337 (1) (2005) 76–84.
[61] K. Van Doorslaer, Evolution of the papillomaviridae, Virology 445 (1) (2013) 11–20.
[62] S. García-Vallvé, Á. Alonso, I.G. Bravo, Papillomaviruses: different genes have different histories, Trends Microbiol. 13 (11) (2005) 514–521.
[63] I.G. Bravo, A. Alonso, Mucosal human papillomaviruses encode four different E5 proteins whose chemistry and phylogeny correlate with malignant or benign growth, J. Virol. 78 (24) (2004) 13613–13626.
[64] R. Nath, et al., Analyses of variant human papillomavirus type-16 E5 proteins for their ability to induce mitogenesis of murine fibroblasts, Cancer Cell Int. 6 (2006) 1–9.
[65] A. Venuti, et al., Papillomavirus E5: the smallest oncoprotein with many functions, Mol. Cancer 10 (2011) 1–18.
[66] L. Petti, L.A. Nilson, D. DiMaio, Activation of the platelet-derived growth factor receptor by the bovine papillomavirus E5 transforming protein, EMBO J. 10 (4) (1991) 845–855.
[67] A. Siddiqa, J. Broniarczyk, L. Banks, Papillomaviruses and endocytic trafficking, Int. J. Mol. Sci. 19 (9) (2018) 2619.
[68] V. Bouvard, et al., The human papillomavirus type 16 E5 gene cooperates with the E7 gene to stimulate proliferation of primary cells and increases viral gene expression, Virology 203 (1) (1994) 73–80.
[69] G.F. Valle, L. Banks, The human papillomavirus (HPV)-6 and HPV-16 E5 proteins co-operate with HPV-16 E7 in the transformation of primary rodent cells, J. Gen. Virol. 76 (5) (1995) 1239–1245.
[70] M.C. Stöppler, et al., The E5 gene of HPV-16 enhances keratinocyte immortalization by full-length DNA, Virology 223 (1) (1996) 251–254.
[71] E. Krawczyk, et al., Koilocytosis: a cooperative interaction between the human papillomavirus E5 and E6 oncoproteins, Am. J. Pathol. 173 (3) (2008) 682–688.
[72] J.P. Maufort, et al., A role for HPV16 E5 in cervical carcinogenesis, Cancer Res. 70 (7) (2010) 2924–2931.
[73] A.C. de Freitas, et al., hrHPV E5 oncoprotein: immune evasion and related immunotherapies, J. Exp. Clin. Cancer Res. 36 (2017) 1–15.
[74] C.L. Halbert, D.A. Galloway, Identification of the E5 open reading frame of human papillomavirus type 16, J. Virol. 62 (3) (1988) 1071–1075.
[75] D. DiMaio, L.M. Petti, The E5 proteins, Virology 445 (1) (2013) 99–114.
[76] E. Krawczyk, et al., The human papillomavirus type 16 E5 oncoprotein translocates calpactin I to the perinuclear region, J. Virol. 85 (21) (2011) 10968–10975.
[77] E. Krawczyk, et al., Membrane orientation of the human papillomavirus type 16 E5 oncoprotein, J. Virol. 84 (4) (2010) 1696–1703.
[78] B. Chen, et al., Advances in molecular mechanism of HPV16 E5 oncoprotein carcinogenesis, Arch. Biochem. Biophys. 745 (2023) 109716.
[79] L. Carrasco, Modification of membrane permeability by animal viruses, Adv. Virus Res. 45 (1995) 61–112.
[80] L.F. Wetherill, et al., High-risk human papillomavirus E5 oncoprotein displays channel-forming activity sensitive to small-molecule inhibitors, J. Virol. 86 (9) (2012) 5341–5351.
[81] J.L. Nieva, V. Madan, L. Carrasco, Viroporins: structure and biological functions, Nat. Rev. Microbiol. 10 (8) (2012) 563–574.
[82] J. Royle, et al., Emerging Roles of viroporins Encoded by DNA viruses: novel Targets for antivirals? Viruses 7 (10) (2015) 5375–5387.
[83] S.R. Sudarshan, R. Schlegel, X. Liu, Two conserved amino acids differentiate the biology of high-risk and low-risk HPV E5 proteins, J. Med. Virol. 94 (9) (2022) 4565–4575.
[84] M. Eilers, et al., Internal packing of helical membrane proteins, Proc. Natl. Acad. Sci. USA 97 (11) (2000) 5796–5801.
[85] H. Sun, et al., The preference of tryptophan for membrane interfaces: insights from N-methylation of tryptophans in gramicidin channels, J. Biol. Chem. 283 (32) (2008) 22233–22243.
[86] Y. Wang, et al., Impact of histidine residues on the transmembrane helices of viroporins, Mol. Membr. Biol. 30 (7) (2013) 360–369.
[87] P. Venkataraman, R.A. Lamb, L.H. Pinto, Chemical rescue of histidine selectivity filter mutants of the M2 ion channel of influenza A virus, J. Biol. Chem. 280 (22) (2005) 21463–21472.
[88] M. Saidijam, S. Azizpour, S.G. Patching, Comprehensive analysis of the numbers, lengths and amino acid compositions of transmembrane helices in prokaryotic, eukaryotic and viral integral membrane proteins of high-resolution structure, J. Biomol. Struct. Dyn. 36 (2) (2018) 443–464.
[89] M. Conrad, V. Bubb, R. Schlegel, The human papillomavirus type 6 and 16 E5 proteins are membrane-associated proteins which associate with the 16-kilodalton pore-forming protein, J. Virol. 67 (10) (1993) 6170–6178.
[90] M. Müller, et al., YIP1 family member 4 (YIPF4) is a novel cellular binding partner of the papillomavirus E5 proteins, Sci. Rep. 5 (1) (2015) 12523.
[91] B. Zhang, et al., The E5 protein of human papillomavirus type 16 perturbs MHC class II antigen maturation in human foreskin keratinocytes treated with interferon-γ, Virology 310 (1) (2003) 100–108.
[92] K. Crusius, et al., The human papillomavirus type 16 E5-protein modulates ligand-dependent activation of the EGF receptor family in the human epithelial cell line HaCaT, Exp. Cell Res. 241 (1) (1998) 76–83.
[93] K. Kabsch, A. Alonso, The human papillomavirus type 16 E5 protein impairs TRAIL-and FasL-mediated apoptosis in HaCaT cells by different mechanisms, J. Virol. 76 (23) (2002) 12162–12172.
[94] I. Bravo, K. Crusius, A. Alonso, The E5 protein of the human papillomavirus type 16 modulates composition and dynamics of membrane lipids in keratinocytes, Arch. Virol. 150 (2005) 231–246.
[95] F. Suprynowicz, et al., HPV-16 E5 oncoprotein upregulates lipid raft components caveolin-1 and ganglioside GM1 at the plasma membrane of cervical cells, Oncogene 27 (8) (2008) 1071–1078.
[96] G.H. Ashrafi, et al., E5 protein of human papillomavirus 16 downregulates HLA class I and interacts with the heavy chain via its first hydrophobic domain, Int. J. Cancer 119 (9) (2006) 2105–2112.
[97] E. Schwarz, et al., Structure and transcription of human papillomavirus sequences in cervical carcinoma cells, Nature 314 (6006) (1985) 111–114.
[98] E. Androphy, et al., Identification of the HPV-16 E6 protein from transformed mouse cells and human cervical carcinoma cell lines, EMBO J. 6 (4) (1987) 989–992.
[99] L. Banks, et al., Identification of human papillomavirus type 18 E6 polypeptide in cells derived from human cervical carcinomas, J. Gen. Virol. 68 (5) (1987) 1351–1359.
[100] V. Band, et al., Loss of p53 protein in human papillomavirus type 16 E6-immortalized human mammary epithelial cells, J. Virol. 65 (12) (1991) 6671–6676.
[101] Z. Liu, et al., The E6 gene of human papillomavirus type 16 is sufficient for transformation of baby rat kidney cells in cotransfection with activated Ha-ras, Virology 201 (2) (1994) 388–396.
[102] S. Sedman, et al., The full-length E6 protein of human papillomavirus type 16 has transforming and trans-activating activities and cooperates with E7 to immortalize keratinocytes in culture, J. Virol. 65 (9) (1991) 4860–4866.
[103] A. Storey, L. Banks, Human papillomavirus type 16 E6 gene cooperates with EJ-ras to immortalize primary mouse cells, Oncogene 8 (4) (1993) 919–924.
[104] D. Pim, et al., Mutational analysis of HPV-18 E6 identifies domains required for p53 degradation in vitro, abolition of p53 transactivation in vivo and immortalisation of primary BMK cells, Oncogene 9 (7) (1994) 1869–1876.
[105] D.E. Wazer, et al., Immortalization of distinct human mammary epithelial cell types by human papilloma virus 16 E6 or E7, Proc. Natl. Acad. Sci. USA 92 (9) (1995) 3687–3691.
[106] C. Halbert, G. Demers, D. Galloway, The E7 gene of human papillomavirus type 16 is sufficient for immortalization of human epithelial cells, J. Virol. 65 (1) (1991) 473–478.
[107] S. Song, et al., Human papillomavirus types 16 E6 and E7 contribute differently to carcinogenesis, Virology 267 (2) (2000) 141–150.
[108] K. Münger, et al., The E6 and E7 genes of the human papillomavirus type 16 together are necessary and sufficient for transformation of primary human keratinocytes, J. Virol. 63 (10) (1989) 4417–4421.
[109] H.L. Howie, R.A. Katzenellenbogen, D.A. Galloway, Papillomavirus E6 proteins, Virology 384 (2) (2009) 324–334.

[110] Y. Nominé, et al., Structural and functional analysis of E6 oncoprotein: insights in the molecular pathways of human papillomavirus-mediated pathogenesis, Mol. Cell 21 (5) (2006) 665–678.
[111] K. Zanier, et al., Solution structure analysis of the HPV16 E6 oncoprotein reveals a self-association mechanism required for E6-mediated degradation of p53, Structure 20 (4) (2012) 604–617.
[112] S. Cole, O. Danos, Nucleotide sequence and comparative analysis of the human papillomavirus type 18 genome: phylogeny of papillomaviruses and repeated structure of the E6 and E7 gene products, J. Mol. Biol. 193 (4) (1987) 599–608.
[113] M.S. Barbosa, D.R. Lowy, J.T. Schiller, Papillomavirus polypeptides E6 and E7 are zinc-binding proteins, J. Virol. 63 (3) (1989) 1404–1407.
[114] T. Kanda, et al., Human papillomavirus type 16 E6 proteins with glycine substitution for cysteine in the metal-binding motif, Virology 185 (2) (1991) 536–543.
[115] L. Sherman, R. Schlegel, Serum-and calcium-induced differentiation of human keratinocytes is inhibited by the E6 oncoprotein of human papillomavirus type 16, J. Virol. 70 (5) (1996) 3269–3279.
[116] M.J. Plevin, M.M. Mills, M. Ikura, The LxxLL motif: a multifunctional binding sequence in transcriptional regulation, Trends Biochem. Sci. 30 (2) (2005) 66–69.
[117] P. Madhu, N.E. Davey, Y. Ivarsson, How viral proteins bind short linear motifs and intrinsically disordered domains, Essays Biochem. 66 (7) (2022) 935–944.
[118] J.M. Huibregtse, M. Scheffner, P.M. Howley, Localization of the E6-AP regions that direct human papillomavirus E6 binding, association with p53, and ubiquitination of associated proteins, Mol. Cell Biol. 13 (8) (1993) 4918–4927.
[119] J.J. Chen, et al., Identification of an α helical motif sufficient for association with papillomavirus E6, J. Biol. Chem. 273 (22) (1998) 13537–13544.
[120] S.B.V. Pol, M.C. Brown, C.E. Turner, Association of bovine papillomavirus type 1 E6 oncoprotein with the focal adhesion protein paxillin through a conserved protein interaction motif, Oncogene 16 (1) (1998) 43–52.
[121] J. Bohl, et al., Competitive binding to a charged leucine motif represses transformation by a papillomavirus E6 oncoprotein, Virology 271 (1) (2000) 163–170.
[122] N. Brimer, C.M. Drews, S.B. Vande Pol, Association of papillomavirus E6 proteins with either MAML1 or E6AP clusters E6 proteins by structure, function, and evolutionary relatedness, PLoS Pathog. 13 (12) (2017) e1006781.
[123] C.M. Drews, S. Case, S.B. Vande Pol, E6 proteins from high-risk HPV, low-risk HPV, and animal papillomaviruses activate the Wnt/β-catenin pathway through E6AP-dependent degradation of NHERF1, PLoS Pathog. 15 (4) (2019) e1007575.
[124] N. Brimer, et al., Cutaneous papillomavirus E6 oncoproteins associate with MAML1 to repress transactivation and NOTCH signaling, Oncogene 31 (43) (2012) 4639–4646.
[125] J.M. Meyers, J.M. Spangle, K. Munger, The human papillomavirus type 8 E6 protein interferes with NOTCH activation during keratinocyte differentiation, J. Virol. 87 (8) (2013) 4762–4767.
[126] M.J.A. Tan, et al., Cutaneous β-human papillomavirus E6 proteins bind Mastermind-like coactivators and repress Notch signaling, Proc. Natl. Acad. Sci. USA 109 (23) (2012) E1473–E1480.
[127] C.M. Drews, N. Brimer, S.B. Vande Pol, Multiple regions of E6AP (UBE3A) contribute to interaction with papillomavirus E6 proteins and the activation of ubiquitin ligase activity, PLoS Pathog. 16 (1) (2020) e1008295.
[128] Z. Songyang, et al., Recognition of unique carboxyl-terminal motifs by distinct PDZ domains, Science 275 (5296) (1997) 73–77.
[129] K. Ganti, et al., The human papillomavirus E6 PDZ binding motif: from life cycle to malignancy, Viruses 7 (7) (2015) 3530–3551.
[130] A.Y. Hung, M. Sheng, PDZ domains: structural modules for protein complex assembly, J. Biol. Chem. 277 (8) (2002) 5699–5702.
[131] F. Ye, M. Zhang, Structures and target recognition modes of PDZ domains: recurring themes and emerging pictures, Biochem. J. 455 (1) (2013) 1–14.
[132] S.S. Lee, R.S. Weiss, R.T. Javier, Binding of human virus oncoproteins to hDlg/SAP97, a mammalian homolog of the Drosophila discs large tumor suppressor protein, Proc. Natl. Acad. Sci. USA 94 (13) (1997) 6670–6675.
[133] T. Kiyono, et al., Binding of high-risk human papillomavirus E6 oncoproteins to the human homologue of the Drosophila discs large tumor suppressor protein, Proc. Natl. Acad. Sci. USA 94 (21) (1997) 11612–11616.
[134] M. Thomas, et al., Analysis of multiple HPV E6 PDZ interactions defines type-specific PDZ fingerprints that predict oncogenic potential, PLoS Pathog. 12 (8) (2016) e1005766.
[135] M. Scheffner, et al., The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53, Cell 63 (6) (1990) 1129–1136.
[136] T. Crook, J.A. Tidy, K.H. Vousden, Degradation of p53 can be targeted by HPV E6 sequences distinct from those required for p53 binding and trans-activation, Cell 67 (3) (1991) 547–556.
[137] C.H. Storrs, S.J. Silverstein, PATJ, a tight junction-associated PDZ protein, is a novel degradation target of high-risk human papillomavirus E6 and the alternatively spliced isoform 18 E6, J. Virol. 81 (8) (2007) 4080–4090.
[138] P. Kuballa, K. Matentzoglu, M. Scheffner, The role of the ubiquitin ligase E6-AP in human papillomavirus E6-mediated degradation of PDZ domain-containing proteins, J. Biol. Chem. 282 (1) (2007) 65–71.
[139] P. Massimi, et al., HPV E6 degradation of p53 and PDZ containing substrates in an E6AP null background, Oncogene 27 (12) (2008) 1800–1804.
[140] D. Pim, et al., HPV E6 targeted degradation of the discs large protein: evidence for the involvement of a novel ubiquitin ligase, Oncogene 19 (6) (2000) 719–725.
[141] H.S. Grm, L. Banks, Degradation of hDlg and MAGIs by human papillomavirus E6 is E6-AP-independent, J. Gen. Virol. 85 (10) (2004) 2815–2819.
[142] D. Martinez-Zapien, et al., Structure of the E6/E6AP/p53 complex required for HPV-mediated degradation of p53, Nature 529 (7587) (2016) 541–545.
[143] J.C.K. Wang, et al., Structure of the p53 degradation complex from HPV16, Nat. Commun. 15 (1) (2024) 1842.
[144] Z. Wang, et al., Structural insights into the functional mechanism of the ubiquitin ligase E6AP, Nat. Commun. 15 (1) (2024) 3531.
[145] M. Eckhardt, et al., Multiple routes to oncogenesis are promoted by the human papillomavirus-host protein network, Cancer Discov. 8 (11) (2018) 1474–1489.
[146] M. Grace, K. Munger, Proteomic analysis of the gamma human papillomavirus type 197 E6 and E7 associated cellular proteins, Virology 500 (2017) 71–81.
[147] O. Rozenblatt-Rosen, et al., Interpreting cancer genomes using systematic host network perturbations by tumour virus proteins, Nature 487 (7408) (2012) 491–495.
[148] S.W. Strickland, et al., Human Papillomavirus E6 interaction with cellular PDZ domain proteins modulates YAP nuclear localization, Virology 516 (2018) 127–138.
[149] J. Wang, et al., Sterile α motif domain containing 9 is a novel cellular interacting partner to low-risk type human papillomavirus E6 proteins, PLoS One 11 (2) (2016) e0149859.
[150] E.A. White, et al., Comprehensive analysis of host cellular interactions with human papillomavirus E6 proteins identifies new E6 binding partners and reflects viral diversity, J. Virol. 86 (24) (2012) 13174–13186.
[151] E.A. White, P.M. Howley, Proteomic approaches to the study of papillomavirus–host interactions, Virology 435 (1) (2013) 57–69.
[152] G. Neveu, et al., Comparative analysis of virus–host interactomes with a mammalian high-throughput protein complementation assay based on Gaussia princeps luciferase, Methods 58 (4) (2012) 349–359.
[153] J. Poirson, et al., Mapping the interactome of HPV E6 and E7 oncoproteins with the ubiquitin-proteasome system, FEBS J. 284 (19) (2017) 3171–3201.
[154] B. Xue, et al., Disordered interactome of human papillomavirus, Curr. Pharmaceut. Des. 20 (8) (2014) 1274–1292.
[155] K. Munger, E.A. White, What are the essential determinants of human papillomavirus carcinogenesis? mBio (2024) e0046224.
[156] J. Skelin, V. Tomaić, Comparative analysis of alpha and beta HPV E6 oncoproteins: insights into functional distinctions and divergent mechanisms of pathogenesis, Viruses 15 (11) (2023) 2253.
[157] Q. Peng, et al., HPV E6/E7: insights into their regulatory role and mechanism in signaling pathways in HPV-associated tumor, Cancer Gene Ther. 31 (1) (2024) 9–17.
[158] A. Vats, et al., Human papillomavirus E6 and E7: what remains? Tumour Virus Research 11 (2021) 200213.
[159] A. Vats, J.V. Thatte, L. Banks, Molecular dissection of the E6 PBM identifies essential residues regulating Chk1 phosphorylation and subsequent 14-3-3 recognition, Tumour Virus Res 15 (2023) 200257.
[160] K. Nagasaka, et al., PDZ domains and viral infection: versatile potentials of HPV-PDZ interactions in relation to malignancy, BioMed Res. Int. 2013 (2013) 369712.
[161] S.S. Boon, L. Banks, High-risk human papillomavirus E6 oncoproteins interact with 14-3-3ζ in a PDZ binding motif-dependent manner, J. Virol. 87 (3) (2013) 1586–1595.
[162] C. KuËhne, et al., Differential regulation of human papillomavirus E6 by protein kinase A: conditional degradation of human discs large protein by oncogenic E6, Oncogene 19 (51) (2000) 5884–5891.
[163] J. Thatte, et al., The human papillomavirus E6 PDZ binding motif links DNA damage response signaling to E6 inhibition of p53 transcriptional activity, J. Virol. 92 (16) (2018), https://doi.org/10.1128/jvi.00465-18.
[164] V. Obsilova, T. Obsil, Structural insights into the functional roles of 14-3-3 proteins, Front. Mol. Biosci. 9 (2022).
[165] H.L. Howie, et al., Beta-HPV 5 and 8 E6 promote p300 degradation by blocking AKT/p300 association, PLoS Pathog. 7 (8) (2011) e1002211.
[166] K.L. Pennington, et al., The dynamic and stress-adaptive signaling hub of 14-3-3: emerging mechanisms of regulation and context-dependent protein–protein interactions, Oncogene 37 (42) (2018) 5587–5604.
[167] J.C. Higareda-Almaraz, et al., Proteomic patterns of cervical cancer cell lines, a network perspective, BMC Syst. Biol. 5 (2011) 96.
[168] A. Vats, et al., Loss of the E6AP ubiquitin ligase induces p53-dependent phosphorylation of human papillomavirus 18 E6 in cells derived from cervical cancer, J. Virol. 96 (6) (2022) e0150321.
[169] L. Olmedo-Nieva, et al., The role of E6 spliced isoforms (E6*) in human papillomavirus-induced carcinogenesis, Viruses 10 (1) (2018) 45.
[170] S.V. Graham, A.A.A. Faizo, Control of human papillomavirus gene expression by alternative splicing, Virus Res. 231 (2017) 83–95.
[171] T. Mesplède, et al., p53 degradation activity, expression, and subcellular localization of E6 proteins from 29 human papillomavirus genotypes, J. Virol. 86 (1) (2012) 94–107.
[172] Z.-M. Zheng, et al., Splicing of a cap-proximal human papillomavirus 16 E6E7 intron promotes E7 expression, but can be restrained by distance of the intron from its RNA 5′ cap, J. Mol. Biol. 337 (5) (2004) 1091–1108.
[173] S. Tang, et al., The E7 oncoprotein is translated from spliced E6* I transcripts in high-risk human papillomavirus type 16-or type 18-positive cervical cancer cell lines via translation reinitiation, J. Virol. 80 (9) (2006) 4249–4263.
[174] S.N. Stacey, et al., Leaky scanning is the predominant mechanism for translation of human papillomavirus type 16 E7 oncoprotein from E6/E7 bicistronic mRNA, J. Virol. 74 (16) (2000) 7284–7297.
[175] O. del Moral-Hernández, et al., The HPV-16 E7 oncoprotein is expressed mainly from the unspliced E6/E7 transcript in cervical carcinoma C33-A cells, Arch. Virol. 155 (2010) 1959–1970.
[176] D. Pim, L. Banks, HPV-18 E6* I protein modulates the E6-directed degradation of p53 by binding to full-length HPV-18 E6, Oncogene 18 (52) (1999) 7403–7408.

[177] D. Pim, V. Tomaic, L. Banks, The human papillomavirus (HPV) E6* proteins from high-risk, mucosal HPVs can direct degradation of cellular proteins in the absence of full-length E6 protein, J. Virol. 83 (19) (2009) 9863–9874.
[178] M. Filippova, et al., The small splice variant of HPV16 E6, E6*, reduces tumor formation in cervical carcinoma xenografts, Virology 450 (2014) 153–164.
[179] M. Filippova, et al., Complexes of human papillomavirus type 16 E6 proteins form pseudo-death-inducing signaling complex structures during tumor necrosis factor-mediated apoptosis, J. Virol. 83 (1) (2009) 210–227.
[180] Y. Wang, et al., Alternative splicing in the genome of HPV and its regulation, Front. Cell. Infect. Microbiol. 14 (2024) 1443868.
[181] J.O. Muñoz-Bello, et al., HPV-18 E6 oncoprotein and its spliced isoform E6* I regulate the Wnt/β-catenin cell signaling pathway through the TCF-4 transcriptional factor, Int. J. Mol. Sci. 19 (10) (2018) 3153.
[182] V.M. Williams, et al., Human papillomavirus type 16 E6* induces oxidative stress and DNA damage, J. Virol. 88 (12) (2014) 6751–6761.
[183] P. Paget-Bailly, et al., Comparative RNA sequencing reveals that HPV16 E6 abrogates the effect of E6* I on ROS metabolism, Sci. Rep. 9 (1) (2019) 5938.
[184] C. Artaza-Irigaray, et al., E6/E7 and E6* from HPV16 and HPV18 upregulate IL-6 expression independently of p53 in keratinocytes, Front. Immunol. 10 (2019) 1676.
[185] S. Yasumoto, et al., Human papillomavirus type 16 DNA-induced malignant transformation of NIH 3T3 cells, J. Virol. 57 (2) (1986) 572–577.
[186] M.A. Bedell, et al., Identification of human papillomavirus type 18 transforming genes in immortalized and primary cells, J. Virol. 63 (3) (1989) 1247–1255.
[187] T. Kanda, A. Furuno, K. Yoshiike, Human papillomavirus type 16 open reading frame E7 encodes a transforming gene for rat 3Y1 cells, J. Virol. 62 (2) (1988) 610–613.
[188] W.C. Phelps, et al., The human papillomavirus type 16 E7 gene encodes transactivation and transformation functions similar to those of adenovirus E1A, Cell 53 (4) (1988) 539–547.
[189] A. Tanaka, et al., Identification of a transforming gene of human papillomavirus type 16, J. Virol. 63 (3) (1989) 1465–1469.
[190] K.H. Vousden, et al., The E7 open reading frame of human papillomavirus type 16 encodes a transforming gene, Oncogene Res. 3 (2) (1988) 167–175.
[191] M. Yutsudo, Y. Okamoto, A. Hakura, Functional dissociation of transforming genes of human papillomavirus type 16, Virology 166 (2) (1988) 594–597.
[192] P. Hawley-Nelson, et al., HPV16 E6 and E7 proteins cooperate to immortalize human foreskin keratinocytes, EMBO J. 8 (12) (1989) 3905–3910.
[193] J.B. Hudson, et al., Immortalization and altered differentiation of human keratinocytes in vitro by the E6 and E7 open reading frames of human papillomavirus type 18, J. Virol. 64 (2) (1990) 519–526.
[194] H. zur Hausen, Papillomaviruses and cancer: from basic studies to clinical application, Nat. Rev. Cancer 2 (5) (2002) 342–350.
[195] Y. Imai, et al., Purification and characterization of human papillomavirus type 16 E7 protein with preferential binding capacity to the underphosphorylated form of retinoblastoma gene product, J. Virol. 65 (9) (1991) 4966–4972.
[196] S.N. Boyer, D.E. Wazer, V. Band, E7 protein of human papilloma virus-16 induces degradation of retinoblastoma protein through the ubiquitin-proteasome pathway, Cancer Res. 56 (20) (1996) 4620–4624.
[197] N. Dyson, et al., The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product, Science 243 (4893) (1989) 934–937.
[198] K. Münger, et al., Complex formation of human papillomavirus E7 proteins with the retinoblastoma tumor suppressor gene product, EMBO J. 8 (13) (1989) 4099–4105.
[199] G.W. Demers, C.L. Halbert, D.A. Galloway, Elevated wild-type p53 protein levels in human epithelial cell lines immortalized by the human papillomavirus type 16 E7 gene, Virology 198 (1) (1994) 169–174.
[200] D.L. Jones, D.A. Thompson, K. Münger, Destabilization of the RB tumor suppressor protein and stabilization of p53 contribute to HPV type 16 E7-induced apoptosis, Virology 239 (1) (1997) 97–107.
[201] J.T. Thomas, L.A. Laimins, Human papillomavirus oncoproteins E6 and E7 independently abrogate the mitotic spindle checkpoint, J. Virol. 72 (2) (1998) 1131–1137.
[202] D.L. Jones, K. Münger, Analysis of the p53-mediated G1 growth arrest pathway in cells expressing the human papillomavirus type 16 E7 oncoprotein, J. Virol. 71 (4) (1997) 2905–2912.
[203] M. Fischer, et al., Human papilloma virus E7 oncoprotein abrogates the p53-p21-DREAM pathway, Sci. Rep. 7 (1) (2017) 2603.
[204] S. Sharma, K. Munger, KDM6A-Mediated expression of the long noncoding RNA DINO causes TP53 tumor suppressor stabilization in human papillomavirus 16 E7-expressing cells, J. Virol. 94 (12) (2020), https://doi.org/10.1128/jvi.02178-19.
[205] A. Vats, et al., Regulation of HPV E7 stability by E6-associated protein (E6AP), J. Virol. 96 (16) (2022) e00663, 22.
[206] S.B. Vande Pol, A.J. Klingelhutz, Papillomavirus E6 oncoproteins, Virology 445 (1–2) (2013) 115–137.
[207] O. Ohlenschläger, et al., Solution structure of the partially folded high-risk human papilloma virus 45 oncoprotein E7, Oncogene 25 (44) (2006) 5953–5959.
[208] A. Dahiya, et al., Role of the LXCXE binding site in Rb function, Mol. Cell Biol. 20 (18) (2000) 6799–6805.
[209] X. Liu, R. Marmorstein, Structure of the retinoblastoma protein bound to adenovirus E1A reveals the molecular basis for viral oncoprotein inactivation of a tumor suppressor, Genes Dev. 21 (21) (2007) 2711–2716.
[210] J.O. Lee, A.A. Russo, N.P. Pavletich, Structure of the retinoblastoma tumour-suppressor pocket domain bound to a peptide from HPV E7, Nature 391 (6670) (1998) 859–865.
[211] M. Singh, et al., Molecular determinants for the complex formation between the retinoblastoma protein and LXCXE sequences, J. Biol. Chem. 280 (45) (2005) 37868–37876.
[212] E.J. Roth, et al., Metal thiolate coordination in the E7 proteins of human papilloma virus 16 and cottontail rabbit papilloma virus as expressed in Escherichia coli, J. Biol. Chem. 267 (23) (1992) 16390–16395.
[213] M.C. McIntyre, et al., Human papillomavirus type 18 E7 protein requires intact Cys-X-X-Cys motifs for zinc binding, dimerization, and transformation but not for Rb binding, J. Virol. 67 (6) (1993) 3142–3150.
[214] G. Pahel, et al., Structural and functional characterization of the HPV16 E7 protein expressed in bacteria, J. Biol. Chem. 268 (34) (1993) 26018–26025.
[215] F. Lipari, et al., Purification and biophysical characterization of a minimal functional domain and of an N-terminal Zn2+-binding fragment from the human papillomavirus type 16 E6 protein, Biochemistry 40 (5) (2001) 1196–1204.
[216] Y. Nominé, et al., Domain substructure of HPV E6 oncoprotein: biophysical characterization of the E6 C-terminal DNA-binding domain, Biochemistry 42 (17) (2003) 4909–4917.
[217] Y. Liu, et al., Determinants of stability for the E6 protein of papillomavirus type 16, J. Mol. Biol. 386 (4) (2009) 1123–1137.
[218] K. Van Doorslaer, R.D. Burk, Evolution of human papillomavirus carcinogenicity, Adv. Virus Res. 77 (2010) 41–62.
[219] S.T. Cole, O. Danos, Nucleotide sequence and comparative analysis of the human papillomavirus type 18 genome. Phylogeny of papillomaviruses and repeated structure of the E6 and E7 gene products, J. Mol. Biol. 193 (4) (1987) 599–608.
[220] L.H. Herbst, et al., Genomic characterization of two novel reptilian papillomaviruses, Chelonia mydas papillomavirus 1 and Caretta caretta papillomavirus 1, Virology 383 (1) (2009) 131–135.
[221] K. Van Doorslaer, et al., Identification of unusual E6 and E7 proteins within avian papillomaviruses: cellular localization, biophysical characterization, and phylogenetic analysis, J. Virol. 83 (17) (2009) 8759–8770.
[222] I. Suarez, G. Trave, Structural insights in multifunctional papillomavirus oncoproteins, Viruses 10 (1) (2018).
[223] E.A. White, et al., Systematic identification of interactions between host cell proteins and E7 oncoproteins from diverse human papillomaviruses, Proc. Natl. Acad. Sci. USA 109 (5) (2012) E260–E267.
[224] E.A. White, P.M. Howley, Proteomic approaches to the study of papillomavirus-host interactions, Virology 435 (1) (2013) 57–69.
[225] Q.u.A. Farooq, et al., Inferring Virus-Host relationship between HPV and its host Homo sapiens using protein interaction network, Sci. Rep. 10 (1) (2020) 8719.
[226] K.L.Z. Van Doorslaer, S. Xirasagar, P. Maes, D. Kaminsky, et al., The papillomavirus Episteme: a major update to the papillomavirus sequence database, Nucleic Acids Res. 45 (2017) D499–D506.
[227] A. Kumar, et al., Human papillomavirus oncoprotein E6 inactivates the transcriptional coactivator human ADA3, Mol. Cell Biol. 22 (16) (2002) 5801–5812.
[228] D. Patel, et al., The E6 protein of human papillomavirus type 16 binds to and inhibits co-activation by CBP and p300, EMBO J. 18 (18) (1999) 5061–5072.
[229] H. Zimmermann, et al., The human papillomavirus type 16 E6 oncoprotein can down-regulate p53 activity by targeting the transcriptional coactivator CBP/p300, J. Virol. 73 (8) (1999) 6209–6219.

Introduction to research project

Thesis Project and Objective

Protein interaction networks encompass the entirety of a cell's protein-protein interactions (PPIs) creating a dynamic network of interactions that are temporal, often transient, and contextual by nature. These networks, or interactomes, determine cellular functions and their regulation therefore perturbations within them underpin the development of disease. This places interactomes at a critical crossroads linking genotype and phenotype, holding key insights into pathogenesis. Human Papilloma Viruses (HPVs) have often been subject to interactomic studies as they are highly diverse genotypically and phenotypically, resulting in clinical presentations that range from transient subclinical to persistent and cancer causing. Moreover, cancer progression induced by high-risk HPVs (hr-HPVs) is driven by a set of viral oncoproteins— E5, E6, and E7— of which E6 has been established as a promiscuous interactor, targeting a myriad of host proteins. Despite previous characterization of E6 by interactomic methods, these studies are limited to qualitative techniques that only provide descriptions or lists of interactions and lack molecular detail.

Previous studies by the Trave lab employed quantitative fragmentomic methods to measure E6 interactions at high throughput. However, these have limited applications because the affinities of fragments do not always correspond to their full-length counterparts. Yet, building on these principles of high throughput quantitative interactomics the team developed an assay, the native holdup (nHU), for measuring the affinities of full-length-to-full-length interactions in a pseudo-native state, or the affinity interactome, from total cell extracts. Utilizing this technique, I set out to meet the following research objective:

Measure the affinity interactome of the prototypical HPV16 E6 with nHU.

In having such a direct and tangible research objective, this opened the project to follow up through several pathways such as seeking to understand mechanisms of 16E6 binding, investigate novel partners identified, and attempt to understand how the global interactomic network of 16E6 underpins HPV-related pathogenesis. Here, I present the culmination of my work which has stemmed from my efforts to meet this research objective beginning with the methodologies. The main results of my work follow presented in the format of a prepared manuscript which is to be submitted for publication. Additional experimental efforts and side projects that developed from the original thesis project are adjoined in the subsequent sections. A final analysis discussing the work in its entirety follows with propositions for further study.

Measuring the HPV16 E6 Affinity Interactome: Materials and Methods

The following section outlines the materials and metods (as they appear in the drafted manuscript) followed by the data and results presented in the form of a drafted manuscript. These represent the methodologies and results of the "main project". In following sections additional experimental procedures and results will be presented.

Materials and Methods

Production and purification of biotinylated proteins

Biotinylated 16E6-MBP bait

The 16E6-MBP construct with solubilization supporting mutations was cloned into a pET vector harboring an Avitag to generate an AviTag-His_6-MBP-TEV-16E6 construct bearing a biotin acceptor site[87,88]. The produced 16E6-MBP construct was co-transformed with biotin ligase, BirA (Pet21a-BirA, Addgene, no. 20857) and expressed in *Escherichia coli* BL21 (DE3) cells grown in Terrific Borth (TB) + 0.2% glucose. Upon induction with 500 µM isopropyl-β-D-thiogalactopyranoside (IPTG) cultures were supplemented with 100 µM $ZnSO_4$ and 200 µM Biotin (dissolved in 100% DMSO). Bacterial cultures were expressed at 16°C overnight with shaking at 200 RMP. Subsequent purification of 16E6-MBP was carried out according to the protocol reported in Sidi et al., with some modifications[89]. Despite the mutation of several Cysteine residues to Serine, the remaining Cysteine content of the 16E6-MBP construct was 8.9% (14 residue) leaving the protein highly susceptible to the effects of protein oxidation. To circumvent this, all purification buffers were extensively degassed and argonized. Bacterial pellets were resuspended in lysis buffer (50 mM Tris-HCl pH 6.8, 250 mM NaCl, 5% glycerol, 1 mM TCEP, DNase [0.25 µg/µL], RNase [0.25 µg/µL], lysozyme, and cOmplete EDTA-free protease inhibitor cocktail [1 tablet/50 mL lysis buffer; Roche, Basel, Switzerland]) using a 1:10 biomass-to-buffer ratio and gently mixing at 4°C. Cells were lysed via repeated rounds of sonication using 40% amplitude and 1-minute long pulses of 0.5s on/0.5s off and 1-minute probe rest between cycles. The clarified lysate was then loaded onto an amylose column (NEB, #E8022L) and eluted with 10 mM maltose (Elution buffer: 50 mM Tris-HCl pH 6.8, 250 mM NaCl, 10 mM maltose). The elution was pooled, and soluble aggregates were removed via overnight ultra-centrifugation at 40000 RPM using the SW41 Ti Swinging-Bucket rotor (Beckman Coulter). The elution was recovered from the ultra-centrifuge tubes via surgical pipette leaving a remainder of 0.5-1 mL protein solution in each tube as to not disrupt the pellet or take up any aggregates entering re-suspension. The recovered elution was pooled and concentrated using Vivaspin Turbo15 Centrifugal Concentrators (Sartorius, V515T02) in preparation for size exclusion chromatography. The monomeric form of 16E6-MBP was resolved using a HiLoad s200 16/600 column (Cytiva). The purified 16E6-MBP was then divided into 500 µL aliquots at 20-30 µM concentration, flash-frozen in liquid nitrogen and stored in its final buffer composition (50 mM Tris-HCl pH 6.8, 300 mM NaCl, 1 mM TCEP) at -80°C until further use.

Biotinylated Maltose Binding Protein (MBP) non-binding control

The pET-Avitag vector containing MBP was co-transformed with BirA for expression in BL21(DE3) cells grown in LB + 0.2% glucose. Upon induction with 1 mM IPTG, cultures were supplemented with 200 µM biotin. Cultures were expressed at 16°C overnight with shaking at 200 RPM. Cells were harvested by centrifugation (7000 RPM for 20 minutes at 4°C) and stored at -20°C until further use. Thawed cells were resuspended in Lysis Buffer (50 mM Tris-HCl pH 7.2, 150 mM NaCl, 1% TritonX-100, DNase [0.25 µg/µL], RNase [0.25 µg/µL], lysozyme, and cOmplete EDTA-free protease inhibitor cocktail) using a 1:10 biomass-to-buffer ratio. Cells were lysed via sonication following the protocol above and the lysate was clarified by centrifugation at 40000xg for 25 minutes at 4°C. The supernatant was recovered and loaded onto a 5 mL Ni-Sepharose HisTrap HP column (Cytiva, 17524801). The column was washed with 10 column volumes (CVs) 50 mM Imidazole (Wash buffer: 50 mM Tris-HCl pH 7.2, 150 mM NaCl, 50 mM Imidazole). Due to BirA and MBP both possessing His-tags, both proteins were captured during this purification step and eluted from the resin with 200 mM Imidazole (Elution buffer: 50 mM Tris-HCl pH 7.2, 150 mM NaCl, 200 mM Imidazole). To isolate MBP, the elution was loaded directly onto two chained 5 mL MBPTrap HP columns (Cytiva, 28918779). Proteins were eluted with 10 mM maltose (Elution Buffer: 50 mM Tris-HCl, 150 mM NaCl, 10 mM maltose), incubating the column in the elution buffer for 2 minutes between each 2 mL fraction collected. Fractions containing pure MBP were pooled and divided into 500 µL aliquots at ≥100 µM concentration, flash frozen in liquid nitrogen, and stored at -80°C until further use.

Verification of recombinant protein biotinylation

The biotinylation efficiency of purified 16E6-MBP and MBP was assessed by capturing each respective protein on an excess of Streptavidin resin (Cytiva, 17-5113-01)[90,91]. Using a 1:4 ratio of resin volume to protein volume, the resin was incubated with a 5 µM solution of either 16E6-MBP or MBP for one hour at room temperature with occasional agitation. The unbound fraction, or flow through, of the protein solution was removed by centrifugation and biotinylation efficiency was assessed by one of two methods: (1) SDS-PAGE. Samples of the initial 5 µM protein solution and flow through were prepared by mixing one part protein solution with three parts 4x SDS-PAGE Sample Buffer (120 mM Tris-HCl pH 7.0, 8% SDS, 100 mM DTT, 31.3% glycerol, 0.004% Bromophenol blue, and 1% β-MercaptoEthanol). Resin-elution samples were prepared by mixing the beads with the sample buffer in a 1:4 ratio of dry beads to sample buffer. The reducing agents in conjunction with subsequent heating steps work to disrupt the Streptavidin disulfide bonds, thus freeing monomeric Streptavidin and the protein from the bead surface. All samples were boiled at 95°C for 10 minutes and briefly centrifuged (10 seconds at 12000 RPM) prior to loading equal volumes onto a polyacrylamide gel. The optical density of Coomassie stained protein bands were measured by the Amersham ImageQaunt TM 800 (Cytivia) and densitometry calculations were performed in Fiji[92]. And/or (2)

estimating protein concentration by measuring the ultraviolet absorbance at wavelength 280 nM via NanoDrop. The protein concentration of each sample was calculated using the estimated extinction coefficient[93] for each construct. In both methods the biotinylation efficiency was estimated by comparing the ratio of depleted protein from the flow through to the amount of protein in the initial 5 µM sample. In the case of SDS-PAGE the amount of protein captured on the resin could also be directly estimated by the resin-elution sample (Fig. S1).

Cell culture and preparation of cellular extracts

HeLa cells (Lab of Dr. Walter Schaffner, University of Zurich, Switzerland) were grown in DMEM (1g/L) media completed with 5% fetal calf serum (FCS) and gentamicin (40 µg/mL) and passaged 1:10 every second day. HaCaT cells (Deutsches Krebsforschungszentrum (DFKZ), Heidelberg, Germany) were grown in DMEM (1g/L) media completed with 10% FCS and gentamicin (40 µg/mL) and passaged 1:4 every third day. All cells were kept at 37°C and 5% CO_2.

Whole cell extracts were prepared by seeding cells onto 100 mm plates using their respective dilution factor for passaging. Cells were grown until confluency and harvested on ice by first removing the media and washing cells with 1x phosphate buffered saline (PBS). Ice-cold cell lysate buffer (50 mM HEPES-KOH pH 7.5, 150 mM NaCl, 1% TritonX, 1 [per 10 mL] cOmplete EDTA-free protease inhibitor cocktail tablet, 5 mM TCEP, 10% glycerol) was added to the plate and cells were removed by physical disruption using a cell lifter (SPL Life Sciences, 90040) to vigorously scrape the plate surface. The collected cells were then lysed via sonication with 4 x 20 second cycles of 1.0 second pulses (1s on/1s off) at 40% amplitude with 1 minute probe rest between each cycle. Lysed cells were incubated at 4°C with gentle agitation for 30 minutes and then clarified by centrifugation for 20 minutes at 12000 RPM and 4°C. The supernatant was recovered and measured for protein concentration by Bradford assay[94] (Bio-Rad, Protein assay dye reagent concentrated, cat **#**5000006; Shimadzu, UV1700 PharmaSpec). The final total protein concentration of the lysates were adjusted with cell lysis buffer to 2 mg/mL before aliquoting and flash freezing with liquid nitrogen. The cell lysates were stored at -80°C until further use.

Native holdup (nHU) assay

Streptavidin Sepharose High Performance affinity resin (Cytiva, 17-5113-01) was removed from the manufacturer bottle, washed out with nHU buffer without reducing agents (50 mM Tris-HCl pH 7.5, 300 mM NaCl), aliquoted, and left to settle overnight to determine the resin bed volume. Once the volume of the resin bed was obtained, the resin was centrifuged, and the buffer was removed and replaced with the appropriated volume of better to create a final slurry concentration of 25 µL resin/100 µL slurry. This step was to correct any discrepancies in resin volume/slurry concentration that may have occurred throughout the manufacturing process as well as to limit any additional pipetting error we may incur. The nHU assay[51] of 16E6-

MBP was carried out according to the principles described by Zambo and colleagues. Resin was incubated with 20 CVs of 15-35 µM biotinylated 16E6-MBP. As biotinylated MBP could be produced in great quantities and stored at high concentration without jeopardizing solubility or structural integrity, 10 CVs of ≥100 µM MBP was used to saturate non-binding control resins. The resins were incubated with their respective bait proteins for 1 hour at room temperature with occasional mixing (approximately every 10 minutes). The protein solution was then removed from the resin via brief centrifugation (15 seconds at 6000 RPM, Thermo Fisher Scientific, mySPIN™ 6, 7500061) and replaced with 1 CV of 1 mM biotin in nHU buffer. The resin was incubated for 10 minutes at room temperature with the biotin to saturate any remaining unoccupied binding sites. Following this so-called “depletion” step, the biotin was removed by brief centrifugation and the resin were washed 3-5 times with 10-20 CVs of nHU buffer containing 1 mM TCEP. Washed resins were then resuspended in a 1:4 ratio of resin volume-to-nHU buffer, creating a slurry concentration of 20 µL resin/100 µL nHU buffer. For single point measurements, which would be analyzed by mass spectrometry, equal volumes of the resin saturated to a final concentration of 42 µM 16E6-MBP or MBP non-binding control resin were aliquoted into 500 µL microcentrifuge tubes. For eight-point titration series, which would be analyzed by Western Blot, a serial dilution was created by mixing equal volumes of bait and control saturated resin while keeping a constant total volume for each sample point. Once the appropriate volumes of resins were distributed and the correct bait concentrations were reached, the nHU buffer was removed and replaced by either HeLa or HaCaT extracts [2 mg/mL] in a 1:4 ratio of resin-to-analyte. The resin-cell extract mixes were incubated for two hours at 4°C with gentle agitation until a supposed binding equilibrium was reached. The depleted analyte samples were recovered via centrifugation. To avoid contamination of the sample with resin, the full volume was not recovered, but instead an excess of ≥10 µL was left as a remainder. Samples were prepared by mixing one-part 4x SDS-PAGE sample buffer with three-parts supernatant, boiled at 95°C for five minutes, and briefly centrifuged (10 seconds at 12000 RPM).

Verification of bait concentration in nHU

The bait concentration of 16E6-MBP presented on the resin surface was verified by eluting the protein bait from bait saturated resins that were not exposed to cell extracts. This can easily be performed in parallel with the nHU assay by using an excess of resin during the bait saturation step and setting it aside for bait concentration analysis. The buffer was removed from the resin and the beads were resuspended directly in 4x SDS-PAGE sample buffer in a 1:4 ratio of beads to sample buffer. Samples were then boiled for 10 minutes at 95°C and briefly centrifuged before loading the liquid fraction onto an SDS-PAGE gel alongside calibrating samples composed of known MBP concentrations. The optical density of the Coomassie stained bands were measured by the Amersham ImageQuant™ 800 (Cytivia) and densitometry calculations were performed in Fiji. The final estimated 16E6-MBP concentration was determined to be 42 µM (Fig. S1).

Mass spectroscopy measurements and analysis

Sample digestion and LC tandem mass spec analysis was performed by the IGBMC proteomics platform. All mass spec experiments and data analysis were performed according to the principles and procedures described in Zambo et al[51] with initial data processing performed in Perseus[95].

Western Blotting

Samples from eight-point titration series of nHU were prepared as described above. Equivalent volumes of samples were loaded onto 8 or 10% polyacrylamide gels for separation. Gels were transferred to PVDF membranes using the semi-dry Transblot Turbo Transfer System and Trans-Blot Turbo RTA Transfer kit (Bio-Rad, no. 1704273). Following transfer, membranes were blocked for 1 hour at room temperature in either 5% milk TBS-T or 5% BSA TBS-T depending on the buffer composition of the primary antibody (Table S2). The following antibodies and dilutions were used: APOBEC3B (E9A2G, CST, 1:1000), GAPDH (MAB374, Sigma, 1:5000), IRF3 (1132-1-AP, Proteintech, 1:5000), MSH6 (66172-1-Ig, Proteintech, 1:1000), PXN (SAB4502553, Sigma, 1:1000), PTPN14 (D5T6Y, CST, 1:1000), SCRIB (PA5-54821, Thermo, 1:1000), SNX27 (MA5-27854, Thermo, 1:1000), TP53 (2527S, CST, 1:1000), UBE3A (clone 3E5, Sigma, 1:1000), USP13 (12577S, CST, 1:1000), XIAP (66800-1-Ig, Proteintech, 1:5000). After overnight incubation with the primary antibodies, membranes were washed three times with TBS-T. The washed membranes were then incubated for 1 hour at room temperature with the respective species of HRP-conjugated secondary antibody (AB_2307392 and AB_2313567, Jackson ImmunoResearch) in a 1:10000 dilution prepared in the corresponding primary antibody buffer. The membranes were washed again three times with TBS-T and developed using the SuperSignal™ West (34580, Thermo Fisher Scientific) ECL detection substrate. Following exposure, membranes were immediately imaged using the Amersham ImageQuant™ 800 imaging system. Signal quantification was performed using Fiji and fitting of binding models was performed in QTiPlot (IONDEV SRL, Bucharest, Romania). Membranes that were re-probed as indicated (Fig. S2) underwent stripping with 15% H_2O_2 for antibodies from different species or mild stripping buffer (1.5% glycine, 0.1% SDS, 1% TWEEN-20 pH 2.2; AbCam) for antibodies of the same species for 30 minutes.

LxxLLome peptide library design

A naïve library of 15-mer peptides centered around a LxxLL motif was curated using SLiMSearch[96-98]. Position specific restrictions of the consensus sequence were relaxed in our search to allow different hydrophobic residues in p0, p3, and p4 which are typically regarded as fully conserved Leucine (L) residues. The following Regular Expression (RegEx) was used for our search:

([ILVM]xx[ILVM][ILVM])|([FWY]xx[ILVM][ILVM])|([ILVM]xx[F][ILVM])|([ILVM]xx[ILVM][FY])

Other search parameters included the sequence length, a disorder cutoff of ≤ 0.7, and an overall net charge ≥ -2 in the flanking region. From this search 86 putative LxxLL motifs were identified, 21 of which were detected in our nHU experiments, and 10 were previously described in the literature as binding to 16E6. Biotinylated peptides of these sequences were produced by Søren Østergaard and lyophilized. Peptides were resuspended, using ammonia as needed, in preparation for HU and BLI experiments.

Holdup (HU) assay of LxxLLome peptide library

The fragementomic affinities between the biotinylated LxxLL peptides and the full-length non-biotinylated 16E6 were measured by the HU assay following the principles and protocols described in Gogl et al[50]. The 86 peptides and 10 biotin controls were arranged in a 384 well filter plate such that each peptide or control would be measured in quadruplicate. Plates were prepared using an E1-ClipTip channel pipette (Thermo Fisher, Waltham, Massachusetts), a Flexdrop IV precision reagent dispenser (Perkin Elmer, Waltham, Massachusetts), and a vacuum manifold. Streptavidin resin was distributed into each well, creating a resin bed volume of 5 µL. Resin was saturated with 3 µL of peptide or 1 mM biotin, incubated at room temperature for 15 minutes with shaking at 500 RPM and incubated for an additional 15 minutes at 4°C. Following resin saturation, any unoccupied binding sites were saturated with biotin through a depletion step by adding 4x the bed volume of 1 mM biotin in HU buffer (50 mM Tris-HCl pH 7.5, 300 mM NaCl, 1 mM TCEP). The resin was then washed to remove any free biotin. Non-biotinylated 16E6-MBP was mixed with mCherry and fluorescein, fluorescent controls used for peak intensity normalization, producing a 3.6 µM 16E6-MBP/100 nM mCherry/50 nM fluorescein solution. Twice the bed volume (2 CVs), or 10 µL of this 16E6-MBP/mCherry/fluorescein solution was added to each well followed by rapid filtration via centrifugation (700 RPM for 2 minutes) to remove the unbound fractions. The fluorescent intensity of the flow through was measured by PHERAstar (BMG Labtech, Offenburg, Germany) microplate reader using 485 ± 10 nM (fluorescein), 575 ± 10 nM – 620 ± 10 nM (mCherry), and 295 ± 10 nM – 350 ± 10 nM (Tryptophan fluorescence) band-pass filters. Plates were recycled using 1M NaCl, 2M Urea, and HU Buffer for replicate measurements.

Biolayer Interferometry (BLI)

From the HU experiments, 11 peptides were identified as the top LxxLL binders of 16E6 and selected for cross validation with BLI. Experiments were performed with the Sartorius Octet® R8 system and Octet® High Precision Streptavidin (SAX) Biosensors (18-5117). Sensors were hydrated in BLI buffer (50 mM Tris-HCl pH 7.5, 300 mM NaCl, 1 mM TCEP, 0.05% TWEEN-20) and loaded until saturation with prepared 1:1000 dilutions of LxxLL peptides, acting as the ligands. Following saturation, the sensors were washed for two minutes in buffer and baselined again. An eight-point titration series of the analyte, non-biotinylated 16E6-MBP, was created by serial dilution of a 26 µM 16E6-MBP solution. For each analyte concentration, a 90 second association step followed by a 180 second disassociation step was performed. The last ten seconds of each association-disassociation cycle was taken for kinetic measurements.

Competition assays with the nHU Spike-in

We performed several competition assays with nHU where cellular extracts were supplemented, or spiked-in, with an excess of a competitor to sequester 16E6 binding. In our first series of nHU spike-in experiments, we aimed to determine the oligomerization state of active UBE3A necessary for 16E6 binding and recruitment of TP53. To this end HaCaT extracts were supplemented with 50 µM recombinantly purified UBE3A-MBP in either monomeric or oligomeric states (provided by Jia Wenn, IGBMC). The nHU experiments were performed as described above, however, just prior to the addition of cell extracts recombinant UBE3A-MBP was spiked-in by diluting the extracts 1:1 with a 100 µM UBE3A-MBP solution in cell lysis buffer to give a final concentration of 50 µM UBE3A-MBP and 1 mg/mL total protein. Results of the nHU-UBE3A spike-in assay were analyzed by Western Blot to evaluate the binding activity of UBE3A and TP53.

In our second experiment we sought to identify the potential 16E6 interaction interfaces mediating contacts with newly identified 16E6-binders. With this in mind HaCaT and HeLa extracts were supplemented with 100 µM of either a purified PDZ domain MBP fusion (SCRIB1-MBP) or an LxxLL peptide (UBE3A_402-417). Extracts were prepared in the same manner as described above, mixing either a 200 µM stock of PDZ-MBP or LxxLL peptide in cell lysis buffer in a 1:1 ratio with the lysates to give a final concentration of 100 µM competitor and 1 mg/mL total protein. The 16E6-MBP concentration on the resin was fixed at approximately 10 µM by mixing one part 42 µM 16E6-MBP saturated resin with three parts MBP saturated resin. Results of the nHU assay were analyzed by Western Blot. Using SNX27 and PXN binding as internal controls for PDZ and LxxLL binding respectively, disruptions in binding were assessed using Student's T test in GraphPad Prism version 10.3.1 for Windows (GraphPad Software, Boston, Massachusetts USA).

AlphaFold 3 modelling and analysis of structures

Structural predictions for PTPN14 (Q15678) and USP13 (Q92995) in complex with HPV16 E6 (P03126) were performed with AlphaFold 3[99,100] using default parameters. In cases where there were several isoforms of a partner, the canonical sequence was selected for modelling. Interactions predicted by the AlphaFold models were then analyzed using the Predictomes tool[101].

Measuring the HPV16 E6 Affinity Interactome: Data and results presented in the form of a manuscript

Weimer et al., (2025) Modelling HPV 16E6-host protein complex formation with quantitative interactomics. (*in preparation*; for the purpose of the thesis manuscript main figures and their legends have been embedded directly into the body text of the article manuscript and line numbers have been removed.)

Title: Modelling HPV 16E6-host protein complex formation with quantitative interactomics

Authors: Weimer, Kathleen[1]; Gergo Gogl[1]; Boglarka Zambo[1]; Bastien Morlet[1]; Luc Negroni[1]; Soren Ostegaard[2]; Trave, Gilles[1]

Affiliations:

1 – IGBMC – CBI: Institut de génétique et de biologie moléculaire et cellulaire - Centre de biologie intégrative 1 rue Laurent Fries - BP 10142 - 67404 Illkirch-Graffenstaden, France

2 – Novo Nordisk A/S, Global Research Technologies, Novo Nordisk Research Park, 2760, Maaloev, Denmark

Corresponding author: Weimer, Kathleen

weimerk@igbmc.fr, weimerk030412@gmail.com

Abstract

Sitting proximal to the phenotypic level of molecular organization, interactomes encompass all molecular interactions of a cell, creating a complex network that is deterministic of cellular function and regulation. Furthermore, associations between genetic mutations and distinct interactomic profiles producing distinguishable phenotypes indicates that interactomes serve as a link between genotype and phenotype putting them at a critical crossroads holding key insights into pathogenesis. Current practices in interactomics, however, sacrifices quantitation and molecular detail to gain higher throughput. This trade-off has led to lack of consensus between studies and methodologies limiting interpretations that can be drawn from interactomes. To remedy this, we take a high throughput quantitative approach, measuring the interaction strength, or affinities, of entire interactomic networks. Here, we mapped the full-length affinity interactome of 16E6, a virally encoded oncoprotein from HPV type 16 which has been implicated in carcinogenesis. Using the native holdup method coupled to mass spectroscopy approximately 4500 affinities were measured directly from total extracts of HPV immortalized HeLa and HPV free keratinocyte HaCat cells. Identifying nearly 100 16E6 binders, their affinities were combined with the quantified proteome thus modeling binary cellular complex formation, or as we call it, the *complexome*. This represents a data informed molecular fishing approach allowing selection of individual interactions that are potential key targets at different stages of transformation for further characterization while still considering the whole network targeted by the viral component for cellular rewiring. With this we can now view interactomes at a depth and breadth as never seen before.

(Manuscript in preparation)

Keywords: HPV, E6, PPIs, interactomics, oncoprotein, carcinogenesis

Introduction

As we have entered the multi-omics era, we have gained the ability to integrate multiple layers of biological data to construct more complete pictures of cellular systems. Furthermore, the potential for more stratified, precision, and even personalized medicine grows as we begin to uncover missing links connecting genotype and phenotype[1,2]. However, understanding deterministic factors of phenotypes is confounded by the increasing complexity of cellular diversity occurring over each level of molecular organization[3]. Until we begin to understand the biochemical and biophysical mechanisms governing cellular systems our understanding of pathogenesis will remain incomplete. To this end, we propose a quantitative interactomics approach.

Sitting proximal to the phenotypic level of molecular organization, interactomes represent a complex network of molecular interactions that are deterministic of cellular function and regulation serving as a direct link to phenotypes. Interactomic studies have demonstrated that mutations drastically altering interactomes are often associated with genes and that various mutations within the same gene can produce different interactomic profiles that are associated with distinguishable phenotypes[4]. However, to date, the field of interactomics is dominated by qualitative methods that lack molecular and mechanistic details often resulting in little overlap among studies and methodologies[5,6]. Here, we measure the strength of 16E6-host interactions, combining them with the quantified proteome thus modelling cellular complex formation. This results in a data informed molecular fishing approach that allows us to select individual interactions that may represent critical targets at different stages of transformation for further characterization as well as consider the whole network targeted by the viral component for cellular rewiring.

Human Papillomaviruses (HPVs) are small, non-enveloped, icosahedral viruses that exhibit tropism for epithelial cells[7]. Currently accounting for the most widespread sexually transmitted infection worldwide[8], clinical presentations range from asymptomatic and subclinical, to benign lesions such as papillomas (warts) and condylomas, to anogenital cancers[9]. Of the over 400 HPV isolates characterized, some 200 of these have been identified to infect humans, and of these 15 are currently recognized as high-risk HPVs (hr-HPVs) classified by their oncogenic potential[10-12]. Taken together, these hr-HPVs are responsible for >99.7% of cervical cancers with HPV16 implicated in >50% of these cases[13-15]. At the crux of carcinogenesis sits the three HPV encoded oncoproteins E5, E6, and E7 working in concert to create the optimal cellular environment for viral replication through prolongation of proliferation and disruption of differentiation[16]. The E6 protein has been established as a driver of cancer progression and malignancy targeting a multitude of cellular host factors[17]. These attributes of E6 combined with the high

genotypic and phenotypic diversity of HPVs has garnered much interest for interactomic studies[18-21], particularly for the prototypical 16E6[22-25]. Despite the fact that these studies are limited to qualitative interactomic approaches, they serve as good benchmarks for validating our methods as we map the full-length quantitative affinity interactome of 16E6.

Materials and Methods

Production and purification of biotinylated proteins

Biotinylated 16E6-MBP bait

The 16E6-MBP construct with solubilization supporting mutations was cloned into a pET vector harboring an Avitag to generate an AviTag-His_6-MBP-TEV-16E6 construct bearing a biotin acceptor site[26,27]. The produced 16E6-MBP construct was co-transformed with biotin ligase, BirA (Pet21a-BirA, Addgene, no. 20857) and expressed in *Escherichia coli* BL21 (DE3) cells grown in Terrific Borth (TB) + 0.2% glucose. Upon induction with 500 µM isopropyl-β-D-thiogalactopyranoside (IPTG) cultures were supplemented with 100 µM $ZnSO_4$ and 200 µM Biotin (dissolved in 100% DMSO). Bacterial cultures were expressed at 16°C overnight with shaking at 200 RMP. Subsequent purification of 16E6-MBP was carried out according to the protocol reported in Sidi et al., with some modifications[28]. Despite the mutation of several Cysteine residues to Serine, the remaining Cysteine content of the 16E6-MBP construct was 8.9% (14 residue) leaving the protein highly susceptible to the effects of protein oxidation. To circumvent this, all purification buffers were extensively degassed and argonized. Bacterial pellets were resuspended in lysis buffer (50 mM Tris-HCl pH 6.8, 250 mM NaCl, 5% glycerol, 1 mM TCEP, DNase [0.25 µg/µL], RNase [0.25 µg/µL], lysozyme, and cOmplete EDTA-free protease inhibitor cocktail [1 tablet/50 mL lysis buffer; Roche, Basel, Switzerland]) using a 1:10 biomass-to-buffer ratio and gently mixing at 4°C. Cells were lysed via repeated rounds of sonication using 40% amplitude and 1-minute long pulses of 0.5s on/0.5s off and 1-minute probe rest between cycles. The clarified lysate was then loaded onto an amylose column (NEB, #E8022L) and eluted with 10 mM maltose (Elution buffer: 50 mM Tris-HCl pH 6.8, 250 mM NaCl, 10 mM maltose). The elution was pooled, and soluble aggregates were removed via overnight ultra-centrifugation at 40000 RPM using the SW41 Ti Swinging-Bucket rotor (Beckman Coulter). The elution was recovered from the ultra-centrifuge tubes via surgical pipette leaving a remainder of 0.5-1 mL protein solution in each tube as to not disrupt the pellet or take up any aggregates entering re-suspension. The recovered elution was pooled and concentrated using Vivaspin Turbo15 Centrifugal Concentrators (Sartorius, V515T02) in preparation for size exclusion chromatography. The monomeric form of 16E6-MBP was resolved using a HiLoad s200 16/600 column (Cytiva). The purified 16E6-MBP was then divided into 500 µL aliquots at 20-30 µM concentration, flash-frozen in liquid nitrogen and stored in its final buffer composition (50 mM Tris-HCl pH 6.8, 300 mM NaCl, 1 mM TCEP) at -80°C until further use.

(Manuscript in preparation)

Biotinylated Maltose Binding Protein (MBP) non-binding control

The pET-Avitag vector containing MBP was co-transformed with BirA for expression in BL21(DE3) cells grown in LB + 0.2% glucose. Upon induction with 1 mM IPTG, cultures were supplemented with 200 µM biotin. Cultures were expressed at 16°C overnight with shaking at 200 RPM. Cells were harvested by centrifugation (7000 RPM for 20 minutes at 4°C) and stored at -20°C until further use. Thawed cells were resuspended in Lysis Buffer (50 mM Tris-HCl pH 7.2, 150 mM NaCl, 1% TritonX-100, DNase [0.25 µg/µL], RNase [0.25 µg/µL], lysozyme, and cOmplete EDTA-free protease inhibitor cocktail) using a 1:10 biomass-to-buffer ratio. Cells were lysed via sonication following the protocol above and the lysate was clarified by centrifugation at 40000xg for 25 minutes at 4°C. The supernatant was recovered and loaded onto a 5 mL Ni-Sepharose HisTrap HP column (Cytiva, 17524801). The column was washed with 10 column volumes (CVs) 50 mM Imidazole (Wash buffer: 50 mM Tris-HCl pH 7.2, 150 mM NaCl, 50 mM Imidazole). Due to BirA and MBP both possessing His-tags, both proteins were captured during this purification step and eluted from the resin with 200 mM Imidazole (Elution buffer: 50 mM Tris-HCl pH 7.2, 150 mM NaCl, 200 mM Imidazole). To isolate MBP, the elution was loaded directly onto two chained 5 mL MBPTrap HP columns (Cytiva, 28918779). Proteins were eluted with 10 mM maltose (Elution Buffer: 50 mM Tris-HCl, 150 mM NaCl, 10 mM maltose), incubating the column in the elution buffer for 2 minutes between each 2 mL fraction collected. Fractions containing pure MBP were pooled and divided into 500 µL aliquots at ≥100 µM concentration, flash frozen in liquid nitrogen, and stored at -80°C until further use.

Verification of recombinant protein biotinylation

The biotinylation efficiency of purified 16E6-MBP and MBP was assessed by capturing each respective protein on an excess of Streptavidin resin (Cytiva, 17-5113-01)[29,30]. Using a 1:4 ratio of resin volume to protein volume, the resin was incubated with a 5 µM solution of either 16E6-MBP or MBP for one hour at room temperature with occasional agitation. The unbound fraction, or flow through, of the protein solution was removed by centrifugation and biotinylation efficiency was assessed by one of two methods: (1) SDS-PAGE. Samples of the initial 5 µM protein solution and flow through were prepared by mixing one part protein solution with three parts 4x SDS-PAGE Sample Buffer (120 mM Tris-HCl pH 7.0, 8% SDS, 100 mM DTT, 31.3% glycerol, 0.004% Bromophenol blue, and 1% β-MercaptoEthanol). Resin-elution samples were prepared by mixing the beads with the sample buffer in a 1:4 ratio of dry beads to sample buffer. The reducing agents in conjunction with subsequent heating steps work to disrupt the Streptavidin disulfide bonds, thus freeing monomeric Streptavidin and the protein from the bead surface. All samples were boiled at 95°C for 10 minutes and briefly centrifuged (10 seconds at 12000

RPM) prior to loading equal volumes onto a polyacrylamide gel. The optical density of Coomassie stained protein bands were measured by the Amersham ImageQaunt TM 800 (Cytivia) and densitometry calculations were performed in Fiji[31]. And/or (2) estimating protein concentration by measuring the ultraviolet absorbance at wavelength 280 nM via NanoDrop. The protein concentration of each sample was calculated using the estimated extinction coefficient[32] for each construct. In both methods the biotinylation efficiency was estimated by comparing the ratio of depleted protein from the flow through to the amount of protein in the initial 5 µM sample. In the case of SDS-PAGE the amount of protein captured on the resin could also be directly estimated by the resin-elution sample (Fig. S1).

Cell culture and preparation of cellular extracts

HeLa cells (Lab of Dr. Walter Schaffner, University of Zurich, Switzerland) were grown in DMEM (1g/L) media completed with 5% fetal calf serum (FCS) and gentamicin (40 µg/mL) and passaged 1:10 every second day. HaCaT cells (Deutsches Krebsforschungszentrum (DFKZ), Heidelberg, Germany) were grown in DMEM (1g/L) media completed with 10% FCS and gentamicin (40 µg/mL) and passaged 1:4 every third day. All cells were kept at 37°C and 5% CO_2.

Whole cell extracts were prepared by seeding cells onto 100 mm plates using their respective dilution factor for passaging. Cells were grown until confluency and harvested on ice by first removing the media and washing cells with 1x phosphate buffered saline (PBS). Ice-cold cell lysate buffer (50 mM HEPES-KOH pH 7.5, 150 mM NaCl, 1% TritonX, 1 [per 10 mL] cOmplete EDTA-free protease inhibitor cocktail tablet, 5 mM TCEP, 10% glycerol) was added to the plate and cells were removed by physical disruption using a cell lifter (SPL Life Sciences, 90040) to vigorously scrape the plate surface. The collected cells were then lysed via sonication with 4 x 20 second cycles of 1.0 second pulses (1s on/1s off) at 40% amplitude with 1 minute probe rest between each cycle. Lysed cells were incubated at 4°C with gentle agitation for 30 minutes and then clarified by centrifugation for 20 minutes at 12000 RPM and 4°C. The supernatant was recovered and measured for protein concentration by Bradford assay[33] (Bio-Rad, Protein assay dye reagent concentrated, cat **#**5000006; Shimadzu, UV1700 PharmaSpec). The final total protein concentration of the lysates were adjusted with cell lysis buffer to 2 mg/mL before aliquoting and flash freezing with liquid nitrogen. The cell lysates were stored at -80°C until further use.

Native holdup (nHU) assay

Streptavidin Sepharose High Performance affinity resin (Cytiva, 17-5113-01) was removed from the manufacturer bottle, washed out with nHU buffer without reducing agents (50 mM Tris-HCl pH 7.5, 300 mM NaCl), aliquoted, and left to settle overnight to determine the resin bed volume. Once the volume of the resin bed was obtained, the resin was centrifuged, and the buffer was removed and replaced with the appropriated volume of better to create a final slurry concentration of 25 µL

resin/100 μL slurry. This step was to correct any discrepancies in resin volume/slurry concentration that may have occurred throughout the manufacturing process as well as to limit any additional pipetting error we may incur. The nHU assay[34] of 16E6-MBP was carried out according to the principles described by Zambo and colleagues. Resin was incubated with 20 CVs of 15-35 μM biotinylated 16E6-MBP. As biotinylated MBP could be produced in great quantities and stored at high concentration without jeopardizing solubility or structural integrity, 10 CVs of ≥100 μM MBP was used to saturate non-binding control resins. The resins were incubated with their respective bait proteins for 1 hour at room temperature with occasional mixing (approximately every 10 minutes). The protein solution was then removed from the resin via brief centrifugation (15 seconds at 6000 RPM, Thermo Fisher Scientific, mySPIN™ 6, 7500061) and replaced with 1 CV of 1 mM biotin in nHU buffer. The resin was incubated for 10 minutes at room temperature with the biotin to saturate any remaining unoccupied binding sites. Following this so-called “depletion” step, the biotin was removed by brief centrifugation and the resin were washed 3-5 times with 10-20 CVs of nHU buffer containing 1 mM TCEP. Washed resins were then resuspended in a 1:4 ratio of resin volume-to-nHU buffer, creating a slurry concentration of 20 μL resin/100 μL nHU buffer. For single point measurements, which would be analyzed by mass spectrometry, equal volumes of the resin saturated to a final concentration of 42 μM 16E6-MBP or MBP non-binding control resin were aliquoted into 500 μL microcentrifuge tubes. For eight-point titration series, which would be analyzed by Western Blot, a serial dilution was created by mixing equal volumes of bait and control saturated resin while keeping a constant total volume for each sample point. Once the appropriate volumes of resins were distributed and the correct bait concentrations were reached, the nHU buffer was removed and replaced by either HeLa or HaCaT extracts [2 mg/mL] in a 1:4 ratio of resin-to-analyte. The resin-cell extract mixes were incubated for two hours at 4°C with gentle agitation until a supposed binding equilibrium was reached. The depleted analyte samples were recovered via centrifugation. To avoid contamination of the sample with resin, the full volume was not recovered, but instead an excess of ≥10 μL was left as a remainder. Samples were prepared by mixing one-part 4x SDS-PAGE sample buffer with three-parts supernatant, boiled at 95°C for five minutes, and briefly centrifuged (10 seconds at 12000 RPM).

Verification of bait concentration in nHU

The bait concentration of 16E6-MBP presented on the resin surface was verified by eluting the protein bait from bait saturated resins that were not exposed to cell extracts. This can easily be performed in parallel with the nHU assay by using an excess of resin during the bait saturation step and setting it aside for bait concentration analysis. The buffer was removed from the resin and the beads were resuspended directly in 4x SDS-PAGE sample buffer in a 1:4 ratio of beads to sample buffer. Samples were then boiled for 10 minutes at 95°C and briefly centrifuged before loading the liquid fraction onto an SDS-PAGE gel alongside

calibrating samples composed of known MBP concentrations. The optical density of the Coomassie stained bands were measured by the Amersham ImageQuant™ 800 (Cytivia) and densitometry calculations were performed in Fiji. The final estimated 16E6-MBP concentration was determined to be 42 µM (Fig. S1).

Mass spectroscopy measurements and analysis

Sample digestion and LC tandem mass spec analysis was performed by the IGBMC proteomics platform. All mass spec experiments and data analysis were performed according to the principles and procedures described in Zambo et al[34] with initial data processing performed in Perseus[35].

Western Blotting

Samples from eight-point titration series of nHU were prepared as described above. Equivalent volumes of samples were loaded onto 8 or 10% polyacrylamide gels for separation. Gels were transferred to PVDF membranes using the semi-dry Transblot Turbo Transfer System and Trans-Blot Turbo RTA Transfer kit (Bio-Rad, no. 1704273). Following transfer, membranes were blocked for 1 hour at room temperature in either 5% milk TBS-T or 5% BSA TBS-T depending on the buffer composition of the primary antibody (Table S2). The following antibodies and dilutions were used: APOBEC3B (E9A2G, CST, 1:1000), GAPDH (MAB374, Sigma, 1:5000), IRF3 (1132-1-AP, Proteintech, 1:5000), MSH6 (66172-1-Ig, Proteintech, 1:1000), PXN (SAB4502553, Sigma, 1:1000), PTPN14 (D5T6Y, CST, 1:1000), SCRIB (PA5-54821, Thermo, 1:1000), SNX27 (MA5-27854, Thermo, 1:1000), TP53 (2527S, CST, 1:1000), UBE3A (clone 3E5, Sigma, 1:1000), USP13 (12577S, CST, 1:1000), XIAP (66800-1-Ig, Proteintech, 1:5000). After overnight incubation with the primary antibodies, membranes were washed three times with TBS-T. The washed membranes were then incubated for 1 hour at room temperature with the respective species of HRP-conjugated secondary antibody (AB_2307392 and AB_2313567, Jackson ImmunoResearch) in a 1:10000 dilution prepared in the corresponding primary antibody buffer. The membranes were washed again three times with TBS-T and developed using the SuperSignal™ West (34580, Thermo Fisher Scientific) ECL detection substrate. Following exposure, membranes were immediately imaged using the Amersham ImageQuant™ 800 imaging system. Signal quantification was performed using Fiji and fitting of binding models was performed in QTiPlot (IONDEV SRL, Bucharest, Romania). Membranes that were re-probed as indicated (Fig. S2) underwent stripping with 15% H_2O_2 for antibodies from different species or mild stripping buffer (1.5% glycine, 0.1% SDS, 1% TWEEN-20 pH 2.2; AbCam) for antibodies of the same species for 30 minutes.

LxxLLome peptide library design

A naïve library of 15-mer peptides centered around a LxxLL motif was curated using SLiMSearch[36-38]. Position specific restrictions of the consensus sequence were relaxed in our search to allow different hydrophobic residues in p0, p3, and p4 which

are typically regarded as fully conserved Leucine (L) residues. The following Regular Expression (RegEx) was used for our search:

([ILVM]xx[ILVM][ILVM])|([FWY]xx[ILVM][ILVM])|([ILVM]xx[F][ILVM])|([ILVM]xx[ILVM][FY])

Other search parameters included the sequence length, a disorder cutoff of ≤ 0.7, and an overall net charge ≥ -2 in the flanking region. From this search 86 putative LxxLL motifs were identified, 21 of which were detected in our nHU experiments, and 10 were previously described in the literature as binding to 16E6. Biotinylated peptides of these sequences were produced by Søren Østergaard and lyophilized. Peptides were resuspended, using ammonia as needed, in preparation for HU and BLI experiments.

Holdup (HU) assay of LxxLLome peptide library

The fragementomic affinities between the biotinylated LxxLL peptides and the full-length non-biotinylated 16E6 were measured by the HU assay following the principles and protocols described in Gogl et al[39]. The 86 peptides and 10 biotin controls were arranged in a 384 well filter plate such that each peptide or control would be measured in quadruplicate. Plates were prepared using an E1-ClipTip channel pipette (ThermoFisher, Waltham, Massachusetts), a Flexdrop IV precision reagent dispenser (Perkin Elmer, Waltham, Massachusetts), and a vacuum manifold. Streptavidin resin was distributed into each well, creating a resin bed volume of 5 µL. Resin was saturated with 3 µL of peptide or 1 mM biotin, incubated at room temperature for 15 minutes with shaking at 500 RPM and incubated for an additional 15 minutes at 4°C. Following resin saturation, any unoccupied binding sites were saturated with biotin through a depletion step by adding 4x the bed volume of 1 mM biotin in HU buffer (50 mM Tris-HCl pH 7.5, 300 mM NaCl, 1 mM TCEP). The resin was then washed to remove any free biotin. Non-biotinylated 16E6-MBP was mixed with mCherry and fluorescein, fluorescent controls used for peak intensity normalization, producing a 3.6 µM 16E6-MBP/100 nM mCherry/50 nM fluorescein solution. Twice the bed volume (2 CVs), or 10 µL of this 16E6-MBP/mCherry/fluorescein solution was added to each well followed by rapid filtration via centrifugation (700 RPM for 2 minutes) to remove the unbound fractions. The fluorescent intensity of the flow through was measured by PHERAstar (BMG Labtech, Offenburg, Germany) microplate reader using 485 ± 10 nM (fluorescein), 575 ± 10 nM – 620 ± 10 nM (mCherry), and 295 ± 10 nM – 350 ± 10 nM (Tryptophan fluorescence) band-pass filters. Plates were recycled using 1M NaCl, 2M Urea, and HU Buffer for replicate measurements.

Biolayer Interferometry (BLI)

From the HU experiments, 11 peptides were identified as the top LxxLL binders of 16E6 and selected for cross validation with BLI. Experiments were performed with the Sartorius Octet® R8 system and Octet® High Precision Streptavidin (SAX) Biosensors (18-5117). Sensors were hydrated in BLI buffer (50

mM Tris-HCl pH 7.5, 300 mM NaCl, 1 mM TCEP, 0.05% TWEEN-20) and loaded until saturation with prepared 1:1000 dilutions of LxxLL peptides, acting as the ligands. Following saturation, the sensors were washed for two minutes in buffer and baselined again. An eight-point titration series of the analyte, non-biotinylated 16E6-MBP, was created by serial dilution of a 26 μM 16E6-MBP solution. For each analyte concentration, a 90 second association step followed by a 180 second disassociation step was performed. The last ten seconds of each association-disassociation cycle was taken for kinetic measurements.

Competition assays with the nHU Spike-in

We performed several competition assays with nHU where cellular extracts were supplemented, or spiked-in, with an excess of a competitor to sequester 16E6 binding. In our first series of nHU spike-in experiments, we aimed to determine the oligomerization state of active UBE3A necessary for 16E6 binding and recruitment of TP53. To this end HaCaT extracts were supplemented with 50 μM recombinantly purified UBE3A-MBP in either monomeric or oligomeric states (provided by Jia Wenn, IGBMC). The nHU experiments were performed as described above, however, just prior to the addition of cell extracts recombinant UBE3A-MBP was spiked-in by diluting the extracts 1:1 with a 100 μM UBE3A-MBP solution in cell lysis buffer to give a final concentration of 50 μM UBE3A-MBP and 1 mg/mL total protein. Results of the nHU-UBE3A spike-in assay were analyzed by Western Blot to evaluate the binding activity of UBE3A and TP53.

In our second experiment we sought to identify the potential 16E6 interaction interfaces mediating contacts with newly identified 16E6-binders. With this in mind HaCaT and HeLa extracts were supplemented with 100 μM of either a purified PDZ domain MBP fusion (SCRIB1-MBP) or an LxxLL peptide (UBE3A_402-417). Extracts were prepared in the same manner as described above, mixing either a 200 μM stock of PDZ-MBP or LxxLL peptide in cell lysis buffer in a 1:1 ratio with the lysates to give a final concentration of 100 μM competitor and 1 mg/mL total protein. The 16E6-MBP concentration on the resin was fixed at approximately 10 μM by mixing one part 42 μM 16E6-MBP saturated resin with three parts MBP saturated resin. Results of the nHU assay were analyzed by Western Blot. Using SNX27 and PXN binding as internal controls for PDZ and LxxLL binding respectively, disruptions in binding were assessed using Student's T test in GraphPad Prism version 10.3.1 for Windows (GraphPad Software, Boston, Massachusetts USA).

AlphaFold 3 modelling and analysis of structures

Structural predictions for PTPN14 (Q15678) and USP13 (Q92995) in complex with HPV16 E6 (P03126) were performed with AlphaFold 3[40,41] using default parameters. In cases where there were several isoforms of a partner, the canonical sequence was selected for modelling. Interactions predicted by the AlphaFold models were then analyzed using the Predictomes tool[42].

(Manuscript in preparation)

Results

Mapping the 16E6 affinity interactome with nHU-MS

To map the 16E6-host affinity interactome, full length biotinylated 16E6-MBP was produced and used as bait in nHU assays against HPV (-) HaCaT and HPV (+) HeLa cell lines (Fig. 1A). Coupling with mass spectrometry analysis, approximately 4500 apparent affinities (pK_{app}) were measured between the two cell lines (Fig. 1B). For the sake of our analysis, we draw a significance threshold at 2σ, considering any point above this threshold as having quantifiable affinities within the detection limits of our assay resulting in the identification of approximately 100 16E6-binders (Figure 1B; Table 1). Of these 16E6-binders many have been previously described in the literature giving us a re-discovery rate of 23%. We have characterized the remaining identified 16E6-binders as novel partners since this appears to be their first mention. Additionally, there are many other 16E6 interactions described by the literature involving preys from our study that appear as interactions below our significance threshold (Fig. 1B). Meta-analysis was performed comparing our nHU results to other 16E6 interaction studies to determine the number of True Positive (TP) 16E6-binders and assess the accuracy of various interactomic methods (Fig. S3). Within our nHU assays, the pK_{apps} measured show good agreement among the two cell lines (Fig. 1B). Next, we sought to analyze these identified 16E6-binders in their cellular context to further understand the biological relevance of these interactions.

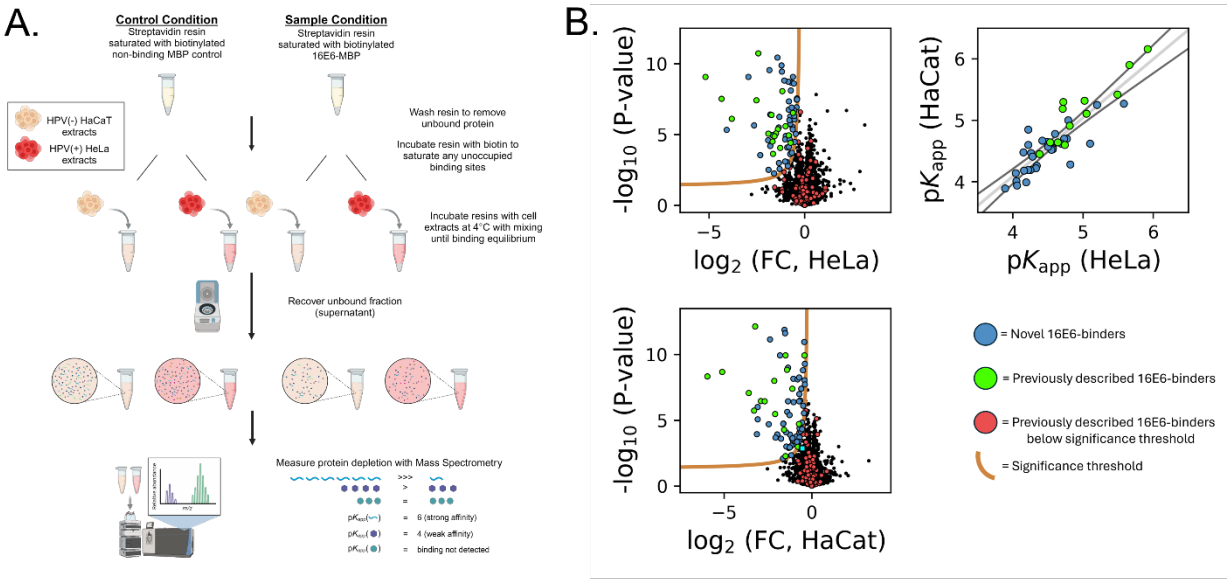


Figure 1 Measuring the 16E6 affinity interactome with nHU.

Using biotinylated 16E6 as bait nHU was performed against HPV(+) and HPV(-) cell lines. Mass spec analysis of duplicate samples in technical triplicate was performed for a total of approximately 4500 apparent affinity (pK_{app}) measurements. A significance threshold at 2 sigma is indicated with a tan line. 100 partners above the significance threshold were identified as previously characterized 16E6 partners in green or novel 16E6 partners in blue. Previously characterized 16E6 partners appearing below the significance threshold are labelled in red. The correlation (1:1) of apparent affinities (pK_{app}) among the two cell lines is shown along the diagonal dashed line.

(Manuscript in preparation)

Table 1 100 16E6-binders identified by nHU-MS.

Accession	Description	Gene Name	pKapp HeLa	pKapp HaCaT	pKapp average
Q53GT1	Kelch-like protein 22	KLHL22	N/D	4.87	4.87
Q8WVV4	Protein POF1B	POF1B	N/D	5.33	5.33
Q15678	Tyrosine-protein phosphatase non-receptor type 14	PTPN14	N/D	5.88	5.88
Q13835	Plakophilin-1	PKP1	N/D	4.81	4.81
P19012	Keratin, type I cytoskeletal 15	KRT15	N/D	4.54	4.54
Q13751	Laminin subunit beta-3	LAMB3	N/D	4.76	4.76
Q12955	Ankyrin-3	ANK3	N/D	5.04	5.04
P19484	Transcription factor EB	TFEB	N/D	5.06	5.06
Q92817	Envoplakin	EVPL	5.00	N/D	5.00
Q8WV41	Sorting nexin-33	SNX33	N/D	4.94	4.94
O43251	RNA binding protein fox-1 homolog 2	RBFOX2	5.01	N/D	5.01
Q9UDY2	Tight junction protein ZO-2	TJP2	5.37	5.23	5.30
Q9Y446	Plakophilin-3	PKP3	N/D	4.51	4.51
Q5T2T1	MAGUK p55 subfamily member 7	MPP7	N/D	5.30	5.30
Q12959	Disks large homolog 1	DLG1	5.34	5.93	5.63
O14907	Tax1-binding protein 3	TAX1BP3	5.67	5.74	5.71
Q9UKK3	Protein mono-ADP-ribosyltransferase PARP4	PARP4	N/D	5.02	5.02
Q86U70	LIM domain-binding protein 1	LDB1	N/D	4.99	4.99
Q96NY8	Nectin-4	NECTIN4	4.87	N/D	4.87
P23142	Fibulin-1	FBLN1	5.33	5.81	5.57
P00734	Prothrombin	F2	4.81	4.61	4.71
O14545	TRAF-type zinc finger domain-containing protein 1	TRAFD1	N/D	4.67	4.67
Q14257	Reticulocalbin-2	RCN2	N/D	5.52	5.52
Q13033	Striatin-3	STRN3	5.43	N/D	5.43
Q9NYL9	Tropomodulin-3	TMOD3	N/D	4.54	4.54
P49023	Paxillin	PXN	5.26	5.27	5.26
Q9UPY3	Endoribonuclease Dicer	DICER1	N/D	5.19	5.19
Q7Z5J4	Retinoic acid-induced protein 1	RAI1	4.96	4.88	4.92

(Manuscript in preparation)

Q8IVL5	Prolyl 3-hydroxylase 2	P3H2	5.60	N/D	5.60
P07996	Thrombospondin-1	THBS1	5.35	5.32	5.34
P63167	Dynein light chain 1, cytoplasmic	DYNLL1	N/D	4.75	4.75
Q8IVD9	NudC domain-containing protein 3	NUDCD3	4.88	4.82	4.85
Q9Y4K1	Beta/gamma crystallin domain-containing protein 1	CRYBG1	4.79	N/D	4.79
P57764	Gasdermin-D	GSDMD	5.31	5.24	5.27
Q9HD26	Golgi-associated PDZ and coiled-coil motif-containing protein	GOPC	6.28	6.53	6.41
Q9NUP9	Protein lin-7 homolog C	LIN7C	5.44	5.54	5.49
Q8NI22	Multiple coagulation factor deficiency protein 2	MCFD2	N/D	5.00	5.00
Q07954	Prolow-density lipoprotein receptor-related protein 1	LRP1	N/D	4.80	4.80
Q86WB0	Nuclear-interacting partner of ALK	ZC3HC1	N/D	4.67	4.67
P68371	Tubulin beta-4B chain	TUBB4B	4.51	N/D	4.51
Q9UH17	DNA dC->dU-editing enzyme APOBEC-3B	APOBEC3B	5.17	N/D	5.17
Q15654	Thyroid receptor-interacting protein 6	TRIP6	5.00	5.07	5.04
Q6WKZ4	Rab11 family-interacting protein 1	RAB11FIP1	4.67	4.63	4.65
Q15057	Arf-GAP with coiled-coil, ANK repeat and PH domain-containing protein 2	ACAP2	5.19	5.18	5.18
Q8WYA6	Beta-catenin-like protein 1	CTNNBL1	4.84	4.81	4.83
Q07666	KH domain-containing, RNA-binding, signal transduction-associated protein 1	KHDRBS1	4.53	N/D	4.53
O95905	Protein ecdysoneless homolog	ECD	4.89	N/D	4.89
Q13425	Beta-2-syntrophin	SNTB2	5.64	5.95	5.79
Q96L92	Sorting nexin-27	SNX27	6.55	6.78	6.67
Q96IF1	LIM domain-containing protein ajuba	AJUBA	5.82	5.88	5.85
Q86W42	THO complex subunit 6 homolog	THOC6	4.77	5.10	4.94
Q8IWX8	Calcium homeostasis endoplasmic reticulum protein	CHERP	4.74	N/D	4.74
Q9H4L5	Oxysterol-binding protein-related protein 3	OSBPL3	4.84	5.22	5.03
P53621	Coatomer subunit alpha	COPA	4.52	4.51	4.51
Q12805	EGF-containing fibulin-like extracellular matrix protein 1	EFEMP1	4.95	4.84	4.90
P52597	Heterogeneous nuclear ribonucleoprotein F	HNRNPF	5.18	5.14	5.16
Q9ULD2	Microtubule-associated tumor suppressor 1	MTUS1	N/D	5.65	5.65
Q9Y697	Cysteine desulfurase, mitochondrial	NFS1	N/D	4.88	4.88
Q9NYU2	UDP-glucose:glycoprotein glucosyltransferase 1	UGGT1	5.42	5.63	5.52
Q96SZ5	2-aminoethanethiol dioxygenase	ADO	4.66	N/D	4.66

(Manuscript in preparation)

P49750	YLP motif-containing protein 1	YLPM1	5.20	5.15	5.18
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	HNRNPA2B1	N/D	4.55	4.55
Q9H6S0	3'-5' RNA helicase YTHDC2	YTHDC2	4.86	N/D	4.86
P55072	Transitional endoplasmic reticulum ATPase	VCP	5.11	5.28	5.19
P13807	Glycogen [starch] synthase, muscle	GYS1	6.20	5.89	6.05
Q13242	Serine/arginine-rich splicing factor 9	SRSF9	5.04	5.29	5.17
Q9NX57	Ras-related protein Rab-20	RAB20	4.90	N/D	4.90
Q93008	Probable ubiquitin carboxyl-terminal hydrolase FAF-X	USP9X	4.78	4.80	4.79
Q9Y4E8	Ubiquitin carboxyl-terminal hydrolase 15	USP15	5.16	5.26	5.21
Q14160	Protein scribble homolog	SCRIB	6.11	6.05	6.08
Q9UGP4	LIM domain-containing protein 1	LIMD1	4.88	5.09	4.99
Q5BKZ1	DBIRD complex subunit ZNF326	ZNF326	N/D	4.83	4.83
P43243	Matrin-3	MATR3	N/D	4.72	4.72
Q00587	Cdc42 effector protein 1	CDC42EP1	4.89	N/D	4.89
Q9NNW5	WD repeat-containing protein 6	WDR6	N/D	4.84	4.84
P07437	Tubulin beta chain	TUBB	4.51	N/D	4.51
O00571	ATP-dependent RNA helicase DDX3X	DDX3X	5.12	5.19	5.16
Q8IXT5	RNA-binding protein 12B	RBM12B	5.72	5.25	5.48
P31942	Heterogeneous nuclear ribonucleoprotein H3	HNRNPH3	5.23	5.40	5.31
Q9Y450	HBS1-like protein	HBS1L	N/D	5.08	5.08
Q96RT1	Erbin	ERBIN	5.10	N/D	5.10
P31689	DnaJ homolog subfamily A member 1	DNAJA1	4.68	4.56	4.62
P46976	Glycogenin-1	GYG1	5.40	5.32	5.36
P52701	DNA mismatch repair protein Msh6	MSH6	4.73	N/D	4.73
P31943	Heterogeneous nuclear ribonucleoprotein H	HNRNPH1	4.91	5.03	4.97
P11498	Pyruvate carboxylase, mitochondrial	PC	4.66	4.77	4.71
Q68D10	Protein SPT2 homolog	SPTY2D1	N/D	5.51	5.51
P01130	Low-density lipoprotein receptor	LDLR	4.73	N/D	4.73
Q9Y5X1	Sorting nexin-9	SNX9	N/D	5.38	5.38
Q8IWR0	Zinc finger CCCH domain-containing protein 7A	ZC3H7A	5.44	4.90	5.17
Q9H078	Caseinolytic peptidase B protein homolog	CLPB	4.84	5.47	5.16
Q8NEY8	Periphilin-1	PPHLN1	N/D	4.82	4.82

Q96MU7	YTH domain-containing protein 1	YTHDC1	N/D	5.15	5.15
Q9Y2H6	Fibronectin type-III domain-containing protein 3A	FNDC3A	4.81	N/D	4.81
P78332	RNA-binding protein 6	RBM6	5.40	N/D	5.40
Q13884	Beta-1-syntrophin	SNTB1	5.30	N/D	5.30
P11532	Dystrophin	DMD	5.22	N/D	5.22
Q9H3G5	Probable serine carboxypeptidase CPVL	CPVL	4.73	N/D	4.73
Q92995	Ubiquitin carboxyl-terminal hydrolase 13	USP13	5.01	N/D	5.01
P98170	E3 ubiquitin-protein ligase XIAP	XIAP	4.82	N/D	4.82

Determining the probability of 16E6-binder binary complex formation in cells

Combining the cellular abundance of each 16E6-binder, provided by our proteomic data, with the calculated strength of the interaction (pK_{app}), we determined the probability of independent 16E6 interactions occurring within their cellular context. In considering the cellular availability of each 16E6-binder when ranking interactions by affinity, we can deduce the likelihood of binary 16E6-binder complex formation or, as we've termed it, the *complexome* (Fig. 2B). When comparing the ratio of 16E6-binder complex formation between the two cell lines, we observe that high affinity 16E6-binders with down-regulated expression in HeLa result in the formation of lower abundances of complex. As such, the converse is also true, and these same 16E6-binders form higher abundances of complex in HaCaT where their expression is up-regulated. To investigate the importance of complex abundance in pathogenic hi-jacking by 16E6 we selected a sample of 16E6-binders across the complexomic spectrum for in-depth analysis. The eleven 16E6-binders selected: SCRIB, SNX27, UBE3A, PTPN14, USP13, IRF3, PXN, XIAP, APOBEC3B, TP53, and MSH6 also represent well characterized and novel 16E6-binders with interactions mediated by the various established 16E6 binding interfaces (Fig. 2A) or interfaces that are unknown. It's important to note that it remains undetermined if all interactions are direct and in the case of TP53 this represents a direct interaction facilitated by a secondary binding event. Finally, members of this selection also perform an array of cellular functions. Starting this inquiry, we aimed to obtain mechanistic insights into the binding events of each selected 16E6-binder using nHU coupled to Western Blot (nHU-WB).

(Manuscript in preparation)

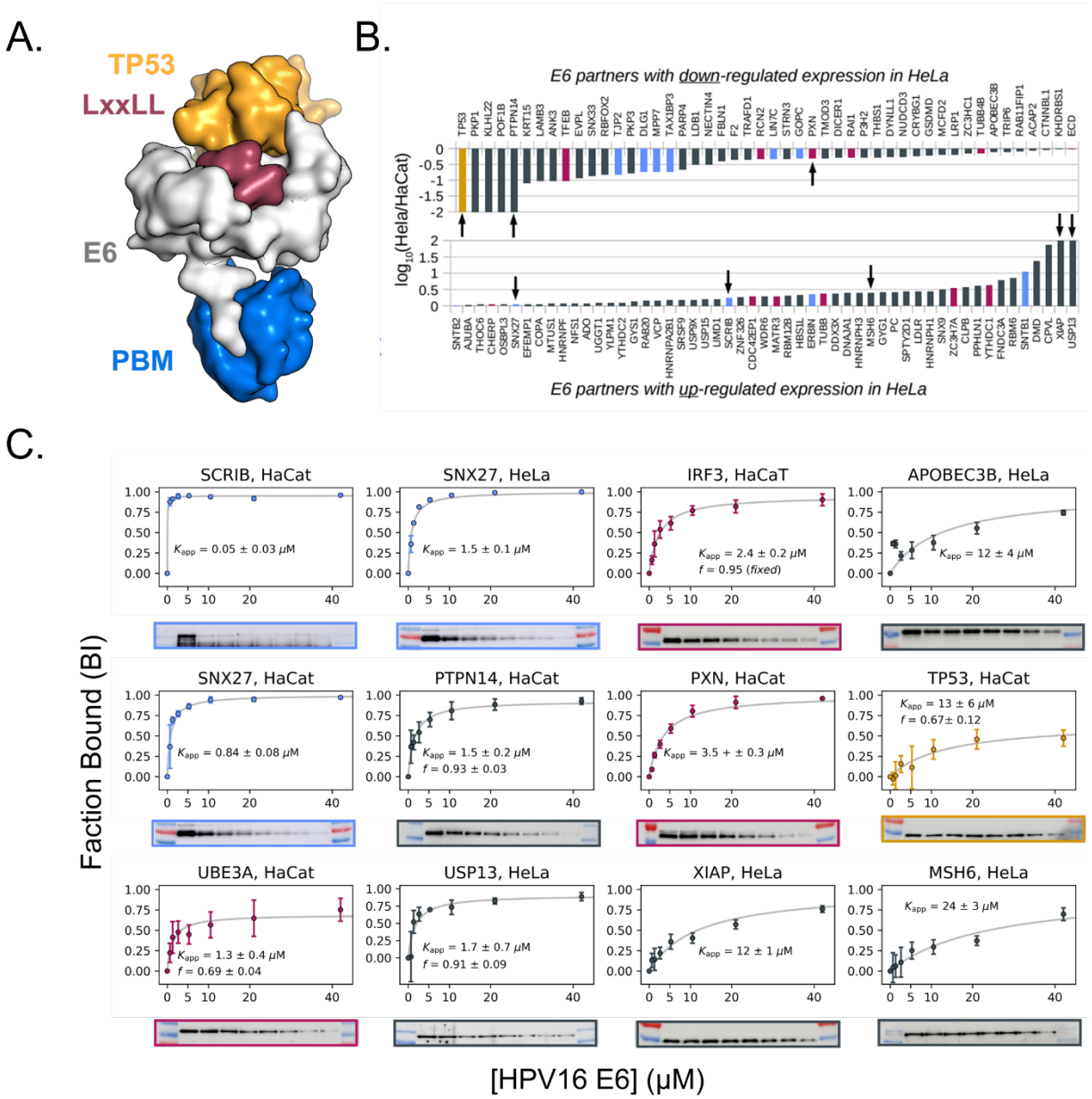


Figure 2 16E6 binding depends on cellular context.

(A) The binding interfaces of 16E6 are highlighted as follows – TP53 binding interface displayed in gold, LxxLL binding pocket displayed in raspberry, the C-terminal PDZ binding motif (PBM) displayed in blue, and remaining 16E6 surface is displayed in grey. (B) 16E6-partner binary complex formation is predicted based on the cellular abundance of each partner, as provided by proteomics, and their apparent affinity (pK_{app}). Following the same color scheme, partners are highlighted according to the 16E6 binding interfaced mediating with unknown interactions are shown in grey. Arrows indicate partners selected across the spectrum of regulation for further characterization with either known or unknown binding mechanisms. UBE3A and IRF3 are not displayed as they were below the detection threshold of MS measurements. (C) nHU-WB of selected 16E6-binders. Eight-point titrations series of the selected 16E6-binders were performed and analyzed via nHU-WB. The term f refers to the protein activity and are only displayed to indicate activities <1.

(Manuscript in preparation)

Investigating interaction mechanisms of 16E6-binders with nHU-WB

Using an eight-point titration series of the 16E6-MBP bait saturated resin, nHU assays were performed and analyzed via Western Blot. Binding models were produced for the eleven 16E6-binders (Fig. 2C) assuming the simplest binding modes[43]. Among the strongest interactions are well characterized 16E6-binders SCRIB, SNX27, and UBE3A as well as novel 16E6-binders PTPN14 and USP13. The PDZ substrates, SCRIB and SNX27, show distinct binding curves despite sharing a binding interface thus demonstrating how we can observe mechanistic differences between avidity mediated interactions, as is the case for SCRIB, versus affinity mediated interactions, such as SNX27 using nHU-WB (Fig. 2C).

Furthermore, clear partial activities are observed for UBE3A and TP53 binding, meaning a portion of these 16E6-binders are not actually accessible for binding in our cellular extracts. In attempts to recover the binding activity of UBE3A, and by proxy TP53, cellular extracts were supplemented with an excess of recombinantly produced UBE3A-MBP in different oligomeric states in a nHU-UBE3A spike-in assay. Despite the ability of the multimeric UBE3A to ameliorate binding to 16E6, TP53 activities were never fully recovered (Fig. 3). Delving further into the mechanisms that mediate 16E6 binding, we revisited each of the known binding interfaces (Fig. 1A) for in-depth study, beginning with the LxxLL binding pocket.

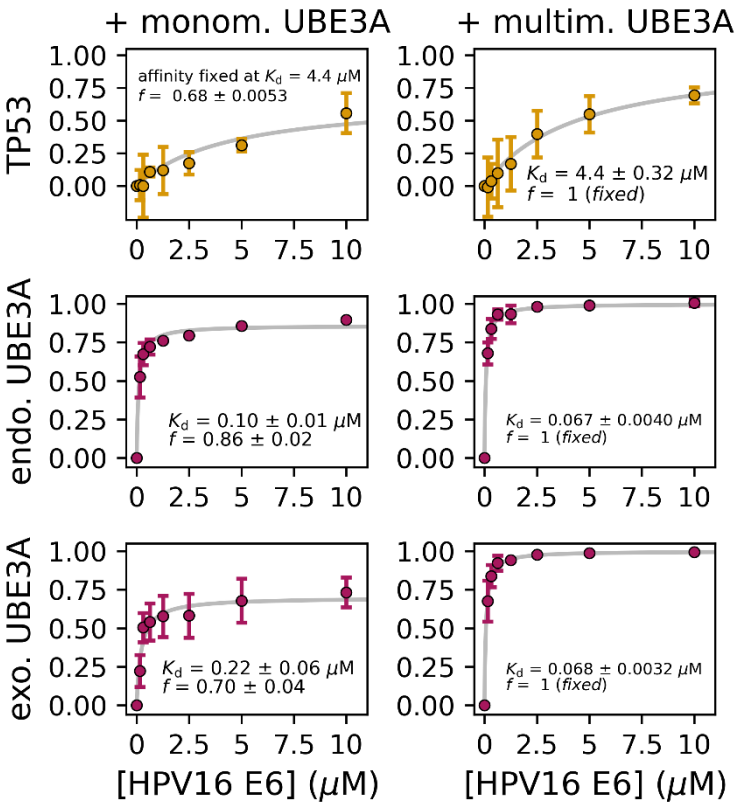


Figure 3 UBE3A activity depends on oligomeric state.

(A) Column one shows nHU experiments performed with cellular extracts supplemented with 50 μM monomeric UBE3A and in column two cellular extracts were supplemented with the same concentration of multimeric UBE3A. Shown in gold is TP53 levels, raspberry endogenous UBE3A levels, and in shaded raspberry exogenous UBE3A-MBP levels. The supplementation of multimeric UBE3A rescues UBE3A binding to 16E6 but does not fully rescue binding to TP53.

Investigating interactions mediated by the 16E6 LxxLL binding pocket

Returning to fragmentomic approaches, we sought to refine and identify key features of the LxxLL motif contributing to 16E6 binding. Using a naïve library of 86 LxxLL motifs discovered through nHU, the literature, and a SLiMFinder search, 28 motifs were identified above our binding threshold in holdup (HU) assays (Fig. 4A). Comparing sequence logo diagrams composed of the 28 binding motifs weighted by affinity to that of the entire library without weighting, we observe the emergence of several sequence features that likely possess a role in binding (Fig. 4C). There is clear conservation of Leucine in the p0 position with more relaxed requirements in p3 and p4, with both positions showing a strong preference for L but also allowing hydrophobic substitutions F/I and M, respectively. Positions p1 and p2, corresponding to the motif's 'xx', favor acidic residues with a distinct preference for D in p1 and E in p2. Directly following the region corresponding to LxxLL, there is a strong propensity for G in p5. Finally, downstream in p8 and p9 we see a subtle re-emergence of a once described -VL motif[44].

(Manuscript in preparation)

The top 11 binders from our HU experiments were cross validated using BLI (Fig. 4D). Further still, when comparing the fragmentomic affinity measurements to the full-length affinity measurements we see good agreement between the affinities for LxxLL-16E6 interactions (Fig. 4B). We turn next to the C-terminal of 16E6 to explore the nuances of PDZ mediated interactions.

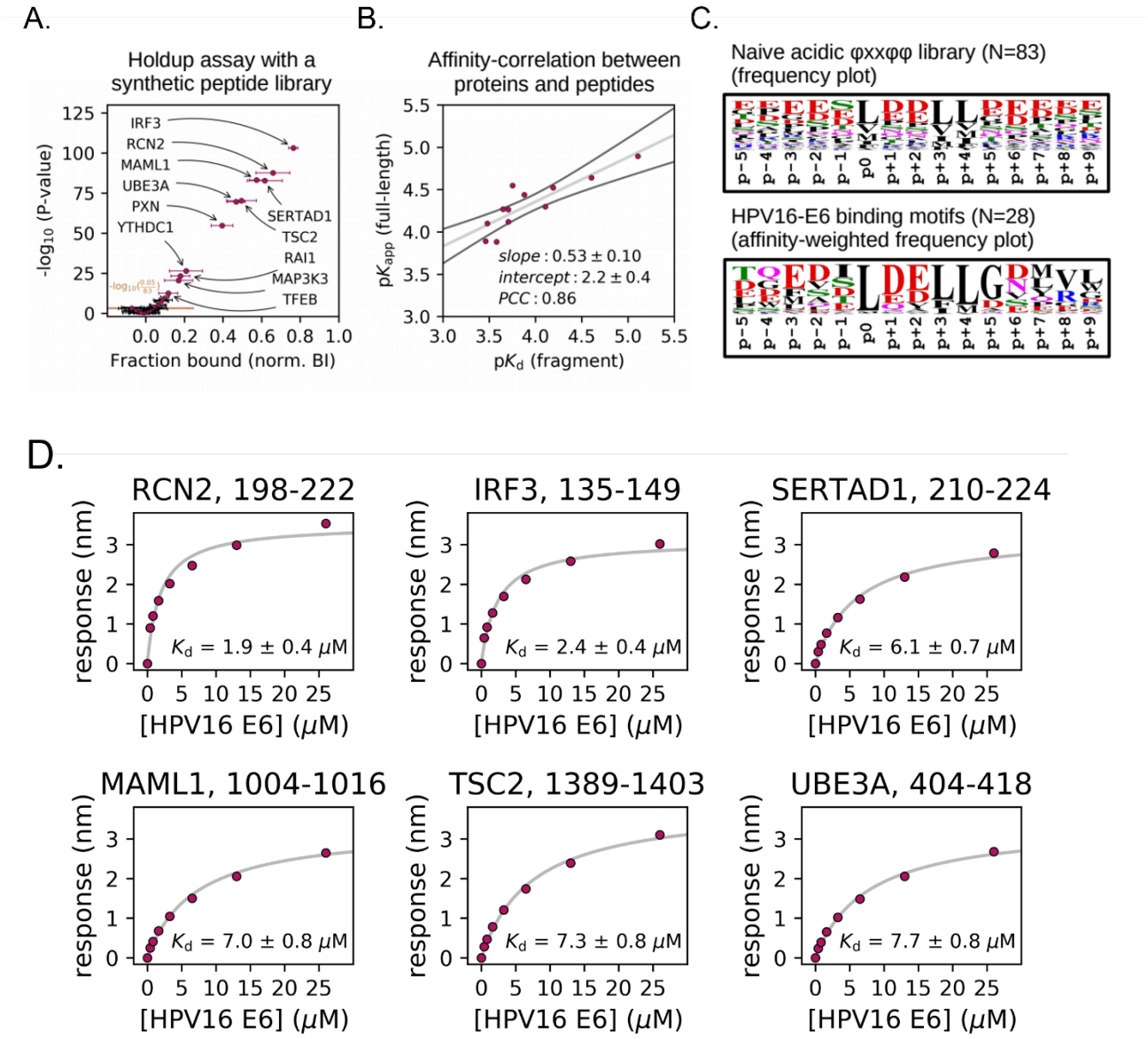


Figure 4 Measuring the 16E6 LxxLLome.

(A) Top binders from HU assays against a peptide library composed of 96 LxxLL motifs. (B) Correlation between the measured apparent affinities of full-length interactions by nHU and measured affinities of full length-peptide interactions by HU show that 16E6-LxxLL interactions are strictly domain-motif mediated interactions with almost no additional contacts. (C) On top is an unweighted frequency plot of the LxxLL peptides included in the panel. On bottom is the frequence plot of the LxxLL peptides that bound 16E6 in the HU assay weighted by affinity (28 of 96 peptides).

(Manuscript in preparation)

Investigating interactions mediated by the 16E6 C-terminal

Performing nHU with a biotinylated peptide corresponding to the 16E6 C-terminal PDZ binding motif (PBM) as bait, we identified 16E6 binders mediated through this interface. The assay showed interactions with PDZ containing partners are retained with the peptide bait (Fig. 5A) yet binding to novel 16E6-binders PTPN14 and USP13 are lost (Fig. 5B). As is the case when using full-length 16E6-MBP as bait, there is good agreement between the affinities measured with the peptide bait in both cell lines. However, comparison of the affinities from full-length interactions to peptide interactions demonstrates less correlation (Fig. 5C). Following our in-depth analysis of the already well-characterized 16E6 interfaces, we began to investigate novel 16E6-binders to determine if interactions occur through already established modes or if there are new points of contact.

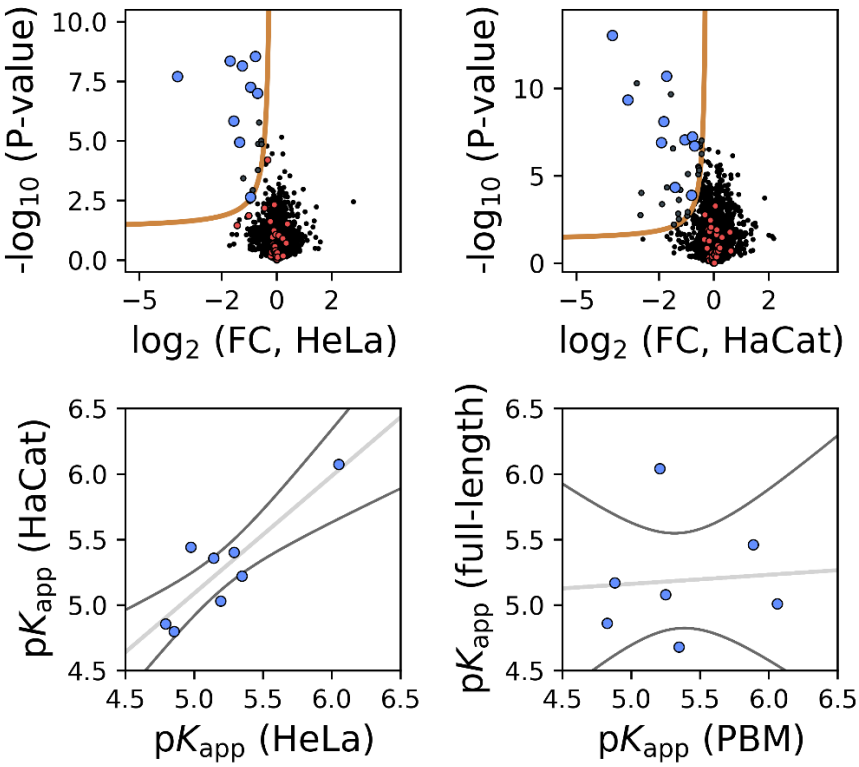


Figure 5 Measuring the 16E6 PDZome.

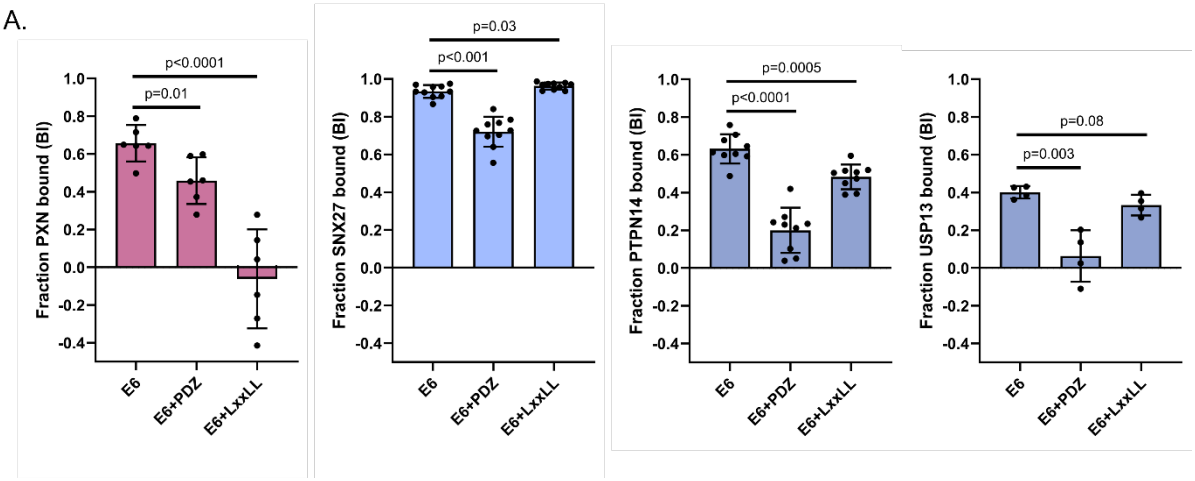
(A) nHU was performed using a biotinylated peptide corresponding to the 16E6 C-terminal PBM as bait against HaCaT and HeLa cell lines. Affinities measured between the two cell lines show good agreement. The nHU affinities using full-length 16E6 and the 16E6 PMB peptide as bait in nHU were compared and show little agreement.

(Manuscript in preparation)

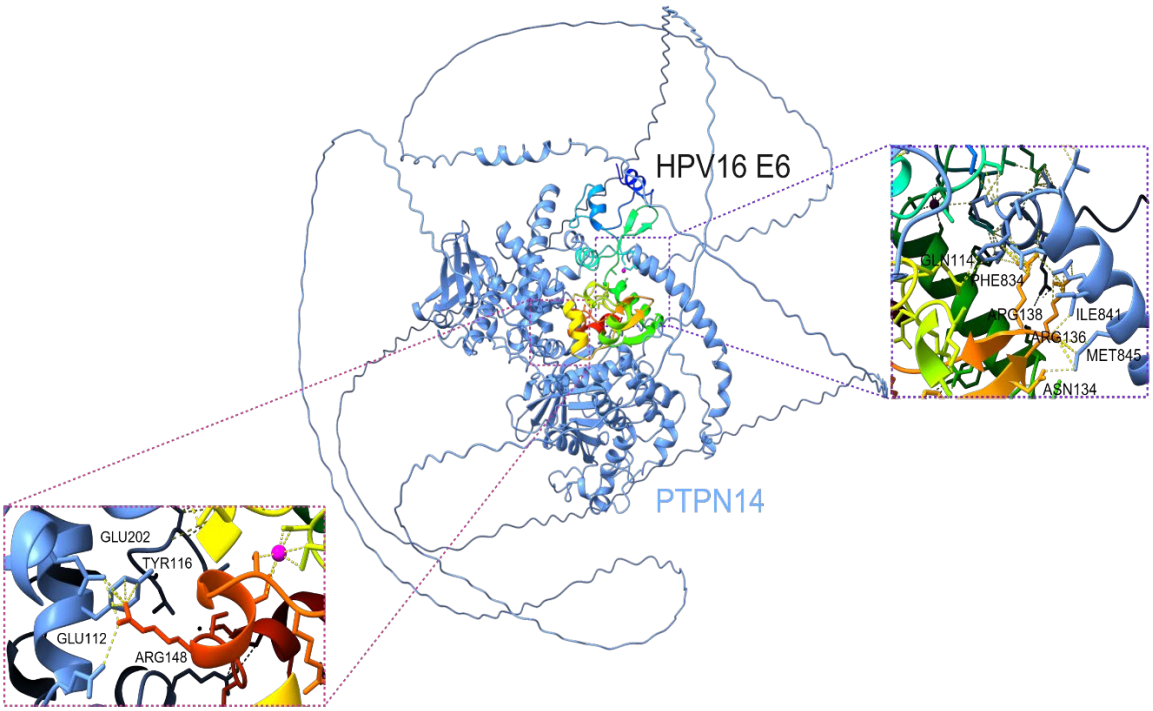
Uncovering 16E6 interfaces mediating interactions with novel 16E6-binders

Still uncertain as to which interfaces may mediate interactions with novel 16E6-binders, we designed a nHU competition assy. Similar to the nHU-UBE3A spike-in experiment, cellular extracts were supplemented with an excess of either a purified PDZ domain fused to MBP (SCRIB-MBP) or an LxxLL peptide acting in this case as competitors to saturate the corresponding 16E6 binding sites. Using SNX27 and PXN as internal standards representing PDZ and LxxLL binding, respectively, binding to novel partners PTPN14 and USP13 were monitored by nHU-WB. In both cases PTPN14 and USP13 exhibit a statistically significant decrease in binding when the C-terminal PBM is sequestered by our PDZ domain, whereas there is no significant change in binding in response to the LxxLL pocket being blocked (Fig. 6A).

A.



B.



(Manuscript in preparation)

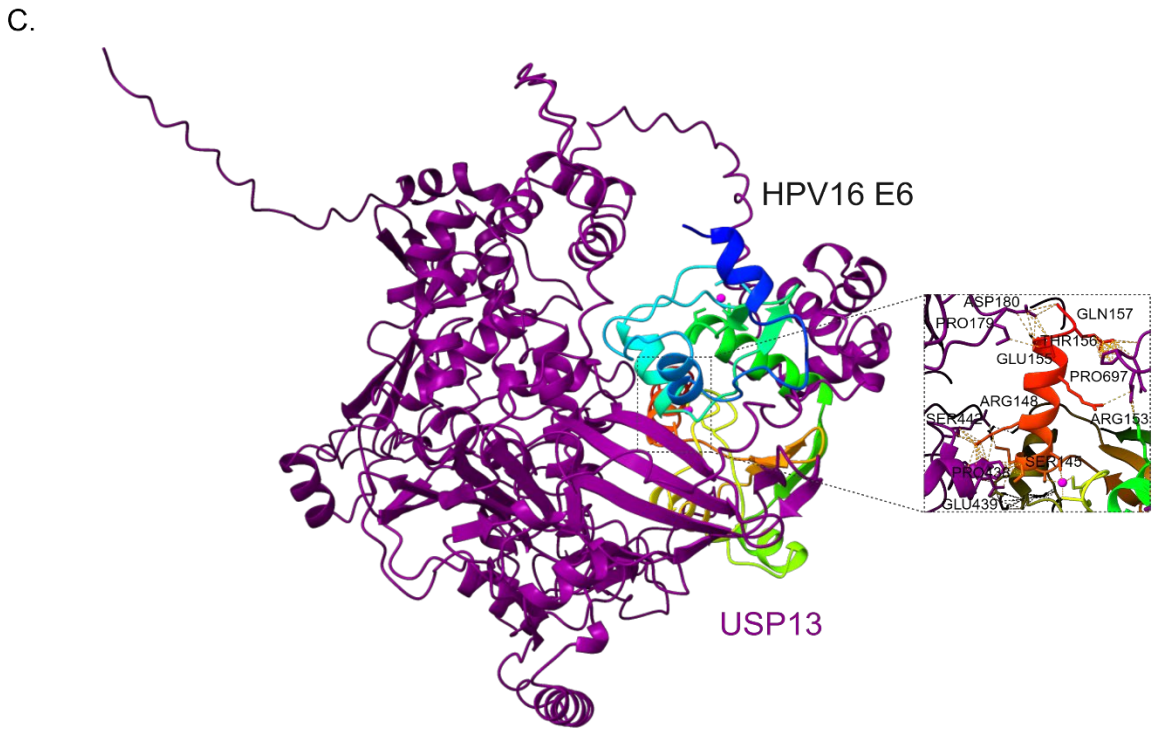


Figure 6 Investigating novel interfaces mediating 16E6-binder interactions.

(A) nHU spike-in experiments were performed using a PDZ-MBP domain and LxxLL peptide to block interaction sites of 16E6. Inhibition of binding with 16E6-binders PTPN14 and USP13 were monitored via nHU-WB and compared to internal standards SNX27 and PXN. Both 16E6-binders show a statistically significant decrease in binding in conditions where the C-terminal PBM is blocked. Statistical analysis was performed in Prism using Student's T Test. (B) AlphaFold3 Prediction of 16E6-PTPN14 complex. PTPN14 is displayed in light blue and demonstrates a folded core with surrounding regions of high disorder. 16E6 is displayed in rainbow coloring. Boxed in dashed hot pink is a predicted C-terminal contact involving residue ARG148 of 16E6. Boxed in dashed purple is another set of predicted contacts involving ASN134, ARG136, and ARG138 of 16E6. (C) AlphaFold3 Prediction of 16E6-USP13 complex. USP13 is displayed in purple and 16E6 is displayed in rainbow coloring. Boxed in dashed black is a set of predicted contacts involving a C-terminal helix from 16E6.

(Manuscript in preparation)

AlphaFold modelling 16E6 interactions with PTPN14 and USP13

Exploring the possibility that novel 16E6-binder interactions may be mediated directly through new interfaces we modelled the 16E6-PTPN14 and 16E6-USP13 interactions with AlphaFold3[40,41](Fig. 6B&C). Both models showed potential interactions with the C-terminal region of 16E6. Local confidence scores varied in both models[45], with the 16E6-USP13 model showing globally high and very high confidence scores (pIDTT >70) with specific regions of low confidence (pIDTT <70). These areas of low confidence correspond to modelled helices which could be indicative that they are more disorder or transient by nature[46]. Conversely the 16E6-PTPN14 model displayed significant regions of low confidence, although this is expected since the protein is highly disordered. The core of the model, however, showed high and very high local confidence scores (pIDTT >70). Both models have a pTM score of approximately 0.5, indicating there is likely similarity with the true structure[47] but conversely they also possess low ipTM scores under 0.6 which is indicative of low confidence in the predictions. The predicted alignment error (PAE) plots for both models indicated the presence of a motif at the 16E6 C-terminal that was likely mediating contact, and this corresponded with the Predictomes analysis which showed several interactions with 16E6 C-terminal residues possessing scores 5<PAE<10.

Discussion

Quantitative interactomics allows us to interpret interactomes in new ways

Determining a standard metric for comparison across orthogonal methods and creating continuity among datasets would increase our confidence in interactomes and how we interpret them. Here, we demonstrate that regardless of cell type the affinities of 16E6-binder interactions remain consistent, therefore the affinities we measure are an intrinsic property of cellular systems (Fig. 1B). It is well known that extrinsic properties, such as protein deprotonation and conformational states, can impact binding thus masking intrinsic affinities during measurements[48,49]. This has necessitated the use of low-throughput methods with precise experimental conditions in traditional biophysical techniques. However, our findings show that despite the heterogeneous and dynamic environment of cell there is a high degree of correlation between affinities measured across cell types. This indicates that interactions in their cellular environment are a tightly regulated process governed by biophysical principles. Further illustrating this point, when comparing the pK_{app} of the 16E6-IRF3 interaction measured from nHU (2.4 ± 0.2 μM) to the pK_D obtained by BLI (2.4 ± 0.4) we see the results of nHU are highly comparable to those of traditional methods. While it is important to note that this example is a comparison of full-length and fragmentomic affinities, it has been previously established that, with the exception of UBE3A, LxxLL-16E6 interactions are typically strictly domain-motif mediated interactions[50] as we also reaffirm in our study (Fig. 4B).

Moreover, in measuring the strength of interactions we uncover biochemical and biophysical insights that allow us to evaluate interactomes in new ways. Unlike other high throughput interactomic methods, such as AP-MS and BioID, nHU disrupts compartmentalization and uses an excess of bait thus capturing all possible physical interactions regardless of biological relevance. This results in the capture of stable and transient, direct and indirect, and interactions over a range of strengths, hence a more complete picture of the interactome compared to methods that only detect stable interactions (AP-MS, IP, Co-fractionation) or direct interactions (Y2H, FRET, BiFC). Discounting any of these categories of interactions limits the perspective of an interactome as either stable or transient complex formation necessitates the realization of function for nearly every protein[51]. Additionally, nHU datasets tend to record more moderate strength interactions as compared to other methods which are enriched with high affinity interactions[5] again providing a more complete picture.

As aforementioned, some interactions captured by nHU may not be pertinent to the cellular context and while this does not dismiss their value in aiding the understanding of biophysical mechanisms, it is important to determine critical cellular interactions to further understanding of cellular systems and pathogenesis. To this end a standard among interactomic studies is to use GO and pathway enrichment analysis[52,53] to determine key interactions of interest. However, this can be subjective and with this approach it can be difficult to predict the downstream impacts of specific interactions or pathways without additional molecular details. This is possible to determine through further functional studies, but this can be expensive in both time and resources and given high overlap between protein functional networks, finding the correct model system can be challenging. To circumvent this we took a novel approach ranking interactions with 16E6-binders by strength, a process known as affinity profiling[54], and combined these profiles with the quantified proteome to determine the probability of cellular complex formation. Our complexomic calculations revealed a clear decrease in the predicted abundance of complexes formed between 16E6 and TP53 in HeLa cells compared to that of HaCaT (Fig. 2B). This mirrors the pathological phenomenon of TP53 degradation in HPV infected cells orchestrated by the 16E6-UBE3A complex[55,56], a critical event in HPV carcinogenesis[57,58]. As HeLa cells endogenously express HPV18 E6[59], TP53 is degraded[60] through this process and thus not available for binding in our assays. When comparing then to HPV (-) HaCaT cells, shifts in the complexome reflect underlying proteomic shifts introduced by HPV therefore giving insight into interactions which may be important during initial stages of transformation.

To further extend these complexomic principles in a data informed molecular fishing approach, novel 16E6-binders PTPN14, USP13, and XIAP were selected for further analysis as they undergo exceptional complexomic shifts. In the case of PTPN14 and USP13 we were able to determine that binding is likely mediated through the 16E6 C-terminus (Fig. 6A), although uncertainty remains as to whether

this is a PDZ mediated process or representative of novel interaction sites as is suggested by our AlphaFold models (Fig. 6B&C). Further biochemical and functional characterization is necessary to evaluate these 16E6-binders, but they hold potential for interesting roles in HPV pathogenesis. The protein tyrosine phosphatase PTPN14 is already implicated in HPV pathologies as a major target of the other HPV oncoprotein E7[61] and it is known that the HPV oncoproteins often exhibit an overlap in their molecular targets. Studies have reported that the deubiquitinase USP13 is upregulated specifically in HPV positive cervical cancers[62] and while USP13 is considered less functionally active compared to other USP family members[63], likely due to a defective zinc binding domain[64], we hypothesize a model where interaction with 16E6 can stimulate USP13 activity through contributing its two zinc fingers. Additionally, 16E6 has a long standing role interacting with the Ubiquitin-Proteosome system[65,66] supporting the identification of USP13 and ubiquitin ligase, XIAP, as players in HPV induced transformation. It's often been suggested that other ubiquitin ligases may be responsible for the degradation of certain PDZ substrates[67,68] and it's been reported that cells expressing 16E6 exhibit upregulation of XIAP[69] implying a potential role for XIAP. While it is clear that 16E6-binders that experience great shifts in complex formation between the two cell lines may resemble key targets at the various stages of pathogenesis, we suppose that complexes that remain stable in levels of formation throughout the transformation progression may hold key roles in maintaining the viral lifecycle and represent interactions that could be targeted therapeutically during any stage of infection[70].

Refining what we know about 16E6 binding

Qualitative interactomic studies of 16E6 have provided comprehensive lists describing interactors, however they lack mechanistic insights even for highly studied interactions such as UBE3A-E6. From our nHU spike-in assays we observed the binding activity of UBE3A is influenced by oligomeric state (Fig.3) agreeing with reported functional[71] and current structural studies[72,73]. Additionally, our results indicate that the dominant form of UBE3A in the cell is monomeric. Taken together, this suggests potential biophysical mechanisms of cellular regulation. Single protein molecules dynamically transition between an array of conformations circulating the average structure, overcoming energy barriers and satisfying free energy constraints as governed by a multidimensional free energy landscape[74]. However, this is only a singular component contributing to the greater energy landscape of cellular networks[75-77]. As protein conformations can be further influenced by complex formation, it stands to reason that there are biophysical mechanisms dictating how and when these interactions occur to maintain a free energy equilibrium just as there are biological mechanisms for maintaining homeostasis. Understanding these biophysical regulatory elements could be exploited for therapeutic strategies, for example, if UBE3A oligomerization is unique to its interaction with E6 its disruption could be targeted to prevent the effects of TP53 degradation while limiting cytotoxic effects.

(Manuscript in preparation)

The LxxLLome

Complementing our analysis with fragmentomic methods, we've shown that interactions with 16E6 facilitated by the LxxLL binding pocket are strictly domain-motif mediated (Fig. 4B). Therefore, there is little or no contribution from secondary binding sites and contacts underpinning binding are restricted to the motif and the regions directly flanking it. The exact consensus motif required for interaction with the 16E6 LxxLL pocket has long been a subject of debate. The original consensus sequence described was an ELLG motif which required the absence of basic residues[44] eventually evolving into an LxxLL motif when a series of mutagenesis studies determined the strongest conservation of hydrophobic residues L_1 and L_4[78,79]. Many of the studies investigating principles of LxxLL interactions with 16E6 only contain a small number of motifs and often focus on specific interactions, such as that of UBE3A[80,81]. By using a large unbiased library of naïve peptides to measure and weight affinities we were able to conclude that while likely most, if not all, LxxLL partners of 16E6 have already been identified, there are several key characteristics that underpin binding to this region. This is especially important when considering that many LxxLL partners contain multiple motifs and many therapeutic strategies against 16E6 involve targeting the LxxLL pocket[82-85].

The PDZome

Titration series of PDZ proteins SCRIB and SNX27 give us insight into binding mechanisms that reflect our knowledge of their structure. The plateau shape of the SCRIB binding curve indicates that this high affinity interaction is mediated by avidity corresponding to its four PDZ domains[86] compared to that of SNX27 which is a high affinity interaction mediated by a singular PDZ domain[87]. In this case it appears that the SCRIB binding mode follows the establishment of a multivalent interaction where with each successive PDZ binding event the overall enthalpy remains constant while loss of entropy is reduced due to the decreasing number of degrees of freedom. This results in a binding constant that decreases with the establishment of each successive contact and a K_d that is globally lower than that of the first binding event[88]. It has be reported that SCRIB also facilitates interactions with other proteins in this manner[89,90]. Additionally, it has been described that PDZ proteins can act as scaffolds meaning these proteins could be responsible for recruitment of indirect 16E6-binders[91]. It is possible that if these indirect interactors form secondary contacts directly with 16E6 it could follow a similar model of positive cooperativity resulting in higher affinity interactions than that would be expected from an otherwise indirect interaction. Furthermore, this could explain the loss of partners with unknown interfaces when using the 16E6 C-terminal PBM as bait as any potential site for secondary contact would be lost, potentially destabilizing the interaction. However, it is difficult to model complex binding modes such as cooperativity, especially without structural insights to interfaces, and therefore it is best to assume simple binding modes until advances in modelling have been made.

(Manuscript in preparation)

Investigating novel 16E6-binders

To clarify the remaining ambiguity surrounding which 16E6 interface mediates interactions with novel 16E6-binders, we performed a nHU spike-in assay where either the C-terminal PBM or LxxLL binding pocket was blocked by a PDZ-MBP domain or LxxLL peptide, respectively. Compared to the SNX27 and PXN internal standards, PTPN14 and USP13 both demonstrated a statistically significant decrease in binding when the C-terminal PBM of 16E6 was sequestered (Fig. 6A). Taken together with the results of nHU experiments using a PBM peptide of 16E6 as bait which reported loss of PTPN14 and USP13 binding (Fig. 5) this indicates these interactions are likely facilitated directly through novel interfaces at the 16E6 C-terminus. However, these results are not sufficient to entirely rule out indirect binding through a PDZ intermediary. It's possible that if these 16E6-binders establish secondary contacts with another interface of 16E6 that they can be recruited through PDZ proteins, but their binding would still be lost when using the peptide bait due to absence of these additional contacts destabilizing the interaction.

Conclusions

The HPV oncoprotein E6 is known to target a myriad of host factors with variable binding profiles based on the HPV type. As such the binding profile of 16E6 is of high interest due to its involvement in cervical cancer development. Here we mapped the full-length affinity interactome of 16E6 with a considerable level of molecular detail as we not only identified novel and previously described partners of 16E6, but we also cross-validated binding modes of some of these highly studied interactions. This gives us a level of confidence in the 16E6-binders discovered through our study while further validation of affinity measurements with other biophysical techniques instills a level of confidence in our affinities measured.

Furthermore, using a data informed molecular fishing approach allowed us to home in on 16E6-binders that likely have significant roles during the initial stages of HPV induced transformation. While further follow-up is necessary to understand the contributions of 16E6-binders such as, PTPN14, USP13, and XIAP, to HPV pathogenesis they hold promise given they have all been previously implicated in HPV pathologies. This represents a new approach in investigating and interpreting interactomes.

Finally, it is necessary to acknowledge the computational boom seen in biological sciences and the grand potential for applications with affinities. We have entered an era of mass data with the ability to produce, process, and integrate data on a scale that was unimaginable just 25 years ago. While the field of interactomics has kept up with the times in regards to developing high throughput methodologies, a need persists for more quantitative techniques to increase confidence across orthogonal methods and improve the overall quality of data. Furthermore, interactomics data represents an untapped resource in computational approaches, such as in machine learning, due to the natural graphical construction of the data.

(Manuscript in preparation)

Affinity predictions, for example, could be possible using a similar architecture to AlphaFold with graphical transformers, an application that is highly desirable in the realm of drug design. On the other hand, the complexome represents important parameters in modeling which could help improve Molecular Dynamics (MD) simulations and other *in silico* models. Both are helpful tools in furthering our understanding of biological concepts and improving experimental design, yet they are dependent on the production of affinity datasets for training and input. Hence, we are encouraging the field of interactomics to adopt quantitative methods, such as the nHU which is simple and adaptable, to further our understanding of biological systems and allow affinities to realize their full potential.

References

1 Petti, M. & Farina, L. Network medicine for patients' stratification: From single-layer to multi-omics. *WIREs Mechanisms of Disease* **15**, e1623, doi:https://doi.org/10.1002/wsbm.1623 (2023).

2 Babu, M. & Snyder, M. Multi-Omics Profiling for Health. *Mol Cell Proteomics* **22**, 100561, doi:10.1016/j.mcpro.2023.100561 (2023).

3 Bludau, I. & Aebersold, R. Proteomic and interactomic insights into the molecular basis of cell functional diversity. *Nature Reviews Molecular Cell Biology* **21**, 327-340, doi:10.1038/s41580-020-0231-2 (2020).

4 Sahni, N. *et al*. Widespread Macromolecular Interaction Perturbations in Human Genetic Disorders. *Cell* **161**, 647-660, doi:10.1016/j.cell.2015.04.013 (2015).

5 Weimer, K., Zambo, B. & Gogl, G. Molecules interact. But how strong and how much? *BioEssays* **45**, 2300007, doi:https://doi.org/10.1002/bies.202300007 (2023).

6 Liu, X., Abad, L., Chatterjee, L., Cristea, I. M. & Varjosalo, M. Mapping protein–protein interactions by mass spectrometry. *Mass Spectrometry Reviews* **n/a**, doi:https://doi.org/10.1002/mas.21887.

7 Moody, C. A. & Laimins, L. A. Human papillomavirus oncoproteins: pathways to transformation. *Nat Rev Cancer* **10**, 550-560, doi:10.1038/nrc2886 (2010).

8 Baseman, J. G. & Koutsky, L. A. The epidemiology of human papillomavirus infections. *Journal of Clinical Virology* **32**, 16-24, doi:https://doi.org/10.1016/j.jcv.2004.12.008 (2005).

9 Magalhães, G. M. *et al*. Update on human papilloma virus - part I: epidemiology, pathogenesis, and clinical spectrum. *Anais Brasileiros de Dermatologia* **96**, 1-16, doi:https://doi.org/10.1016/j.abd.2020.11.003 (2021).

10 Van Doorslaer K, L. Z., Xirasagar S, Maes P, Kaminsky D, et al. The papillomavirus Episteme: a major update to the papillomavirus sequence database. *Nucleic Acids Res* **45**, D499-D506, doi:10.1093/nar/gkw879 (2017).

11 *PaVE The papillomavirus episteme. Available from: pave.niaid.nih.gov*.

12 Doorbar, J., Egawa, N., Griffin, H., Kranjec, C. & Murakami, I. Human papillomavirus molecular biology and disease association. *Rev Med Virol* **25 Suppl 1**, 2-23, doi:10.1002/rmv.1822 (2015).

13 Graham, Sheila V. The human papillomavirus replication cycle, and its links to cancer progression: a comprehensive review. *Clinical Science* **131**, 2201-2221, doi:10.1042/CS20160786 (2017).

(Manuscript in preparation)

14 Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* **68**, 394-424, doi:10.3322/caac.21492 (2018).
15 Dunne, E. F. & Park, I. U. HPV and HPV-Associated Diseases. *Infectious Disease Clinics of North America* **27**, 765-778, doi:https://doi.org/10.1016/j.idc.2013.09.001 (2013).
16 Scarth, J. A., Patterson, M. R., Morgan, E. L. & Macdonald, A. The human papillomavirus oncoproteins: a review of the host pathways targeted on the road to transformation. *J Gen Virol* **102**, doi:10.1099/jgv.0.001540 (2021).
17 Song, S., Liem, A., Miller, J. A. & Lambert, P. F. Human Papillomavirus Types 16 E6 and E7 Contribute Differently to Carcinogenesis. *Virology* **267**, 141-150, doi:https://doi.org/10.1006/viro.1999.0106 (2000).
18 Eckhardt, M. *et al.* Multiple Routes to Oncogenesis Are Promoted by the Human Papillomavirus-Host Protein Network. *Cancer Discov* **8**, 1474-1489, doi:10.1158/2159-8290.Cd-17-1018 (2018).
19 Rozenblatt-Rosen, O. *et al.* Interpreting cancer genomes using systematic host network perturbations by tumour virus proteins. *Nature* **487**, 491-495, doi:10.1038/nature11288 (2012).
20 White, E. A. & Howley, P. M. Proteomic approaches to the study of papillomavirus-host interactions. *Virology* **435**, 57-69, doi:10.1016/j.virol.2012.09.046 (2013).
21 Wang, J., Dupuis, C., Tyring, S. K. & Underbrink, M. P. Sterile α Motif Domain Containing 9 Is a Novel Cellular Interacting Partner to Low-Risk Type Human Papillomavirus E6 Proteins. *PLOS ONE* **11**, e0149859, doi:10.1371/journal.pone.0149859 (2016).
22 Neveu, G. *et al.* Comparative analysis of virus–host interactomes with a mammalian high-throughput protein complementation assay based on Gaussia princeps luciferase. *Methods* **58**, 349-359, doi:https://doi.org/10.1016/j.ymeth.2012.07.029 (2012).
23 White, E. A. *et al.* Comprehensive Analysis of Host Cellular Interactions with Human Papillomavirus E6 Proteins Identifies New E6 Binding Partners and Reflects Viral Diversity. *J Virol* **86**, 13174-13186, doi:doi:10.1128/jvi.02172-12 (2012).
24 Poirson, J. *et al.* Mapping the interactome of HPV E6 and E7 oncoproteins with the ubiquitin-proteasome system. *The FEBS Journal* **284**, 3171-3201, doi:https://doi.org/10.1111/febs.14193 (2017).
25 Xue, B., Ganti, K., Rabionet, A., Banks, L. & Uversky, V. N. Disordered interactome of human papillomavirus. *Curr Pharm Des* **20**, 1274-1292, doi:10.2174/13816128113199990072 (2014).
26 Martinez-Zapien, D. *et al.* Structure of the E6/E6AP/p53 complex required for HPV-mediated degradation of p53. *Nature* **529**, 541-545, doi:10.1038/nature16481 (2016).
27 Beckett, D., Kovaleva, E. & Schatz, P. J. A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Sci* **8**, 921-929, doi:10.1110/ps.8.4.921 (1999).

(Manuscript in preparation)

28 Sidi, A. O. *et al.* Strategies for bacterial expression of protein-peptide complexes: application to solubilization of papillomavirus E6. *Protein Expr Purif* **80**, 8-16, doi:10.1016/j.pep.2011.06.013 (2011).
29 Kay, B. K., Thai, S. & Volgina, V. V. in *High Throughput Protein Expression and Purification: Methods and Protocols* (ed Sharon A. Doyle) 185-198 (Humana Press, 2009).
30 Li, Y. & Sousa, R. Expression and purification of E. coli BirA biotin ligase for in vitro biotinylation. *Protein Expression and Purification* **82**, 162-167, doi:https://doi.org/10.1016/j.pep.2011.12.008 (2012).
31 Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nature Methods* **9**, 676-682, doi:10.1038/nmeth.2019 (2012).
32 Wilkins, M. R. *et al.* Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol* **112**, 531-552, doi:10.1385/1-59259-584-7:531 (1999).
33 Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**, 248-254, doi:https://doi.org/10.1016/0003-2697(76)90527-3 (1976).
34 Zambo, B., Morlet, B., Negroni, L., Trave, G. & Gogl, G. Native holdup (nHU) to measure binding affinities from cell extracts. *bioRxiv*, 2022.2006.2022.497144, doi:10.1101/2022.06.22.497144 (2022).
35 Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods* **13**, 731-740, doi:10.1038/nmeth.3901 (2016).
36 Edwards, R. J., Davey, N. E. & Shields, D. C. SLiMFinder: A Probabilistic Method for Identifying Over-Represented, Convergently Evolved, Short Linear Motifs in Proteins. *PLOS ONE* **2**, e967, doi:10.1371/journal.pone.0000967 (2007).
37 Davey, N. E., Haslam, N. J., Shields, D. C. & Edwards, R. J. SLiMFinder: a web server to find novel, significantly over-represented, short protein motifs. *Nucleic Acids Research* **38**, W534-W539, doi:10.1093/nar/gkq440 (2010).
38 Krystkowiak, I. & Davey, N. E. SLiMSearch: a framework for proteome-wide discovery and annotation of functional modules in intrinsically disordered regions. *Nucleic Acids Research* **45**, W464-W469, doi:10.1093/nar/gkx238 (2017).
39 Gogl, G. *et al.* Quantitative fragmentomics allow affinity mapping of interactomes. *Nature Communications* **13**, 5472, doi:10.1038/s41467-022-33018-0 (2022).
40 Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589, doi:10.1038/s41586-021-03819-2 (2021).
41 Varadi, M. *et al.* AlphaFold Protein Structure Database in 2024: providing structure coverage for over 214 million protein sequences. *Nucleic Acids Res* **52**, D368-d375, doi:10.1093/nar/gkad1011 (2024).
42 Schmid, E. W. & Walter, J. C. Predictomes: A classifier-curated database of AlphaFold-modeled protein-protein interactions. *bioRxiv*, doi:10.1101/2024.04.09.588596 (2024).
43 Vasilief, I., Gadiou, R. & Franke, K. (2019).

(Manuscript in preparation)

44 Elston, R. C., Napthine, S. & Doorbar, J. The identification of a conserved binding motif within human papillomavirus type 16 E6 binding peptides, E6AP and E6BP. *J Gen Virol* **79 (Pt 2)**, 371-374, doi:10.1099/0022-1317-79-2-371 (1998).
45 Guo, H.-B. *et al.* AlphaFold2 models indicate that protein sequence determines both structure and dynamics. *Scientific Reports* **12**, 10696, doi:10.1038/s41598-022-14382-9 (2022).
46 Wilson, C. J., Choy, W. Y. & Karttunen, M. AlphaFold2: A Role for Disordered Protein/Region Prediction? *Int J Mol Sci* **23**, doi:10.3390/ijms23094591 (2022).
47 Xu, J. & Zhang, Y. How significant is a protein structure similarity with TM-score = 0.5? *Bioinformatics* **26**, 889-895, doi:10.1093/bioinformatics/btq066 (2010).
48 Li, J. *et al.* Conformational equilibria and intrinsic affinities define integrin activation. *EMBO J* **36**, 629-645, doi:10.15252/embj.201695803 (2017).
49 Linkuvienė, V., Zubrienė, A. & Matulis, D. Intrinsic affinity of protein – ligand binding by differential scanning calorimetry. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1870**, 140830, doi:https://doi.org/10.1016/j.bbapap.2022.140830 (2022).
50 Poirson, J. *et al.* High-Risk Mucosal Human Papillomavirus 16 (HPV16) E6 Protein and Cutaneous HPV5 and HPV8 E6 Proteins Employ Distinct Strategies To Interfere with Interferon Regulatory Factor 3-Mediated Beta Interferon Expression. *J Virol* **96**, e01875-01821, doi:doi:10.1128/jvi.01875-21 (2022).
51 Song, W. J., Sontz, P. A., Ambroggio, X. I. & Tezcan, F. A. Metals in protein-protein interfaces. *Annu Rev Biophys* **43**, 409-431, doi:10.1146/annurev-biophys-051013-023038 (2014).
52 Durmuş, S. & Ülgen, K. Ö. Comparative interactomics for virus–human protein–protein interactions: DNA viruses versus RNA viruses. *FEBS open bio* **7**, 96-107 (2017).
53 Turkki, P. *et al.* Tensin-2 interactomics reveals interaction with GAPDH and a phosphorylation-mediated regulatory role in glycolysis. *Scientific Reports* **14**, 19862, doi:10.1038/s41598-024-65787-7 (2024).
54 Gogl, G. *et al.* Dual Specificity PDZ- and 14-3-3-Binding Motifs: A Structural and Interactomics Study. *Structure* **28**, 747-759.e743, doi:10.1016/j.str.2020.03.010 (2020).
55 Scheffner, M., Werness, B. A., Huibregtse, J. M., Levine, A. J. & Howley, P. M. The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell* **63**, 1129-1136, doi:10.1016/0092-8674(90)90409-8 (1990).
56 Huibregtse, J. M., Scheffner, M. & Howley, P. M. A cellular protein mediates association of p53 with the E6 oncoprotein of human papillomavirus types 16 or 18. *EMBO J* **10**, 4129-4135 (1991).
57 Thomas, M., Pim, D. & Banks, L. The role of the E6-p53 interaction in the molecular pathogenesis of HPV. *Oncogene* **18**, 7690-7700, doi:10.1038/sj.onc.1202953 (1999).
58 Tommasino, M. *et al.* The role of TP53 in Cervical carcinogenesis. *Hum Mutat* **21**, 307-312, doi:10.1002/humu.10178 (2003).
59 Boshart, M. *et al.* A new type of papillomavirus DNA, its presence in genital cancer biopsies and in cell lines derived from cervical cancer. *EMBO J* **3**, 1151-1157, doi:10.1002/j.1460-2075.1984.tb01944.x (1984).

(Manuscript in preparation)

60 MATLASHEWSKI, G., BANKS, L., PIM, D. & CRAWFORD, L. Analysis of human p53 proteins and mRNA levels in normal and transformed cells. *European Journal of Biochemistry* **154**, 665-672, doi:https://doi.org/10.1111/j.1432-1033.1986.tb09449.x (1986).
61 White, E. A., Münger, K. & Howley, P. M. High-risk human papillomavirus E7 proteins target PTPN14 for degradation. *MBio* **7**, 10.1128/mbio. 01530-01516 (2016).
62 Morgan, E. L. *et al*. The deubiquitinase (DUB) USP13 promotes Mcl-1 stabilisation in cervical cancer. *Oncogene* **40**, 2112-2129, doi:10.1038/s41388-021-01679-8 (2021).
63 Wang, Q. *et al*. Role of USP13 in physiology and diseases. *Frontiers in Molecular Biosciences* **9**, 977122 (2022).
64 Zhang, Y.-H., Zhou, C.-J., Zhou, Z.-R., Song, A.-X. & Hu, H.-Y. Domain analysis reveals that a deubiquitinating enzyme USP13 performs non-activating catalysis for Lys63-linked polyubiquitin. *PloS one* **6**, e29362 (2011).
65 Scheffner, M. & Whitaker, N. J. in *Seminars in cancer biology*. 59-67 (Elsevier).
66 Đukić, A. *et al*. HPV oncoproteins and the ubiquitin proteasome system: a signature of malignancy? *Pathogens* **9**, 133 (2020).
67 Pim, D., Thomas, M., Javier, R., Gardiol, D. & Banks, L. HPV E6 targeted degradation of the discs large protein: evidence for the involvement of a novel ubiquitin ligase. *Oncogene* **19**, 719-725 (2000).
68 Grm, H. S. & Banks, L. Degradation of hDlg and MAGIs by human papillomavirus E6 is E6-AP-independent. *Journal of general virology* **85**, 2815-2819 (2004).
69 Cabeça, T. K. *et al*. HPV-Mediated Resistance to TNF and TRAIL Is Characterized by Global Alterations in Apoptosis Regulatory Factors, Dysregulation of Death Receptors, and Induction of ROS/RNS. *International Journal of Molecular Sciences* **20**, 198 (2019).
70 Ganti, K. *et al*. Interaction of the Human Papillomavirus E6 Oncoprotein with Sorting Nexin 27 Modulates Endocytic Cargo Transport Pathways. *PLoS Pathog* **12**, e1005854, doi:10.1371/journal.ppat.1005854 (2016).
71 Ronchi, V. P., Klein, J. M., Edwards, D. J. & Haas, A. L. The active form of E6-associated protein (E6AP)/UBE3A ubiquitin ligase is an oligomer. *J Biol Chem* **289**, 1033-1048, doi:10.1074/jbc.M113.517805 (2014).
72 Wang, Z. *et al*. Structural insights into the functional mechanism of the ubiquitin ligase E6AP. *Nature Communications* **15**, 3531, doi:10.1038/s41467-024-47586-w (2024).
73 Sandate, C. R., Chakraborty, D., Kater, L., Kempf, G. & Thomä, N. H. Structural insights into viral hijacking of p53 by E6 and E6AP. *bioRxiv*, 2023.2011.2001.565136, doi:10.1101/2023.11.01.565136 (2023).
74 Henzler-Wildman, K. & Kern, D. Dynamic personalities of proteins. *Nature* **450**, 964-972, doi:10.1038/nature06522 (2007).
75 Gentile, F. The free energy landscape of small-world networks of cells. *Journal of Biomechanics* **162**, 111909, doi:https://doi.org/10.1016/j.jbiomech.2023.111909 (2024).
76 Tsai, C.-J. & Nussinov, R. The free energy landscape in translational science: how can somatic mutations result in constitutive oncogenic activation? *Physical Chemistry Chemical Physics* **16**, 6332-6341, doi:10.1039/C3CP54253J (2014).

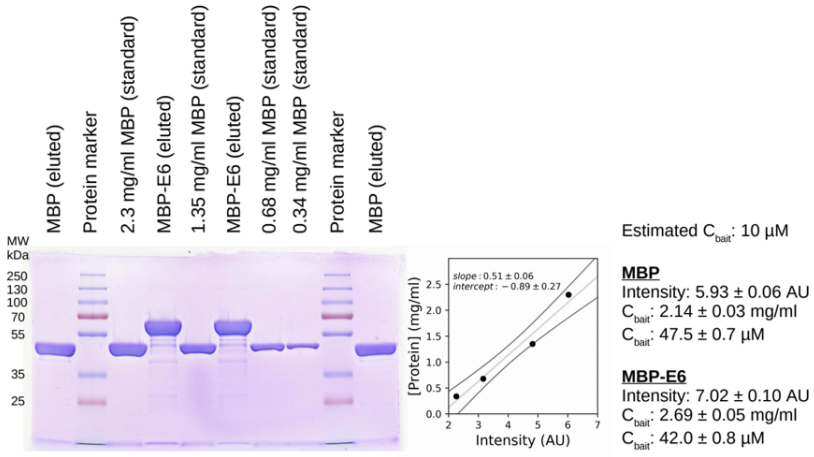
(Manuscript in preparation)

77 Nussinov, R. & Wolynes, P. G. A second molecular biology revolution? The energy landscapes of biomolecular function. *Physical Chemistry Chemical Physics* **16**, 6321-6322, doi:10.1039/C4CP90027H (2014).
78 Bohl, J., Das, K., Dasgupta, B. & Pol, S. B. V. Competitive binding to a charged leucine motif represses transformation by a papillomavirus E6 oncoprotein. *Virology* **271**, 163-170 (2000).
79 Chen, J. J., Hong, Y., Rustamzadeh, E., Baleja, J. D. & Androphy, E. J. Identification of an α helical motif sufficient for association with papillomavirus E6. *Journal of Biological Chemistry* **273**, 13537-13544 (1998).
80 Zanier, K. *et al.* Structural basis for hijacking of cellular LxxLL motifs by papillomavirus E6 oncoproteins. *Science* **339**, 694-698, doi:10.1126/science.1229934 (2013).
81 Ansari, T., Brimer, N. & Pol, S. B. V. Peptide Interactions Stabilize and Restructure Human Papillomavirus Type 16 E6 To Interact with p53. *J Virol* **86**, 11386-11391, doi:doi:10.1128/jvi.01236-12 (2012).
82 Ramirez, J. *et al.* Targeting the two oncogenic functional sites of the HPV E6 oncoprotein with a high-affinity bivalent ligand. *Angewandte Chemie* **127**, 8069-8073 (2015).
83 Rietz, A. *et al.* Molecular Probing of the HPV-16 E6 Protein Alpha Helix Binding Groove with Small Molecule Inhibitors. *PLOS ONE* **11**, e0149845, doi:10.1371/journal.pone.0149845 (2016).
84 Chitsike, L. & Duerksen-Hughes, P. J. PPI Modulators of E6 as Potential Targeted Therapeutics for Cervical Cancer: Progress and Challenges in Targeting E6. *Molecules* **26**, 3004 (2021).
85 Ye, X. *et al.* Discovery of reactive peptide inhibitors of human papillomavirus oncoprotein E6. *Chemical Science* **14**, 12484-12497 (2023).
86 Nakagawa, S. & Huibregtse, J. M. Human scribble (Vartul) is targeted for ubiquitin-mediated degradation by the high-risk papillomavirus E6 proteins and the E6AP ubiquitin-protein ligase. *Mol Cell Biol* **20**, 8244-8253, doi:10.1128/mcb.20.21.8244-8253.2000 (2000).
87 Yong, X. *et al.* SNX27-FERM-SNX1 complex structure rationalizes divergent trafficking pathways by SNX17 and SNX27. *Proceedings of the National Academy of Sciences* **118**, e2105510118, doi:doi:10.1073/pnas.2105510118 (2021).
88 Erlendsson, S. & Teilum, K. Binding Revisited—Avidity in Cellular Function and Signaling. *Frontiers in Molecular Biosciences* **7**, doi:10.3389/fmolb.2020.615565 (2021).
89 How, J. Y., Caria, S., Humbert, P. O. & Kvansakul, M. Crystal structure of the human Scribble PDZ 1 domain bound to the PDZ-binding motif of APC. *FEBS letters* **593**, 533-542 (2019).
90 Lim, K. Y., Gödde, N. J., Humbert, P. O. & Kvansakul, M. Structural basis for the differential interaction of Scribble PDZ domains with the guanine nucleotide exchange factor β-PIX. *Journal of Biological Chemistry* **292**, 20425-20436 (2017).
91 Erlendsson, S. *et al.* Mechanisms of PDZ domain scaffold assembly illuminated by use of supported cell membrane sheets. *eLife* **8**, e39180, doi:10.7554/eLife.39180 (2019).

(Manuscript in preparation)

Supplementary Material

Figure S1 Concentration measurements of immobilized baits.



(Manuscript in preparation)

Figure S2 Meta-analysis of 16E6 interactomic studies

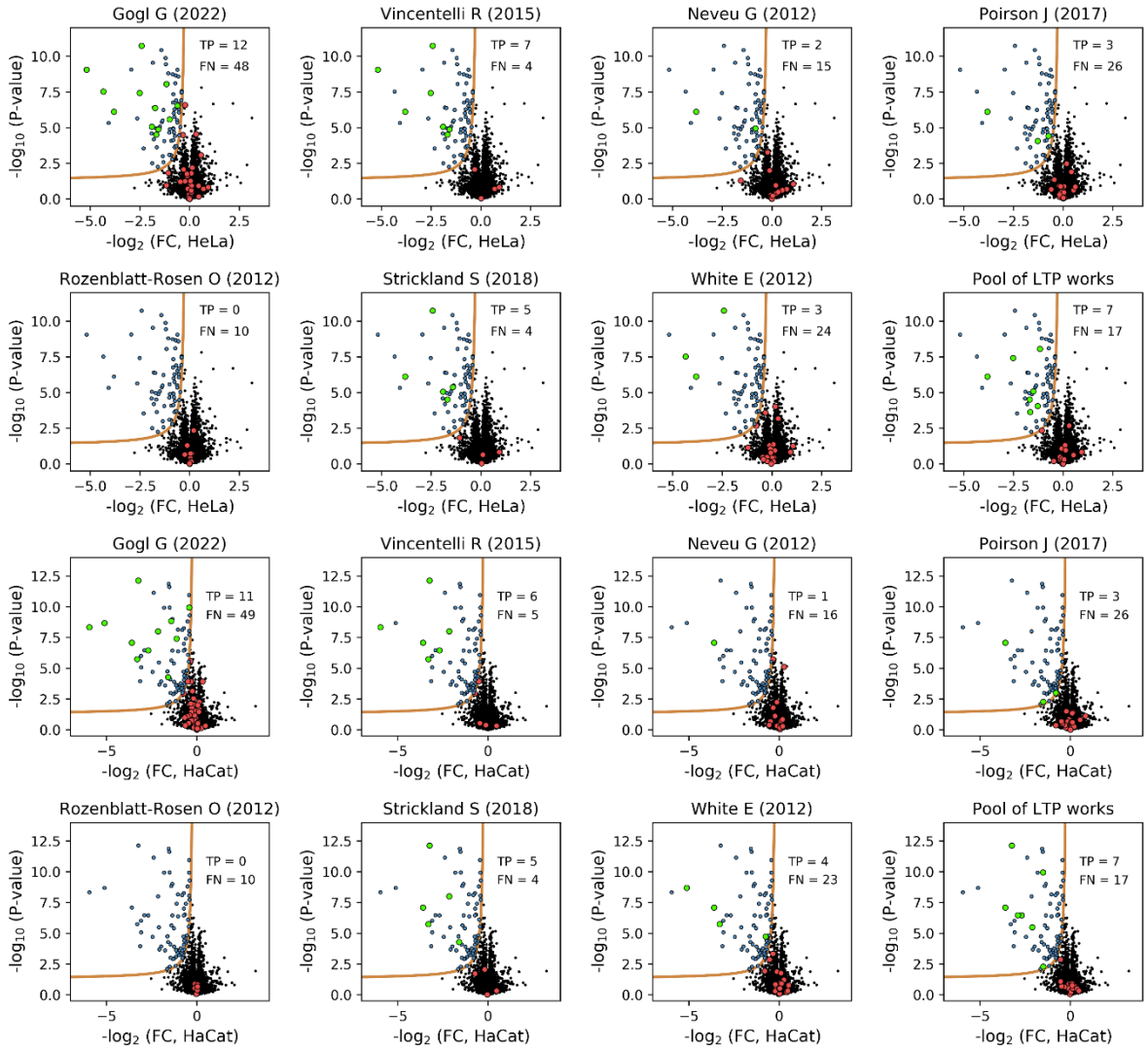
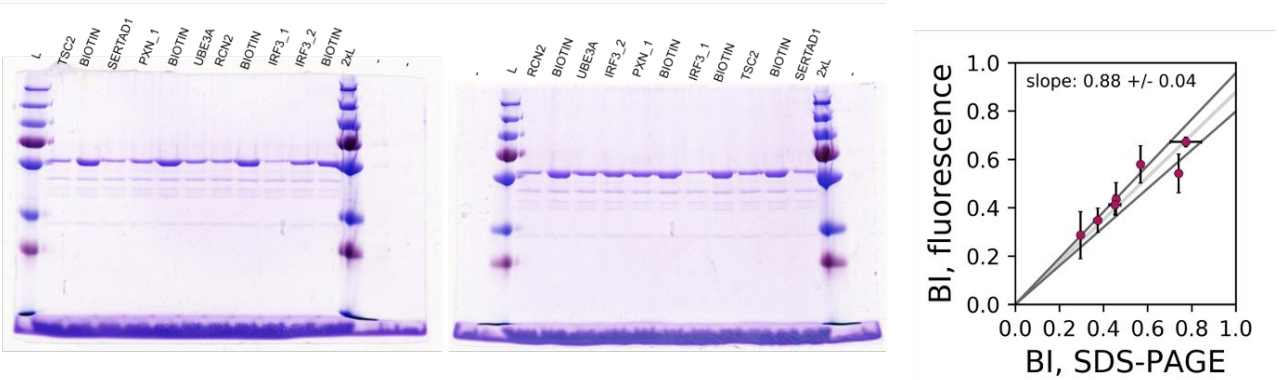


Figure S4 SDS-PAGE validation of HU



(Manuscript in preparation)

Figure S5 XIAP and MSH6 binding interface prism plots

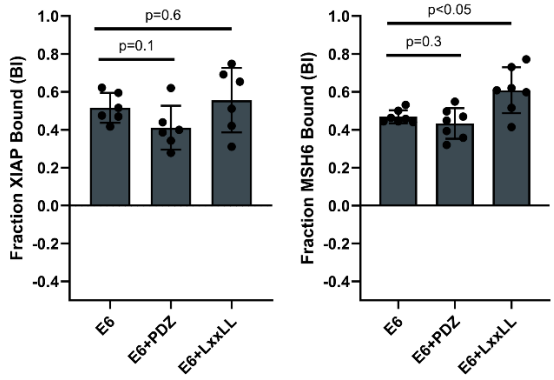


Table S1 Terrific Broth Recipe for 16E6 Expression

Component	**Final Concentration**
Yeast Extract	24 g/L
Tryptone	20 g/L
Glycerol	0.4%
Phosphate Buffer – 0.17 M KH_2PO_4/ 0.72 K_2HPO_4	10%

*TB prepared without pH adjustment (pH ~7.2)

Table S2 Antibodies for nHU-WB

Name	**Reference**	**Manufacturer**	**Species**	**Dilution**	**Buffer: TBS-T**	**Molecular Weight (KDa)**
GAPDH	MAB374	Sigma	ms	1:5000	5% milk	35
IRF3	1132-1-AP	Proteintech	rb	1:5000	5% milk	47-60
MSH6	66172-1-Ig	Proteintech	ms	1:1000	5% milk	160
PXN	SAB4502553	Sigma	rb	1:1000	5% milk	64
PTPN14	D5T6Y	CST	rb	1:1000	5% milk	160
SCRIB	PA5-54821	Thermo	rb	1:1000	5% milk	175-250
SNX27	MA5-27854	Thermo	ms	1:1000	5% milk	61
TP53	2527S	CST	rb	1:1000	5% milk	53
UBE3A	clone 3E5	Sigma	ms	1:1000	5% milk	105
USP13	16840-1-AP	Proteintech	rb	1:1000	5% milk	97-100
XIAP	66800-1-Ig	Proteintech	ms	1:5000	5% milk	57
APOBEC	E1A2G	CST	rb	1:1000	5% BSA	35

(Manuscript in preparation)

Table S3 Biotinylated LxxLL peptides use for LxxLLome HU assays

Accession	Gene Name	Residues	Peptide Name	Peptide Sequence	Comment
Q9NQS1	AVEN	281-295	Q9NQS1-AVEN-281-295	biotin-Ado-Ado-HLEEELDLLLNLDAP-amidated	
Q8NI22	MCFD2	1079-1093	Q8NI22-MCFD2-1079-1093	biotin-Ado-Ado-ELINIIDGVLRDDDK-amidated	
Q9NP74	PALMD	234-248	Q9NP74-PALMD-234-248	biotin-Ado-Ado-LAPVEVEELLRQASE-amidated	
P19484	TFEB	162-176	P19484-TFEB-162-176	biotin-Ado-Ado-NPERELDDVIDNIMR-amidated	
Q969V6	MKL1	471-485	Q969V6-MKL1-471-485	biotin-Ado-Ado-TPGDTFGEMVTSPLT-amidated	
Q86UU0	BCL9L	1321-1335	Q86UU0-BCL9L-1321-1335	biotin-Ado-Ado-IPEFDLSRIIPSEKP-amidated	
Q8IY92	SLX4	563-577	Q8IY92-SLX4-563-577	biotin-Ado-Ado-LMQEPVPPLVPPEHS-amidated	
Q7Z5J4	RAI1	869-883	Q7Z5J4-RAI1-869-883	biotin-Ado-Ado-NISNTVQQLLLSKAA-amidated	
Q13033	STRN3	300-314	Q13033-STRN3-300-314	biotin-Ado-Ado-EALKEFDFLVTAEDG-amidated	
Q68D10	SPTY2D1	596-610	Q68D10-SPTY2D1-596-610	biotin-Ado-Ado-EYDSEMEDFIEDEGE-amidated	
P38398	BRCA1	1409-1423	P38398-BRCA1-1409-1423	biotin-Ado-Ado-QEMAELEAVLEQHGS-amidated	
Q8N7J2	AMER2	550-564	Q8N7J2-AMER2-550-564	biotin-Ado-Ado-YSGDALYDLYADPDG-amidated	
Q9Y5S2	CDC42BPB	1596-1610	Q9Y5S2-CDC42BPB-1596-1610	biotin-Ado-Ado-GPGDGMQVLMDLPLS-amidated	
Q8N4Y2	CRACR2B	206-220	Q8N4Y2-CRACR2B-206-220	biotin-Ado-Ado-EHEREVRALYEETEQ-amidated	
Q14676	MDC1	640-654	Q14676-MDC1-640-654	biotin-Ado-Ado-ISRENLTDLVVDTDT-amidated	
Q8IW19	APLF	440-454	Q8IW19-APLF-440-454	biotin-Ado-Ado-HNTLPVRNVLDEDND-amidated	
Q8WXA3	RUFY2	289-303	Q8WXA3-RUFY2-289-303	biotin-Ado-Ado-HSRQGLDEMYNEARR-amidated	
P49750	YLPM1	202-216	P49750-YLPM1-202-216	biotin-Ado-Ado-QKHTQLQQILQQYQQ-amidated	
Q12955	ANK3	2153-2167	Q12955-ANK3-2153-2167	biotin-Ado-Ado-FHEVPIPPVITETRT-amidated	
Q8N4L1	TMEM151A	Jul-21	Q8N4L1-TMEM151A-7-21	biotin-Ado-Ado-GDGGEVPALIPDGEP-amidated	
P49815	TSC2	1389-1403	P49815-TSC2-1389-1403	biotin-Ado-Ado-PELQTLQDILGDPGD-amidated	
P31943	HNRNPH1	428-442	P31943-HNRNPH1-428-442	biotin-Ado-Ado-SSMSGYDQVLQENSS-amidated	
Q9UHY8	FEZ2	164-178	Q9UHY8-FEZ2-164-178	biotin-Ado-Ado-QVIEEIEEMMQESPD-amidated	
Q5VWK0	NBPF6	430-444	Q5VWK0-NBPF6-430-444	biotin-Ado-Ado-VPEDSVNEVYLTPSV-amidated	
Q8WYA6	CTNNBL1	1-12	Q8WYA6-CTNNBL1-1-12	Biotin-Ado-Ado-MDVGELLSYQPN-amidated	short motif
Q12888	TP53BP1	633-647	Q12888-TP53BP1-633-647	biotin-Ado-Ado-TRSEALSSVLDQEEA-amidated	
Q9BRD0	BUD13	75-89	Q9BRD0-BUD13-75-89	biotin-Ado-Ado-GDLPVVAEFVDERPE-amidated	
Q9UHV2	SERTAD1	210-224	Q9UHV2-SERTAD1-210-224	biotin-Ado-Ado-LDEAELDYLMDVLVG-amidated	
Q9NNW5	WDR6	734-748	Q9NNW5-WDR6-734-748	biotin-Ado-Ado-SEGPDLTDIVITCSE-amidated	
Q14676	MDC1	640-654	Q14676-MDC1-640-654	biotin-Ado-Ado-ISRENLTDLVVDTDT-amidated	

(Manuscript in preparation)

Q13227	GPS2	31-45	Q13227-GPS2-31-45	biotin-Ado-Ado-QEEEEVDKMMEQKMK-amidated	
Q8IXT5	RBM12B	101-115	Q8IXT5-RBM12B-101-115	biotin-Ado-Ado-SGVDSLSNFIESVKE-amidated	
P49023	PXN	1-13	P49023-PXN-1-13	Biotin-Ado-Ado-MDDLDALLADLES-amidated	short motif
Q12955	ANK3	835-849	Q12955-ANK3-835-849	biotin-Ado-Ado-NVPETMNEVLDMSDD-amidated	
Q14004	CDK13	607-621	Q14004-CDK13-607-621	biotin-Ado-Ado-LPPLPLPPMLPEDKE-amidated	
Q8N7U6	EFHB	207-221	Q8N7U6-EFHB-207-221	biotin-Ado-Ado-NTEPQMGLVIEPPQC-amidated	
Q9ULD2	MTUS1	407-421	Q9ULD2-MTUS1-407-421	biotin-Ado-Ado-IEKKSLEDLLSEKQE-amidated	
Q92585	MAML1	1004-1016	Q92585-MAML1-1004-1016	Biotin-Ado-Ado-EWMSDLDDLLGSQ-free COOH	C-terminal motif
Q8NEZ4	KMT2C	1483-1497	Q8NEZ4-KMT2C-1483-1497	biotin-Ado-Ado-LSEEQLDGILSPELD-amidated	
Q8TC90	CCER1	266-280	Q8TC90-CCER1-266-280	biotin-Ado-Ado-EENQSLAPLLVEEEE-amidated	
Q9BSJ8	ESYT1	1005-1019	Q9BSJ8-ESYT1-1005-1019	biotin-Ado-Ado-PPDPYVSLLLLPDKN-amidated	
Q5W5W9	RESP18	113-127	Q5W5W9-RESP18-113-127	biotin-Ado-Ado-EQEEKLQLLFPSETH-amidated	
Q9NYL9	TMOD3	444-458	Q9NYL9-TMOD3-444-458	biotin-Ado-Ado-TELKQLETVLDDLDP-amidated	
Q9Y3T9	NOC2L	108-122	Q9Y3T9-NOC2L-108-122	biotin-Ado-Ado-GPFHSLPDVLEEASE-amidated	
Q05086	UBE3A	404-418	Q05086-UBE3A-404-418	biotin-Ado-Ado-SSELTLQELLGEERR-amidated	
Q9NSK0	KLC4	170-184	Q9NSK0-KLC4-170-184	biotin-Ado-Ado-ATKDSLDDLFPNEEE-amidated	
Q00587	CDC42EP1	277-291	Q00587-CDC42EP1-277-291	biotin-Ado-Ado-LGPSLLSELLGVMSL-amidated	
Q86WK6	AMIGO1	477-491	Q86WK6-AMIGO1-477-491	biotin-Ado-Ado-SDPESVSSVFSDTPI-amidated	
Q15018	ABRAXAS2	372-386	Q15018-ABRAXAS2-372-386	biotin-Ado-Ado-SDDSDYENLIDPTEP-amidated	
Q14257	RCN2	198-222	Q14257-RCN2-198-222	biotin-Ado-Ado-DGFVSLEEFLGDYRW-amidated	
Q5THK1	PRR14L	507-521	Q5THK1-PRR14L-507-521	biotin-Ado-Ado-SEESSFSSLMQIEEA-amidated	
Q8WUY3	PRUNE2	1079-1093	Q8WUY3-PRUNE2-1079-1093	biotin-Ado-Ado-YDDPSMMQLYNETNR-amidated	
Q96JL9	ZNF333	83-97	Q96JL9-ZNF333-83-97	biotin-Ado-Ado-EELPSMQDLLEEASS-amidated	
Q7Z478	DHX29	519-533	Q7Z478-DHX29-519-533	biotin-Ado-Ado-DPEESWENLVSDEDF-amidated	
Q8N108	MIER1	355-369	Q8N108-MIER1-355-369	biotin-Ado-Ado-GVTDYMDRLLDESES-amidated	
Q8IZX4	TAF1L	1756-1770	Q8IZX4-TAF1L-1756-1770	biotin-Ado-Ado-EASVLYEDLLISEGE-amidated	
Q9HCK8	CHD8	475-489	Q9HCK8-CHD8-475-489	biotin-Ado-Ado-RGEQNIPRVLNEDEL-amidated	
Q9NQC1	JADE2	571-585	Q9NQC1-JADE2-571-585	biotin-Ado-Ado-QDEETLLSFMRDPSL-amidated	
P11532	DMD	3651-3665	P11532-DMD-3651-3665	biotin-Ado-Ado-DTSTGLEEVMEQLNN-amidated	
Q9BY07	SLC4A5	93-107	Q9BY07-SLC4A5-93-107	biotin-Ado-Ado-PAAEQLQDILGEEDE-amidated	
Q9NNW5	WDR6	716-730	Q9NNW5-WDR6-716-730	biotin-Ado-Ado-GPEYGVPSFMQPDDL-amidated	
P49815	TSC2	1303-1317	P49815-TSC2-1303-1317	biotin-Ado-Ado-GSPGEVPVLVEPPGL-amidated	
Q6ZV50	RFX8	451-465	Q6ZV50-RFX8-451-465	biotin-Ado-Ado-EDMGTVKEMLPDDPT-amidated	

(Manuscript in preparation)

Q8IWR0	ZC3H7A	281-295	Q8IWR0-ZC3H7A-281-295	biotin-Ado-Ado-VLGDELDDLLDSAPE-amidated
Q9HBT7	ZNF287	115-129	Q9HBT7-ZNF287-115-129	biotin-Ado-Ado-TLVEDLTQILEEEAP-amidated
O95905	ECD	418-432	O95905-ECD-418-432	biotin-Ado-Ado-LSPDQLDQLLQEAVG-amidated
Q9UPY3	DICER1	596-610	Q9UPY3-DICER1-596-610	biotin-Ado-Ado-TGETDIDPVMDDDDV-amidated
Q6ZNG2	DBX2	166-180	Q6ZNG2-DBX2-166-180	biotin-Ado-Ado-PREESMLPLLTQDSN-amidated
Q14653	IRF3	135-149	Q14653-IRF3-135-149	biotin-Ado-Ado-TQEDILDELLGNMVL-amidated
Q9BSK1	ZNF577	395-409	Q9BSK1-ZNF577-395-409	biotin-Ado-Ado-HTSLYMSELIQEQKT-amidated
Q6ZMS4	ZNF852	109-123	Q6ZMS4-ZNF852-109-123	biotin-Ado-Ado-RLESDFLEIIDEDKK-amidated
P49023	PXN	212-226	P49023-PXN-212-226	biotin-Ado-Ado-DVRPSVESLLDELES-amidated
Q14653	IRF3	139-153	Q14653-IRF3-139-153	biotin-Ado-Ado-ILDELLGNMVLAPLP-amidated
P07437	TUBB	410-424	P07437-TUBB-410-424	biotin-Ado-Ado-EAESNMNDLVSEYQQ-amidated
Q14184	DOC2B	78-92	Q14184-DOC2B-78-92	biotin-Ado-Ado-EDDEDVDQLFGAYGS-amidated
Q92995	USP13	404-418	Q92995-USP13-404-418	biotin-Ado-Ado-VKSELIEQVMKEEHK-amidated
Q7Z5J4	RAI1	116-130	Q7Z5J4-RAI1-116-130	biotin-Ado-Ado-EEYSSLCELLGSPEQ-amidated
P68371	TUBB4B	410-424	P68371-TUBB4B-410-424	biotin-Ado-Ado-EAESNMNDLVSEYQQ-amidated
Q92585	MAML1	262-276	Q92585-MAML1-262-276	biotin-Ado-Ado-VPDEDMKDLFNEDFE-amidated
Q8IWX8	CHERP	100-114	Q8IWX8-CHERP-100-114	biotin-Ado-Ado-QGAPSMDELIQQSQW-amidated
Q5SQN1	SNAP47	162-176	Q5SQN1-SNAP47-162-176	biotin-Ado-Ado-TRGEELTGLMAGSQK-amidated
Q96MU7	YTHDC1	11-25	Q96MU7-YTHDC1-11-25	biotin-Ado-Ado-GELNVLDDILTEVPE-amidated
P43243	MATR3	191-205	P43243-MATR3-191-205	biotin-Ado-Ado-DRGPSLNPVLDYDHG-amidated
Q99759	MAP3K3	2-16	Q99759-MAP3K3-2-16	biotin-Ado-Ado-DEQEALNSIMNDLVA-amidated
P19484	TFEB	27-41	P19484-TFEB-27-41	biotin-Ado-Ado-DNIMRLDDVLGYINP-amidated

Table S4 Plate arrangement for LxxLLome HU assays

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24		BIOTIN
A	Biotin		Q9NQS1-AVEN-281-295		Q8NI22-MCFD2-1079-1093		Biotin		Q9NP74-PALMD-234-248		P19484-TFEB-162-176		Q969V6-MKL1-471-485		Q969V6-MKL1-471-485		Q86UU0-BCL9L-1321-1335		Q8IY92-SLX4-563-577		Q7Z5J4-RAI1-869-883		Q13033-STRN3-300-314		A	
B																									B	
C	Q68D10-SPTY2D1-596-610		P38398-BRCA1-1409-1423		Q8N7J2-AMER2-550-564		Q9Y5S2-CDC42BPB-1596-1610		Q8N4Y2-CRACR2B-206-220		Q14676-MDC1-640-654		Q8IW19-APLF-440-454		Biotin		Q8WXA3-RUFY2-289-303		P49750-YLPM1-202-216		Q12955-ANK3-2153-2167		Q8N4L1-TMEM151A-7-21		C	
D																									D	
E	P49815-TSC2-1389-1403		P31943-HNRNPH1-428-442		Q9UHY8-FEZ2-164-178		Q5VWK0-NBPF6-430-444		Biotin		Q8WYA6-CTNNBL1-1-12		Q12888-TP53BP1-633-647		Q9BRD0-BUD13-75-89		Q9UHV2-SERTAD1-210-224		Q9NNW5-WDR6-734-748		Q14676-MDC1-640-654		Q13227-GPS2-31-45		E	
F																									F	
G	Q8IXT5-RBM12B-101-115		P49023-PXN-1-13		Q12955-ANK3-835-849		Q14004-CDK13-607-621		Q8N7U6-EFHB-207-221		Q9ULD2-MTUS1-407-421		Biotin		Q92585-MAML1-1004-1016		Q8NEZ4-KMT2C-1483-1497		Q8TC90-CCER1-266-280		Q9BSJ8-ESYT1-1005-1019		Q5W5W9-RESP18-113-127		G	
H																									H	
I	Q9NYL9-TMOD3-444-458		Q9Y3T9-NOC2L-108-122		Q05086-UBE3A-404-418		Q9NSK0-KLC4-170-184		Q00587-CDC42EP1-277-291		Biotin		Biotin		Q86WK6-AMIGO1-477-491		Q15018-ABRAXAS2-372-386		Q14257-RCN2-198-222		Biotin		Q5THK1-PRR14L-507-521		I	
J																									J	
K	Q8WUY3-PRUNE2-1079-1093		Q96JL9-ZNF333-83-97		Q7Z478-DHX29-519-533		Q8N108-MIER1-355-369		Q8IZX4-TAF1L-1756-1770		Q9HCK8-CHD8-475-489		Q9NQC1-JADE2-571-585		P11532-DMD-3651-3665		Q9BY07-SLC4A5-93-107		Q9NNW5-WDR6-716-730		P49815-TSC2-1303-1317		Biotin		K	
L																									L	
M	Q6ZV50-RFX8-451-465		Q8IWR0-ZC3H7A-281-295		Q9HBT7-ZNF287-115-129		O95905-ECD-418-432		Q9UPY3-DICER1-596-610		Q6ZNG2-DBX2-166-180		Q14653-IRF3-135-149		Q9BSK1-ZNF577-395-409		Q6ZMS4-ZNF852-109-123		P49023-PXN-212-226		Q14653-IRF3-139-153		P07437-TUBB-410-424		M	
N																									N	
O	Q14184-DOC2B-78-92		Q92995-USP13-404-418		Q7Z5J4-RAI1-116-130		P68371-TUBB4B-410-424		Q92585-MAML1-262-276		Q8IWX8-CHERP-100-114		Q5SQN1-SNAP47-162-176		Q96MU7-YTHDC1-11-25		P43243-MATR3-191-205		Q99759-MAP3K3-2-16		P19484-TFEB-27-41		Biotin		O	
P																									P	

(Manuscript in preparation)

Supplemental Materials and Methods, Data, and Results

The following section pertains to additional experimental details from the main project and additional experimental efforts undertaken that were not presented in the article manuscript. Each section is grouped by research objective and includes subsections describing experimental methodologies, data, and results.

Supplemental Materials and Methods, Data, and Results:

Measuring 16E6-UBE3A interactions with holdup (HU)

Materials and methods

HU of 16E6 with UBE3A in different oligomerization states

Direct holdup (HU) of 16E6 with UBE3A was performed by a six-point titration series of 16E6-MBP bait saturated resin [0 – 10 μM] against solutions containing monomeric UBE3A or oligomeric UBE3A-MBP. The MBP tag of monomeric UBE3A was removed via cleavage by TEV protease prior to use in HU. The UBE3A protein solutions were produced and purified by Jia Wen (JW). Following saturation of the resin with the bait according to the same protocol described for the nHU assay (pgs. 46, 57), bait and control saturated resins were exposed to the respective protein solutions of UBE3A for 1 hour at room temperature with periodic gentle agitation. The unbound fraction of UBE3A was then removed via brief benchtop centrifugation and samples were prepared for SDS-PAGE by adding 4x SDS-PAGE sample buffer and boiling at 95° for five minutes (pg. 47, 56). Equal amounts of sample were loaded onto 12% SDS-PAGE gel for analysis via Coomassie staining. The gel was run at 90V until samples entered the resolving gel (approximately 10-15 minutes) then voltage was increased to 120V until samples were fully resolved. The optical density of the Coomassie stained bands were measured by the Amersham ImageQuant™ 800 (Cytivia) and densitometry calculations were performed in Fiji. The fitting of binding models was performed in QTiPlot (IONDEV SRL, Bucharest, Romania).

Data and results

In a pilot experiment prior to the development of the nHU spike-in assay, we used a HU-Coomassie assay to assess the binding activity of UBE3A in monomeric and oligomeric states. A titration series of biotinylated 16E6-MBP bait saturated resin was produced and incubated with a protein solution of either monomeric or oligomeric UBE3A to detect direct binding. A visible depletion of UBE3A (E6AP) is observed in both monomeric and oligomeric samples as 16E6 concentration increases, but it appears this effect is stronger in oligomeric samples than the monomeric (Fig. T9A). Binding curves reveal a partial activity in the binding of monomeric UBE3A to 16E6 (Fig. T9B) which would go on to correlate with finding from the nHU spike-in experiment with UBE3A (Fig. 3; pg. 71).

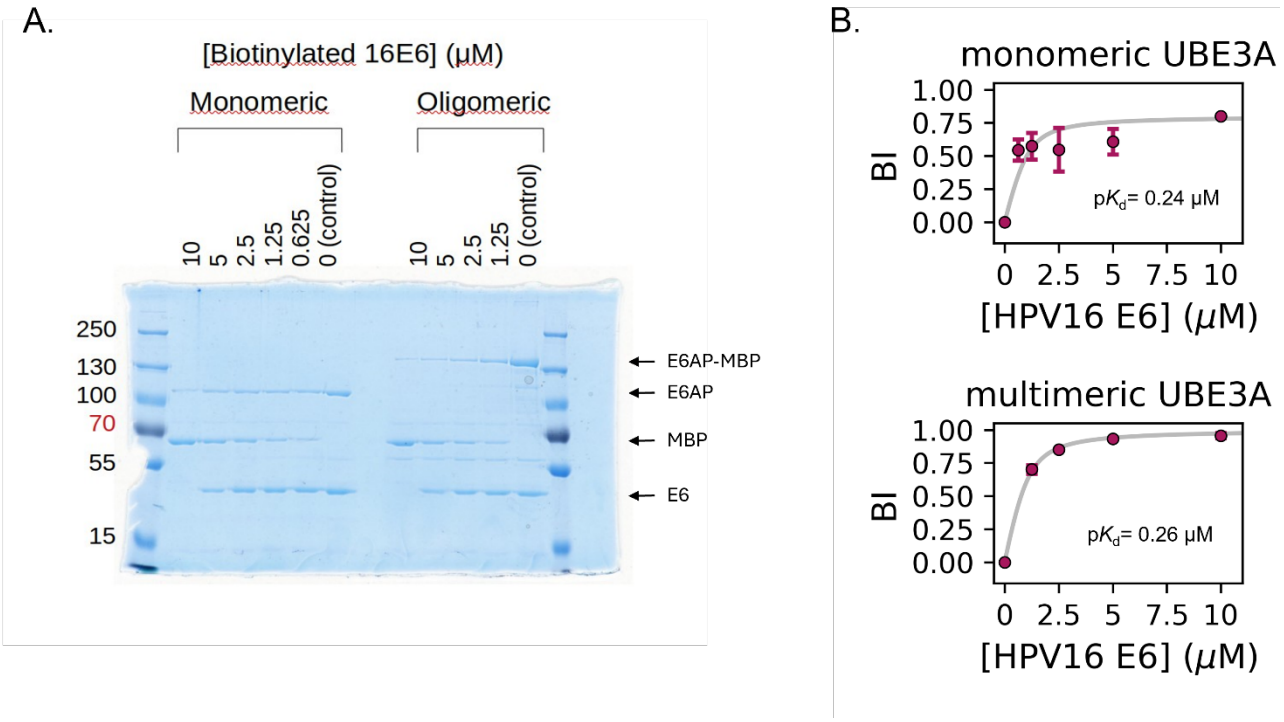


Figure T9 16E6-UBE3A binding is dependent on oligomeric state of UBE3A.

(A) Coomassie read-out of HU assay between 16E6 and UBE3A in different oligomeric states. Bands on the left-hand side of the gel correspond to the titration series performed with binding to monomeric UBE3A. Bands on the right-hand side correspond to the titration series performed with binding to oligomeric UBE3A. Pipetting error introduced to the [0.625 μM] point of oligomeric UBE3A resulted in it being discarded from analysis. The MBP tag of monomeric UBE3A was cleaved via TEV protease digestion prior to HU experiments accounting for the difference in molecular weight that is observed on the gel between the monomeric and oligomeric UBE3A. The presence of MBP and E6 bands on the bottom of the gel represents leakage of the bait from the resin during the experimental procedure. There inversely proportional relationship reflects the titration process used to create the correct bait concentrations in the titration series. (B) Fitting of the binding models demonstrates that while the apparent affinities (pK_{apps}) of the 16E6-UBE3A interaction remain consistent, partial activities are observed monomeric UBE3A.

Supplementary experimental details regarding 16E6 nHU-spike in with UBE3A

The final spike-in concentration of 50 nM monomeric/oligomeric UBE3A was determined by first assuming the minimal protein concentration of cellular TP53 and 16E6. Estimated cellular concentrations of TP53 are 0.06-0.5 μM[102] and cellular 16E6 expression is reported to be low so we assume a minimum 16E6 concentration of 150 nM[103,104]. Anticipating that our prepared cell extracts contain 100-fold less total protein than whole cells, we therefore assume a maximum 5 nM concentration of TP53. To ensure that full UBE3A activity can then be recovered, we supplement cell extracts with a 10x excess as compared to the estimated TP53 concentration.

Mapping the E6* interactome

Materials and methods

nHU with E6*

Using a biotinylated peptide corresponding to the amino acid sequence of 16E6* as bait, single point nHU experiments with HeLa and HaCaT extracts were performed to identify interactions of this 16E6 splice variant. The following peptide was synthesized by the IGBMC Peptide Synthesis Platform:

Biotin-LinkerTTds-MFQDPQERPRKLPQLCTELQTTIHDILECVYCKQQLLRREVY

Samples were sent for mass spec analysis by the IGBMC proteomics platform.

Data and results

To map the affinity interactome of 16E6 splice variant 16E6*, a biotinylated peptide corresponding to the 16E6* sequence was used as bait in nHU assays against HaCaT and HeLa cells. There has been much debate as to whether the E6* transcript is expressed as a protein, but numerous studies have reported biological activity and even its detection in cell lines[105]. Very few 16E6*-binders were fished in these nHU assays, with several of these being carboxylases and likely the result of interactions either with biotin on the surface of the resin or the resin itself[106]. However, the PDZ protein SNTB1 was detected above the binding threshold in the nHU assay against HaCaT cell lines. It's possible that E6* has a much more restrictive protein binding profile than that of its full-length counterpart E6[106].

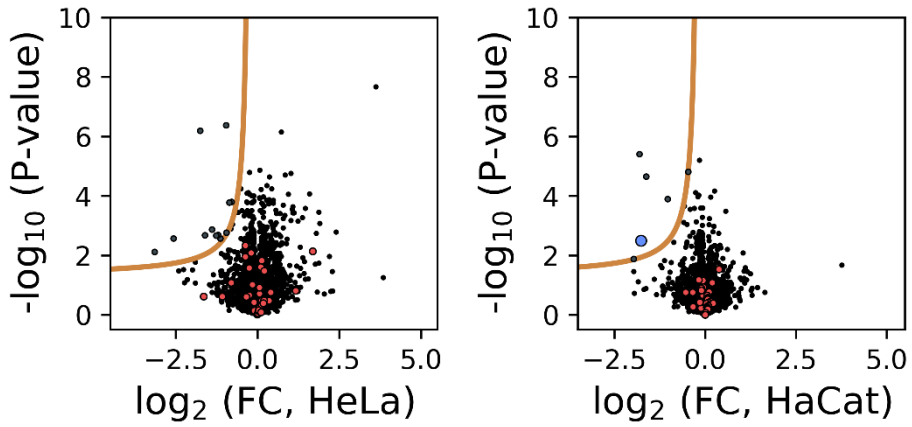


Figure T10 The affinity interactome of E6*. Single point nHU measurements using E6* peptide as bait against HeLa cells (left) and HaCaT cells (right).

Evaluating the 16E6-PTPN14 interaction and its cellular effects

Materials and methods

nHU spike-in with PTPN14 PY motif peptide

A nHU competition assay was performed using a single point concentration of 16E6 bait against HaCaT cell extracts supplemented with 100 μM (final concentration) of PTPN14 peptide corresponding to a PY motif sequence. The following peptide was synthesized by the IGBMC Peptide Synthesis Platform:

Peptide sequence- RHSAIIVPSYRPTPDYETVMRQMKRG

Positions 430-455

Experimental procedures for nHU spike-in with the PTPN14 peptide were carried out as described in the article manuscript (pg. 51, 61)

Monitoring PTPN14 in HEK293T-16E6 cell lines

Samples of HEK293T cells that were genetically manipulated via the sleeping beauty transposon-transposase system to introduce stable expression of 16E6 or 16E6 with a C-terminal PBM deletion (ΔPBM) were provide by Boglarka Zambo (BZ)[107]. The samples were analyzed via western blot according to established protocols (pg. 49, 59) to monitor PTPN14 levels.

Transfection of HEK293T-16E6 cell line with E7

The aforementioned cell line, henceforth referred to as HEK293T-16E6 and HEK293T-16E6ΔPBM, were provide by BZ for the use in transient transfection experiments. A plasmid for mammalian expression of 16E7 was provided by JW (Fig. T11). Transient transfection of HEK293T-16E6, HEK293T-16E6ΔPBM, and HEK293T-IRES (control) cells with 16E7 was preformed using the JetPrime™ lipofection reagent kit (Polyplus transfection, France). Prior to transfection a Poly-D-lysine treated (A3890401, Thermo Fischer Scientific, Waltham, Massachusetts) 24 well plates were seeded with 1x10^5 cells per well (Table T1). The JetPrime reagent kit was used as recommended by the manufacturer following the calculations below (Table T2). After the reagents were prepared, 16E7 DNA was added to the mix to the final desired concentration. Immediately following the addition of DNA, the mix was briefly vortexed and spun down before incubating for 10 minutes at room temperature. Then 50 μL of the appropriate transfection mix was added to each well. Once all DNA was dispersed, the plate was gently mixed and stored at 37°C with 5% CO_2. After approximately 4-5 hours the media in all wells of the plate was exchanged for fresh media to limit extensive cytotoxic effects from the transfection reagents. Plates were stored again at 37°C with 5% CO_2 for 24-48 hours. Cells were then harvested by removing the media, washing once with 1x PBS, and adding 4x SDS-PAGE sample buffer (pg. 57). Cells were then collected using a cell scraper to physically remove them from the plate surface. The recovered samples were boiled

at 95°C for 5 minutes, spun down briefly with a centrifuge, and in the case of exceptionally viscous samples further lysis was performed with a brief round of sonication (pulse: 0.1s on/0.1s off) prior to storage at -20°C until analyzed by Western Blot according to established protocols (pg. 49, 59)

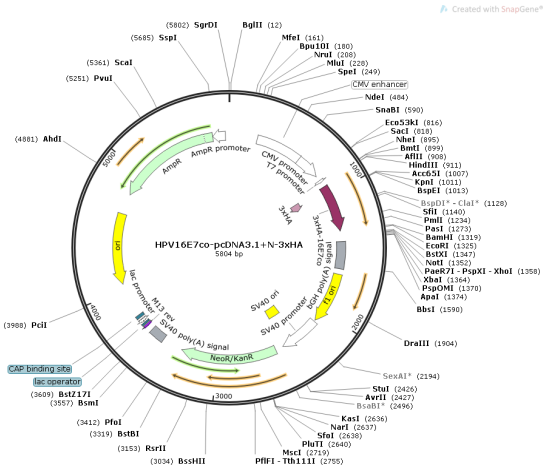


Fig. T11 16E7 plasmid used for transfection experiments.

Plasmid map of 16E7 vector used for transfection in HEK293T cell lines.

Table T1 Plate arrangement for transfection with E7

IRES		E6		E6dPBM	
3xHA	E7	3xHA	E7	3xHA	E7

Table T2 JetPrime transfection with E7

	DNA Conc.	DNA/ well (µg)	Dilution vol. (µl) /	well	JP	Buffer JP	DNA (µl)
3xHA control	0.64	0.5	50	24	24	1200	18.75
3xHA-E7	0.175	0.5	50	24	24	1200	68.5714

Data and results

Investigating 16E6-PTPN14 binding with nHU PY motif peptide spike-in

PTPN14 is known to interact with WW containing proteins, such as MAGI-1, through its PY motif[108]. The PDZ protein MAGI-1 is also an established interaction partner of 16E6[109]. If PTPN14 interaction with 16E6 is mediated indirectly through MAGI-1 (Fig. T12A) then by using the PY motif peptide in a nHU competition assay, the peptide should outcompete full length PTPN14 and saturate the WW domain of MAGI-1 thus disrupting association of both proteins with 16E6. Titrations of PTPN14 and MAGI-1 were performed to assess the capability of MAGI-1 to serve as an intermediary. This can be dependent on affinity or concentration. Assuming that there is no positive cooperativity, the affinity of the intermediary must be stronger than that of PTPN14. However, if the cellular abundance of the intermediary greatly exceeds that of PTPN14 it is possible to see a complete depletion of PTPN14 but not the intermediary upon 16E6 binding.

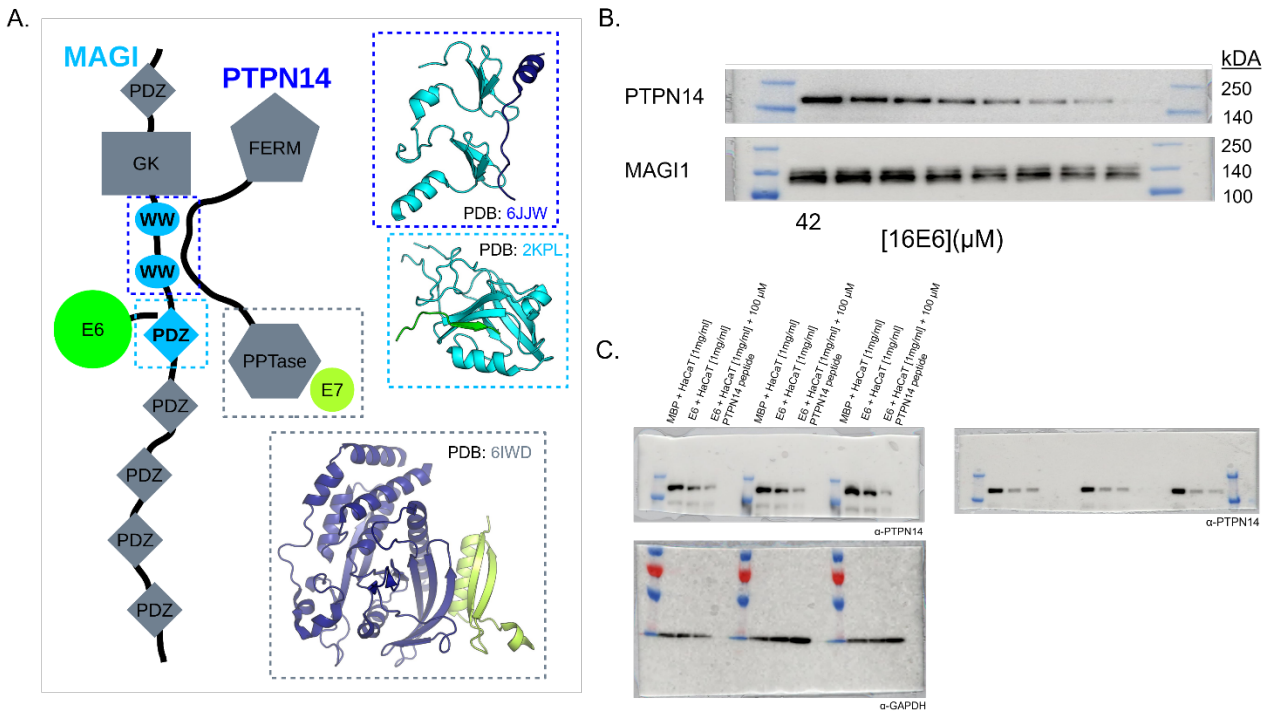


Figure T12 Evaluating MAGI-1 as an intermediary for 16E6-PTPN14 binding.

(A) Schematic of proposed modelling suggesting MAGI-1 facilitates indirect interaction between 16E6 and PTPN14. Boxed in dashed dark blue is the established interaction between WW domains of MAGI-1 and a PY motif found in PTPN14. Boxed in dashed cyan is PDZ mediated interaction between MAGI-1 and 16E6. Boxed in dashed grey is the interaction interface between E7 and PTPN14. (B) nHU titrations of PTPN14 (top) and MAGI-1 (bottom). Ruling out positive cooperativity and assuming simplest binding modes, the affinity of the intermediary for PTPN14 would need to be greater than that of the affinity of the 16E6-PTPN14 interaction. (C) WB of nHU spike-in experiment against PTPN14 (top left and right) and GAPH (bottom left; non-binding and load control). Samples were loaded in triplicate from left to right: MBP (non-binding control), 16E6 (binding control), 16E6 + PTPN14 peptide. No appreciable differences in binding could be viewed between the 16E6 sample and 16E6+PTPN14 sample (top right).

Evaluating cellular effects of 16E6 on PTPN14

To evaluate the cellular effects of 16E6 exerted on PTPN14, we monitored the protein level of PTPN14 in HEK293T-16E6, -16E6 ΔPBM, and -IRES cell lines. The HEK293T-16E6 ΔPBM cell lines expressed a C-terminal deletion of the PBM which could have indicated if the 16E6-PTPN14 interaction involved this interface and therefore an PDZ intermediary. The initial Western Blot showed a band at 130 kDa that corresponds to the predicted molecular weight of PTPN14 (Fig. T13A), however this antibody is reported to detect PTPN14 at 160 kDa which was observed continuously in other experiments using HaCaT cells. It's possible that these cells contain another isoform of PTPN14 but considering RNA transcript levels (nTPM) of

PTPN14 are reported to be up to twice as low as what is found in HaCaT and detection in HaCaT is already weak it is much more likely that this is a non-specific band. The western blot was repeated developing only the portion of the membrane corresponding to 160 kDa and the detected bands showed no difference in PTPN14 expression with 16E6 presence (Fig. T13B).

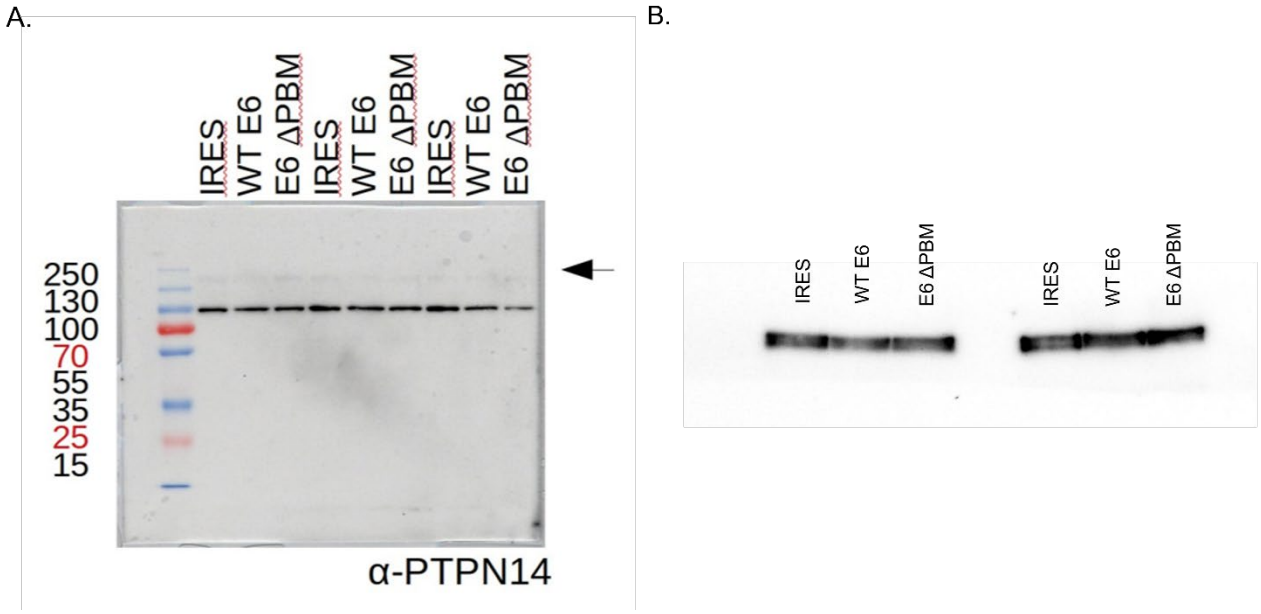


Figure T13 Impact of 16E6 expression on PTPN14 in HEK293T cells.

(A) Western blot against PTPN14 to evaluate the impact of 16E6 on PTPN14 levels. Samples of the following HEK293T cell extracts were analyzed -IRES (negative control; E6 null), -WT 16E6, and -16E6ΔPBM. A band at 130 kDa is likely non-specific binding of the antibody despite corresponding to the estimated molecular weight of PTPN14. A faint band is also seen at 160 kDa corresponding to the expected molecular weight based on previous experimentation. (B) The repeated western blot showed no difference between the PTPN14 levels in -IRES, -WT 16E6, and -E6ΔPBM.

Evaluating cellular effects of 16E6 and E7 on PTPN14 and USP13

Following the results of experiments evaluating the cellular impact of 16E6 on PTPN14 it was supposed that the presence of E7 may be necessary to model any E6 function. In the context of HPV pathologies, PTPN14 is one of the primary targets of HPV E7[110] and E6 and E7 have been describe to elicit combinatorial effects on cells[111]. Furthermore, there has been reported evidence that there may be a direct interaction between E6 and E7[112] which could justify the need for both parties to be present to display E6 functionality. To this end HEK293T-16E6 cell lines were transiently transfected via lipofection with E7 to produce E6/E7 expressing cells and monitor the effects on PTPN14 (Fig. T14A&B). Additionally, USP13 was probed to test for combined cellular effects of E6/E7 expression (Fig. T14C). The expression of E7 in transfected HEK293T-IRES (control), -16E6, and -16E6ΔPBM cell lines demonstrates the procedure was effective.

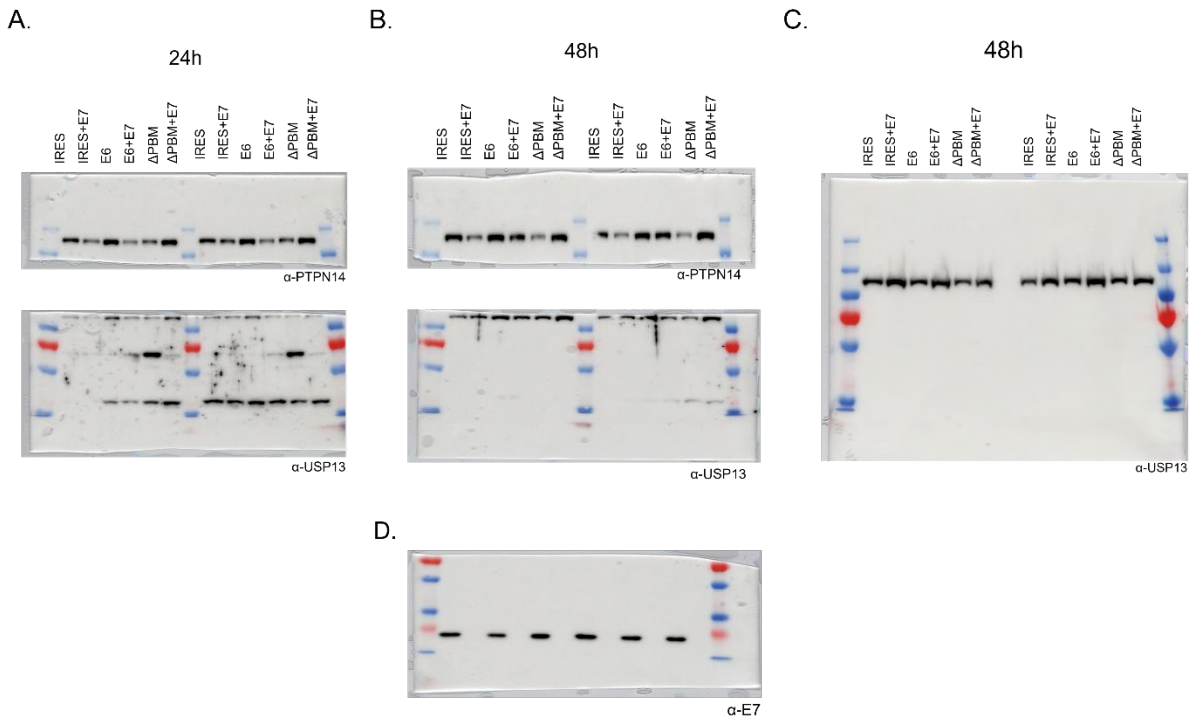


Figure T14 Impact of 16E6 and E7 expression on PTPN14 and USP13 in HEK293T cells.

(A) Western blot against PTPN14 (top) and USP13 (bottom) in HEK293T cells expressing 16E6 and 16E7 for 24 hours. IRES (negative control; E6 null), WT 16E6, and 16E6ΔPBM samples were either transfected with the 16E7 plasmid (+E7) or just the lipofection reagent. (B) Western blot against PTPN14 (top) and USP13 (bottom) in HEK293T cells expressing 16E6 and 16E7 for 48 hours. IRES (negative control; E6 null), WT 16E6, and 16E6 ΔPBM samples were either transfected with the 16E7 plasmid (+E7) or just the lipofection reagent. (C) Western blot against USP13 in HEK293T cells expressing 16E6 and 16E7 for 48 hours. IRES (negative control; E6 null), WT 16E6, and 16E6 ΔPBM samples were either transfected with the 16E7 plasmid (+E7) or just the lipofection reagent. (D) Western blot against 16E7 was performed to validate transfection.

Investigating 16E6 interactions with the ubiquitin proteosome system

Materials and methods

Knocking-down USP13, XIAP, and UBE3A expression in HeLa and HEK293T

siRNA experiments were performed using the JetPrime™ lipofection reagent kit and siRNAs against USP13 (8975, Dharmacon Horizon, Colorado, USA), XIAP (331, Dharmacon Horizon), and UBE3A (7337, Dharmacon Horizon) in HEK293T and HeLa cell lines. Transfections were carried out according to the protocol described above and the calculations below (Table T3&T4). Plates were incubated at 37°C and 5% CO_2 for 48 hours before harvesting cells (see protocol above) based on the protein expression results from siRNA experiments against UBE3A performed by other members of the lab.

Table T3 Plate arrangement for siRNA experiments with USP13, XIAP, and UBE3A

		1	2	3	4	5	6
	A	siC	siC	siC	siUBE3A	siUBE3A	siUBE3A
HeLa	B	siXIAP	siXIAP	siXIAP	siUSP13	siUSP13	siUSP13
	C	siC	siC	siC	siUBE3A	siUBE3A	siUBE3A
HEK293T-E6	D	siXIAP	siXIAP	siXIAP	siUSP13	siUSP13	siUSP13

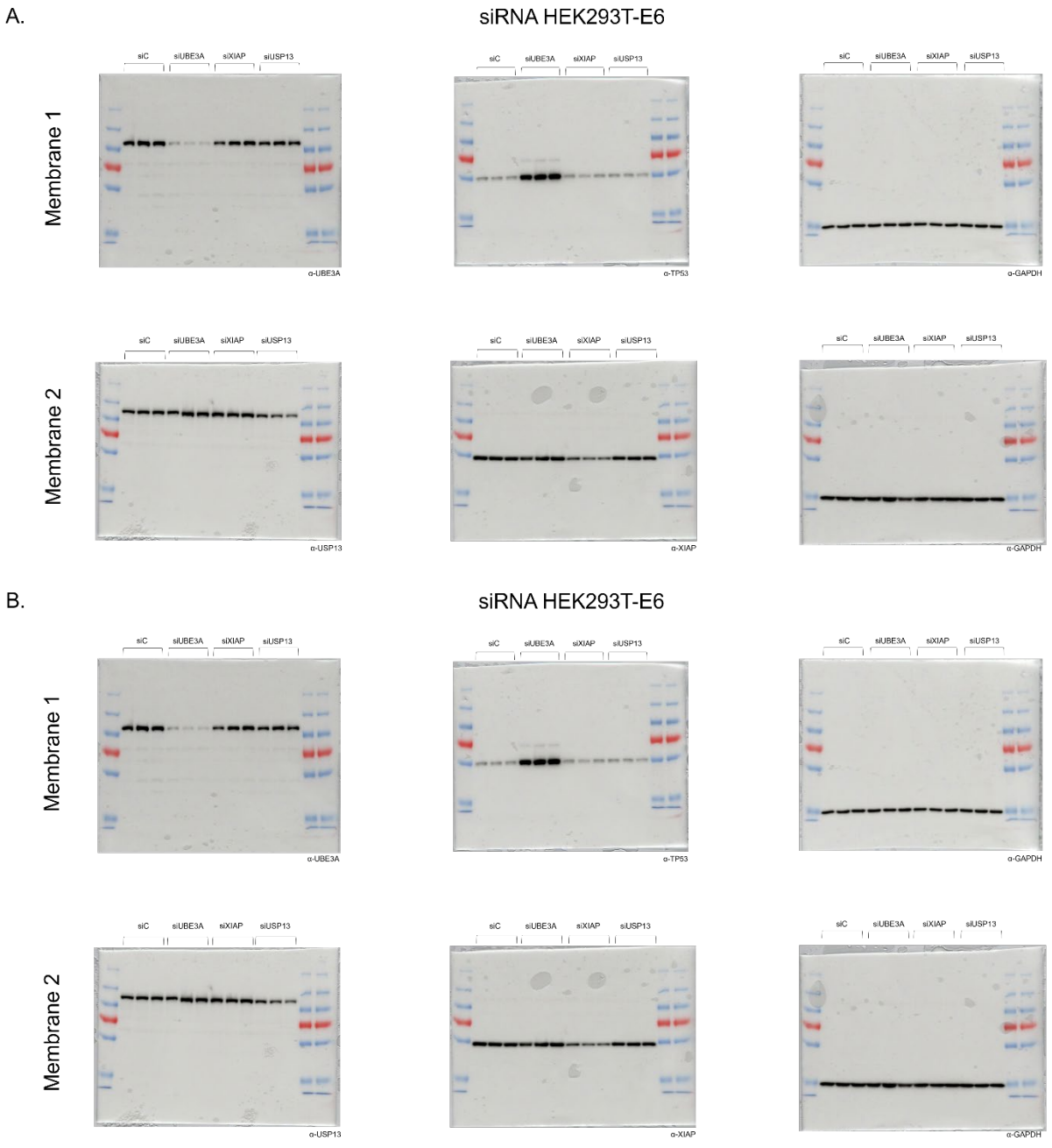
Table T4 JetPrime transfection with USP13, XIAP, and UBE3A siRNA

Samples	RNA Conc. (µM)	Well(s)	Dilution vol. (µl) / well	final conc (pmol)	JP	Buffer	RNA (µl)
siC	5	3	50	20	6	150	12
siXIAP	5	3	50	20	6	150	12
siUBE3A	5	3	50	20	6	150	12
siUSP13	5	3	50	20	6	150	12

Data and results

Investigating 16E6 interactions with the ubiquitin proteosome system

To understand the true extent to which 16E6 targets the ubiquitin proteosome system (UPS), siRNA experiments were performed against 16E6-binders with established roles in the UPS. This included ubiquitin ligases UBE3A and XIAP and the deubiquitinase USP13. Experiments were performed in HEK293T and HeLa cells. Protein levels were then monitored for each of these 16E6-binders and TP53 via Western Blot to assess the impact of silencing. As expected, the silencing of UBE3A resulted in a statistically significant increase in the expression of TP53 (Fig. T15C). Otherwise, the silencing of these 16E6-binders did not elicit any effect on any of the other 16E6-binders being monitored.



C.

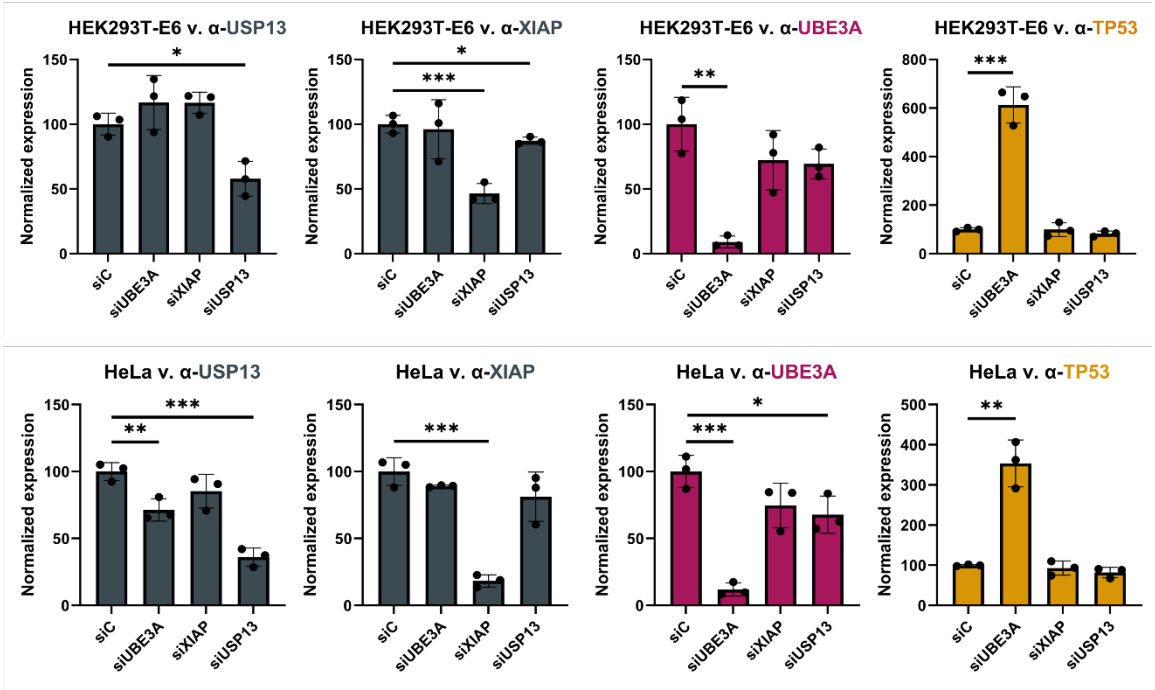


Figure T15 Investigating 16E6 interactions with the ubiquitin proteosome system (UPS).

(A) Western Blots monitoring USP13, XIAP, UBE3A, and TP53 levels of siRNA experiments performed in HeLa cells. (B) Western Blots monitoring USP13, XIAP, UBE3A, and TP53 levels of siRNA experiments performed in HEK293T cells. (C) Statistical analysis of siRNA experiments.

Evaluation of 16E6 inhibitors via nHU

Materials and methods

16E6 treatment with inhibitors and nHU

Assays for the therapeutic evaluation of 16E6 inhibitors follow the same principles of the traditional nHU with additional steps added to perform inhibitor treatment of the 16E6 proceeding its use as bait. Experiments are performed at a single point concentration comparing a non-binding control and a non-treated 16E6 binding control to samples of treated 16E6. Resin saturated with 5 μM 16E6-MBP was treated with 50-100 μM (10-20x) concentrations of inhibitor (dissolved in cell lysis buffer and up to 1% DMSO) for 24 hours at 4°C with agitation. Prior to nHU the resin was washed twice with 400 μL (20x) of nHU buffer to remove any free protein and/or inhibitor. Then, 10 μL of resin was aliquoted into 500 μL microcentrifuge tubes for nHU. The bait saturated resin was incubated with 40 μL (1:4 resin:analyte ratio) of cellular extracts [2 mg/mL] for 2 hours at 4°C with agitation. The unbound fraction was separated by centrifugation and analyzed via Western Blot (pg. 49, 59). Inhibitory activity was determined by probing against SN27 (negative control), PXN (positive control), UBE3A and TP53.

Data and results

Adapting nHU for therapeutic evaluation of E6 inhibitors

There is no current medical intervention for those already infected with HPV and current treatment options for related pathologies such as cutaneous and genital warts and cervical dysplasia don't target infection, but instead remove the infected tissue or treat them with cytotoxic chemicals.[113] While prophylactic vaccination has proven effective in preventing HPV related carcinogenesis, there have been challenges in its implementation including high cost which has contributing to a remaining burden of disease world-wide, especially in low income countries[81]. Targeting the LxxLL-binding pocket of E6 with small molecule inhibitors to prevent p53 degradation has emerged as a viable strategy for therapeutic intervention[114,115]. To this end we have adapted the nHU assay to evaluate E6 inhibitors through monitoring the binding activity of inhibitor treated 16E6 (Fig. T16).

These experiments were set up as proof of concept to demonstrate that nHU can be adapted to evaluate the inhibitory effect of small molecules. This is why no statistical analysis was performed on the results of the binding activities and the Western Blot experiments were not repeated for the monitoring of UBE3A binding. It is important to note that this particular use of the nHU assay is a shift to a pseudo-quantitative analysis of binding as the idea is to only show a decrease in overall binding. An interesting next step in development would be to use the nHU spike-in format to perform a series of assays to test for efficacy and off target effects. This would be done through using untreated 16E6 as a bait in nHU assay and using cell extracts that have been dosed with the inhibitor. Binding activities could be monitored again through Western Blot and following elution of the bait from the resin MS could be used to determine the amount of 16E6 reacted with the inhibitor. This could then be complemented through a reverse assay where UBE3A serves as the bait in a nHU against 16E6 expressing cellular extracts that have been treated with inhibitor. Using Western Blot to monitor 16E6 binding would indicate the efficacy of the inhibitor to target 16E6 in a cellular environment and block binding with UBE3A. To drive these developments this has become the full PhD project of Theo Juncker. I have continued to assist in the development of this technical application by assisting in the establishment of a model system for therapeutic evaluation by nHU. This is done using the inhibitor Nutlin which is a known antagonist of the MDM2-TP53 axis[116,117].

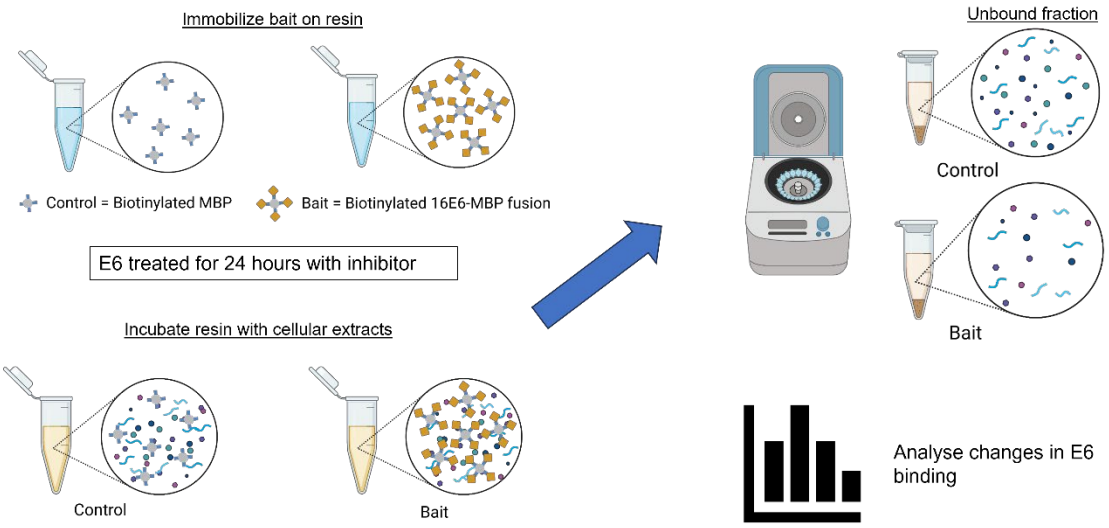


Figure T16 Employing nHU for the evaluation of therapeutics.

Schematic of the nHU assay for inhibitor treated E6. The 16E6 bait is first immobilized on the resin and subsequently treated with an excess of inhibitor. Prior to exposing the bait to preys from cell extracts, the resin is washed extensively to remove any excess free protein and/or inhibitor. Measurements are performed in a single point coupled to Western Blot where changes in binding compared to an untreated control are quantified by measuring the difference in signal.

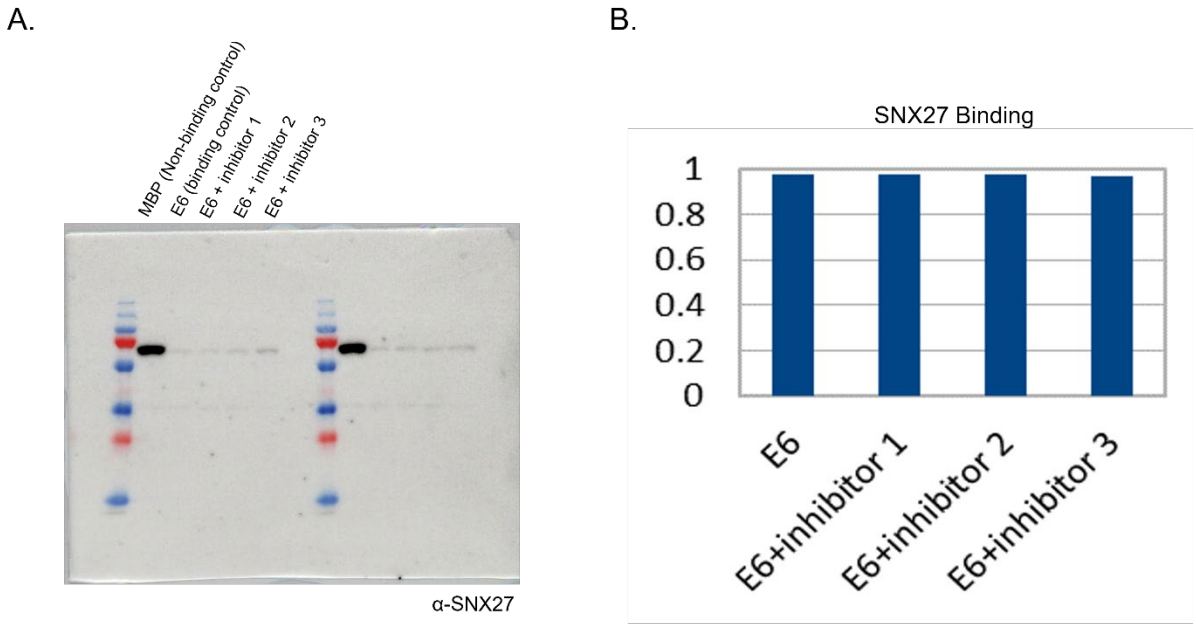
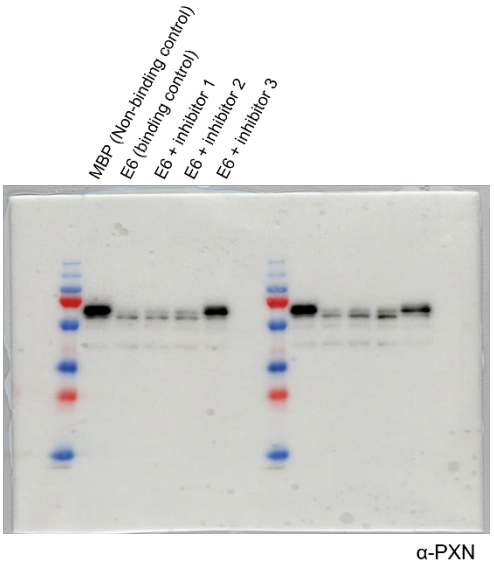


Figure T17 Monitoring 16E6 binding to SNX27 following inhibitor treatment.

As SNX27 interaction with 16E6 is mediated through the C-terminal PBM, it should remain unperturbed even in the presence of the inhibitor. (A) Replicate Western Blot of nHU inhibitor experiment against SNX27. From left to right MBP represents a non-binding control, E6 corresponds to untreated 16E6 and represents a binding control, E6 + inhibitor 1-3 represent the binding activities of 16E6 following treatment with the respective inhibitor. (B) Signal quantification of 16E6 binding. This demonstrates that the inhibitor does not interfere with 16E6 binding to non-LxxLL mediated partners.

A.



B.

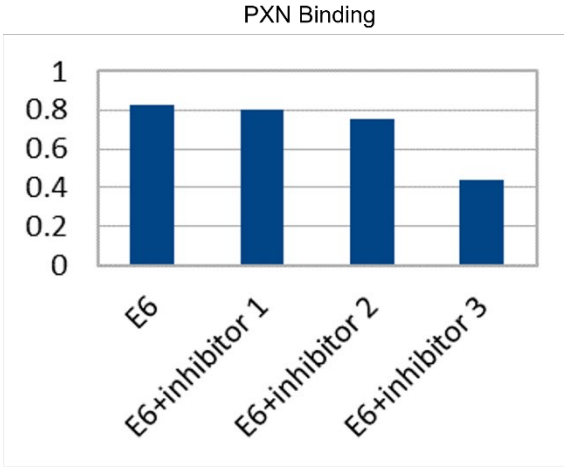


Figure T18 Monitoring 16E6 binding to PXN following inhibitor treatment.

As PXN interaction with 16E6 is mediated through the LxxLL binding pocket therefore 16E6 binding to PXN should be inhibited in the same manner as UBE3A if the inhibitor is effective. (A) Replicate Western Blot of nHU inhibitor experiment against PXN. From left to right MBP represents a non-binding control, E6 corresponds to untreated 16E6 and represents a binding control, E6 + inhibitor 1-3 represent the binding activities of 16E6 following treatment with the respective inhibitor. (B) Signal quantification of 16E6 binding. This demonstrates a possible decrease in binding to PXN following 16E6 treatment with inhibitor 3.

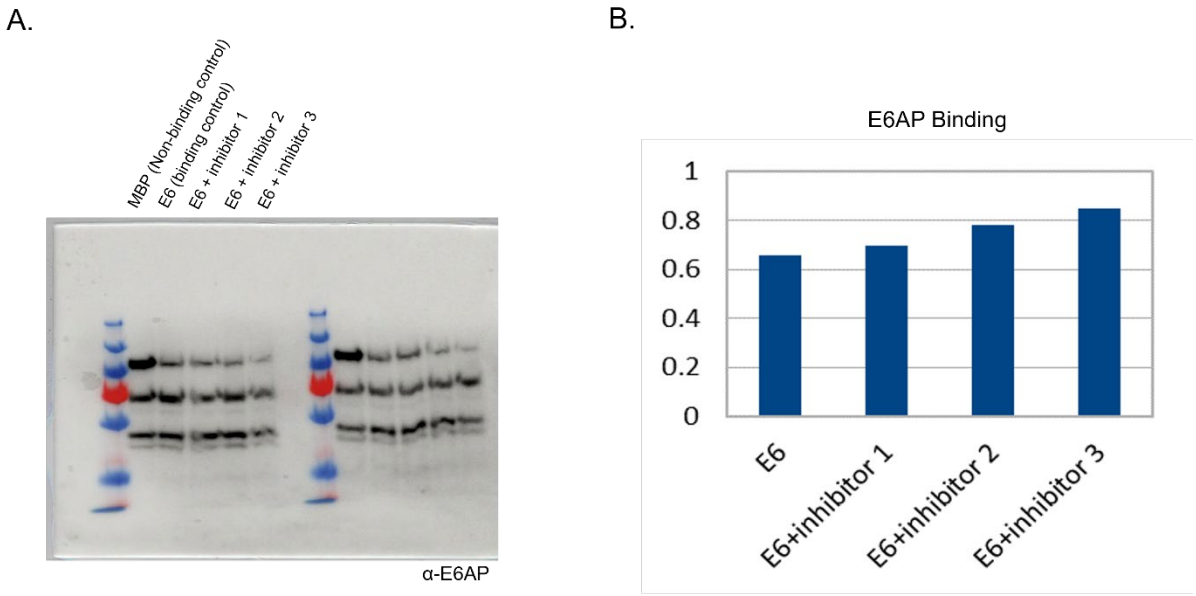


Figure T19 Monitoring 16E6 binding to UBE3A following inhibitor treatment.

Monitoring 16E6 binding activity to UBE3A following inhibitor treatment represents direct measurement of the desired therapeutic effect. (A) Replicate Western Blot of nHU inhibitor experiment against PXN. From left to right MBP represents a non-binding control, E6 corresponds to untreated 16E6 and represents a binding control, E6 + inhibitor 1-3 represent the binding activities of 16E6 following treatment with the respective inhibitor. Extensive artifact on the membrane makes it difficult to conclude reasonable results. (B) Signal quantification of 16E6 binding. While the binding activities show a slight increase with each inhibitor these are most likely due to artifact from the membrane and do not resemble true results.

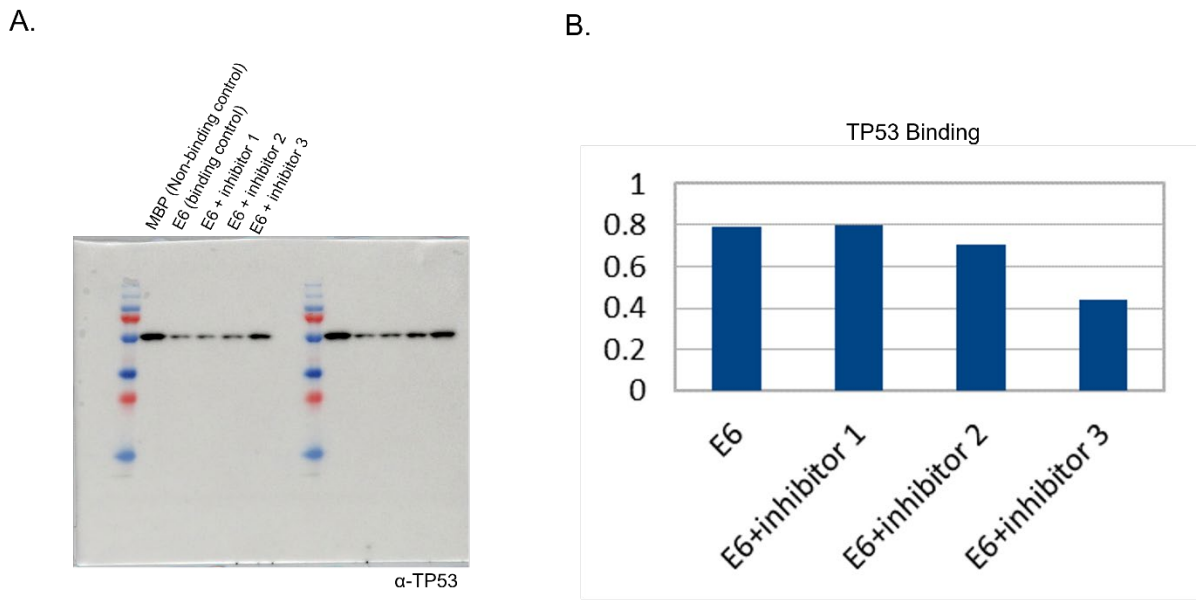


Figure T20 Monitoring 16E6 binding to TP53 following inhibitor treatment.

Monitoring 16E6 binding activity to TP53 following inhibitor treatment represents direct measurement of the desired therapeutic effect. If binding to UBE3A is successfully inhibited, TP53 binding should also be disrupted. (A) Replicate Western Blot of nHU inhibitor experiment against PXN. From left to right MBP represents a non-binding control, E6 corresponds to untreated 16E6 and represents a binding control, E6 + inhibitor 1-3 represent the binding activities of 16E6 following treatment with the respective inhibitor. Extensive artifact on the membrane makes it difficult to conclude reasonable results. (B) Signal quantification of 16E6 binding. This demonstrates a possible decrease in binding to TP53 following 16E6 treatment with inhibitor 3 which correlates with what was observed for our internal control PXN.

<u>Systematic AlphaFold modeling of 100 16E6-binders</u>

Materials and methods

AlphaFold3 predictions

AlphaFold3 was used to systematically model all 100 16E6-binders identified through nHU assays. Predictions were made and analyzed according to the procedure described in the manuscript (pg. 51, 61).

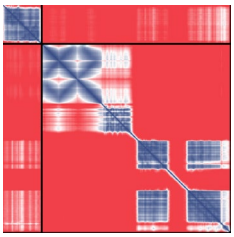
Data and Results

Evaluation of AlphaFold models of 16E6-binders using Predictomes

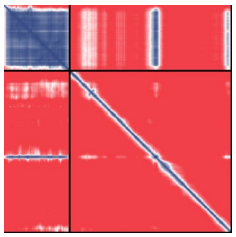
To try and gain insights into 16E6 interaction interfaces mediating contacts with 16E6-binders, structural models of 16E6 in binary complex with all 100 16E6-binders was performed. All predicted AlphaFold models were evaluated using the Predictomes analysis tool. Models with solved crystal structures such as UBE3A (4XR8), SCRIB (8B82), and IRF3 (6SJA) were used as benchmarks to further validate the predictions, but it's important to note that this could also reflect overtraining of the model. Predicted Aligned Error (PAE) plots produced by the Predictomes software are displayed on the subsequent pages. The PAE estimates the expected positional error for each residue in a predicted protein as if they were aligned to the corresponding residues of the true structure and these plots allow assessment of confidence in the relative positions and orientations of different parts (domains) of the prediction. This value has often been extended to evaluate models containing two monomeric protein molecules to infer binding interfaces based on the confidence in the position and orientation of domains. The PAE is represented on a plot where the (X, Y) coordinates represent the predicted error at residue x if the predicted and true structures were aligned on y. A low PAE value, which is represented by dark coloring, indicates high confidence in the predicted positions and orientations while a high PAE value, which is represented by light coloring, indicates low confidence and uncertainty. The Predictomes PAE plots show the PAE of each respective component of the model and how they relate to each other. Blue colored squares in the matrix represent predictions in which at least one pair of interfacial residues is closer than 5Å and meets minimum confidence criteria (PAE < 15 and pLDDT > 50) and are shown on a red background. Grey plots represent protein pairs where no residues meet the above interaction criteria and white plots indicate protein pairs where no prediction was made, usually because the proteins exceed ~3600 amino acids. No white or grey plots were produced in our analysis meaning interaction interfaces were predicted for all 16E6-binders. The top left quadrant, which is sectioned by perpendicular black lines represents the prediction of 16E6. The bottom right side of the plot represents the modelled 16E6 binder. Sections adjacent to the 16E6 plot, across the top and left side of the plot represent the 16E6 residues from C-term to N-term, indicating the potential location of contacts where dark blue squares indicate highly confident predictions.

Predictomes Analysis

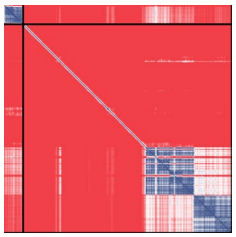
ACAP2 – Q15057



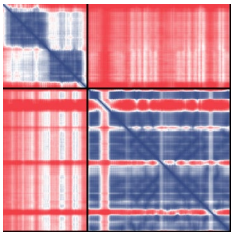
CDC42EP1 – Q00587



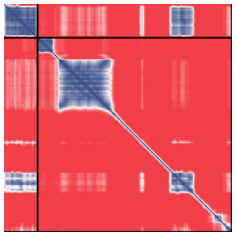
CRYBG1 – Q9Y4K1



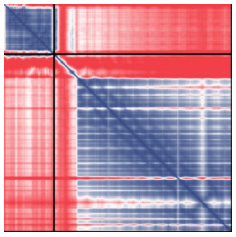
ADO – Q96SZ5



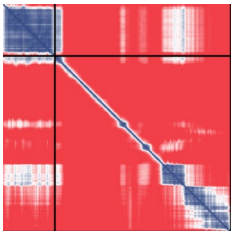
CHERP – Q8IWX8



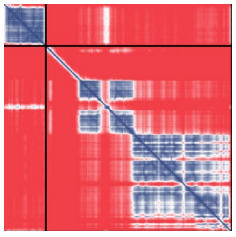
CTNNBL1 – Q8WYA6



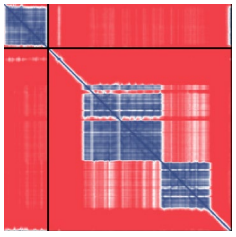
AJUBA – Q96IF1



CLPB – Q9H078



DDX3X – O00571

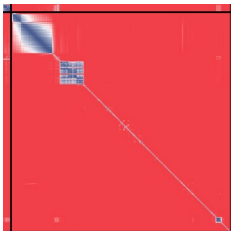


ANK3 – Q12955

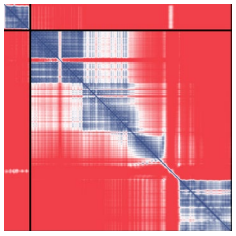
COPA – P53621

DICER1 – Q9UPY3

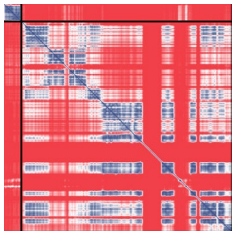
APOBEC3B – Q9UH17

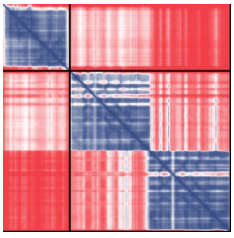


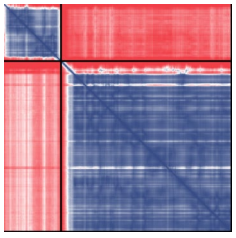
CPVL – Q9H3G5

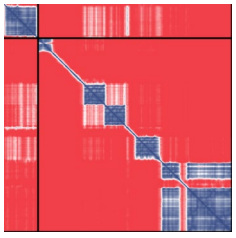


DLG1 – Q12959

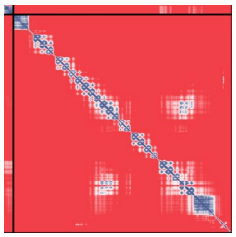




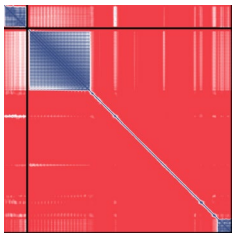




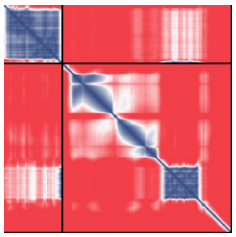
DMD – P11532



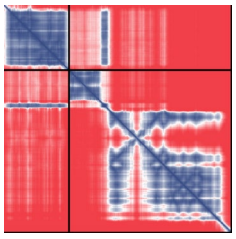
ERBIN – Q96RT1



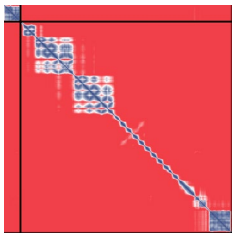
GOPC – Q9HD26



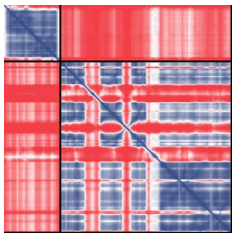
DNAJA1 – P31689



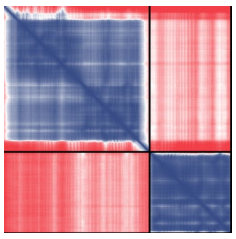
EVPL – Q92817



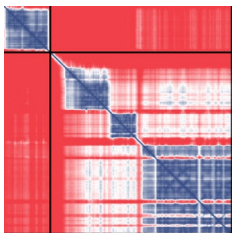
GSDMD - P57764



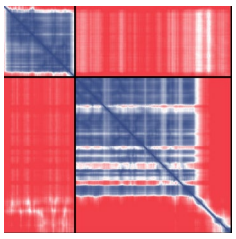
DYNLL1 – P63167



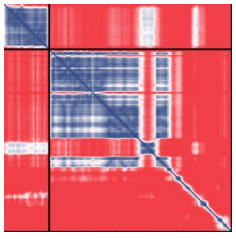
F2 – P00734



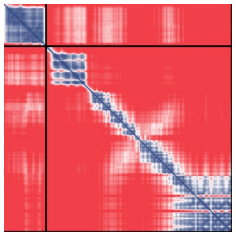
GYG1 – P46976



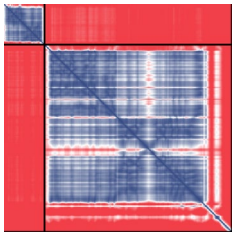
ECD – O95905



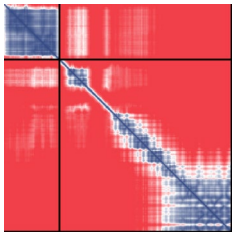
FBLN1 – P23142



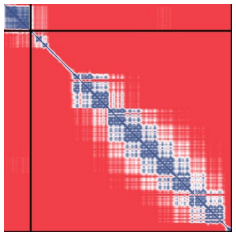
GYS1 – P13807



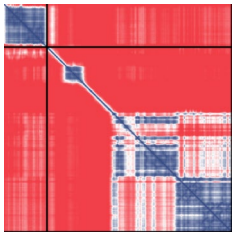
EFEMP1 – Q12805



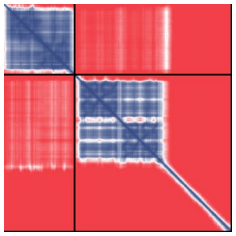
FNDC3A – Q9Y2H6



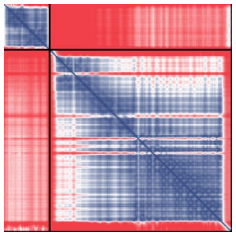
HBS1L – Q9Y450



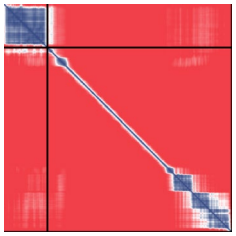
HNRNPA2B1 – P22626



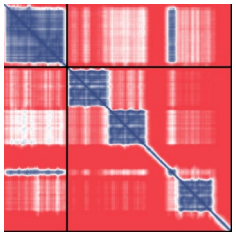
KLHL22 – Q53GT1



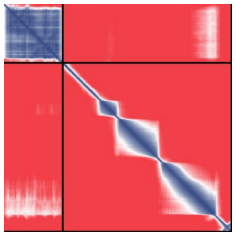
LIMD1 – Q9UGP4



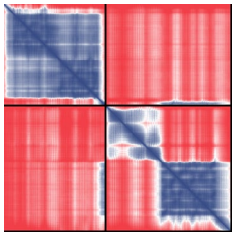
HNRNPF – P52597



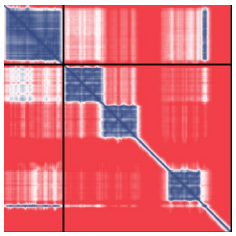
KRT15 – P19012



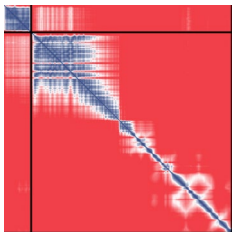
LIN7C – Q9NUP9



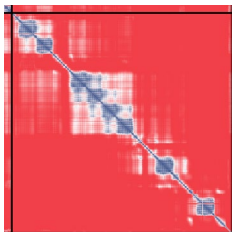
HNRNPH1 – P31943



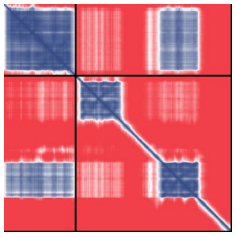
LAMB3 – Q13751



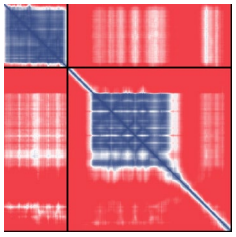
LRP1 – Q07954



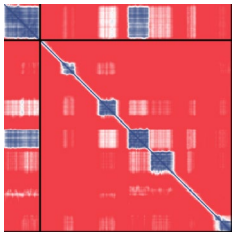
HNRNPH3 – P31942



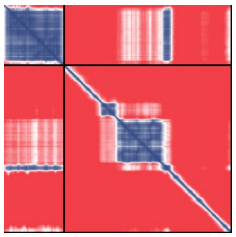
LDB1 – Q86U70



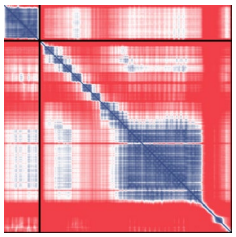
MATR3 – P43243



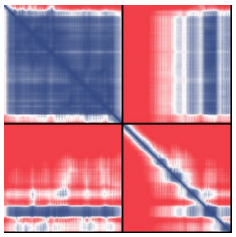
KHDRBS1 – Q07666



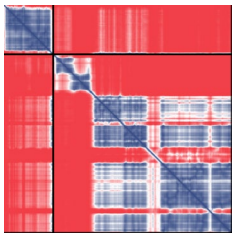
LDLR – P01130



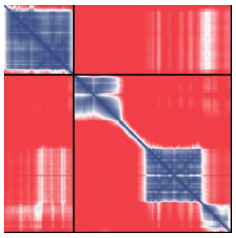
MCFD2 – Q8NI22



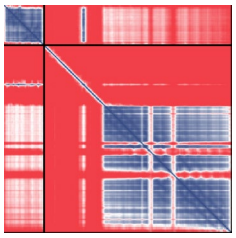
MPP7 – Q5T2T1



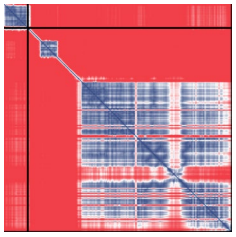
NUDCD3 – Q8IVD9



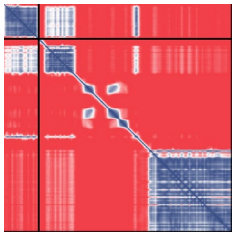
PKP1 - Q13835



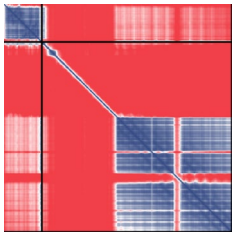
MSH6 – P52701



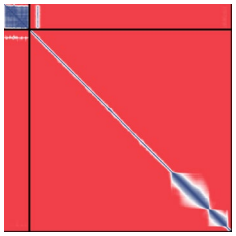
OSBPL3 – Q9H4L5



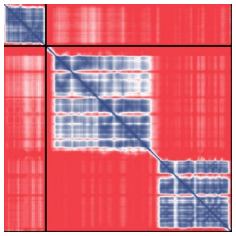
PKP3 – Q9Y446



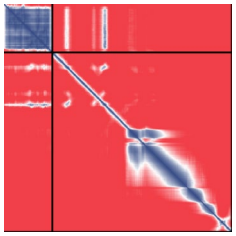
MTUS1 – Q9ULD2



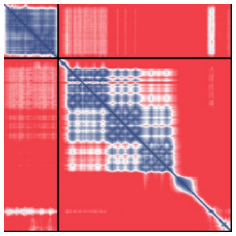
P3H2 – Q8IVL5



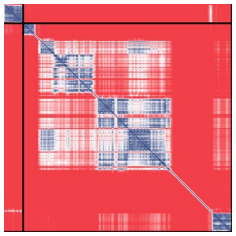
POF1B – Q8WVV4



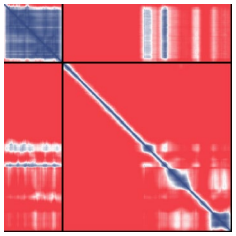
NECTIN4 – Q96NY8



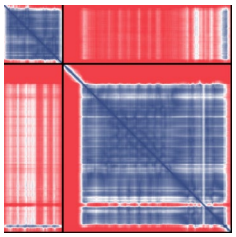
PARP4 – Q9UKK3



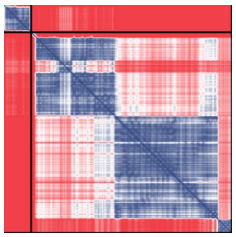
PPHLN1 – Q8NEY8



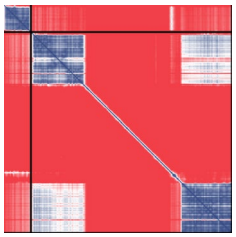
NFS1 – Q9Y697



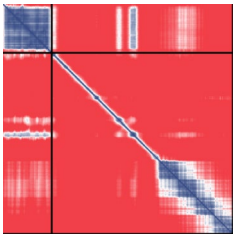
PC – P11498



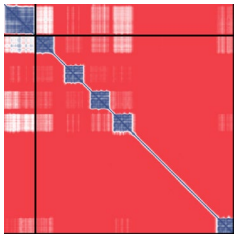
PTPN14 - Q15678



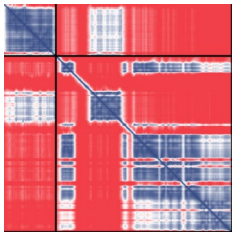
PXN – P49023



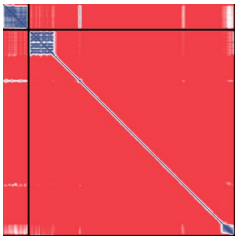
RBM12B – Q8IXT5



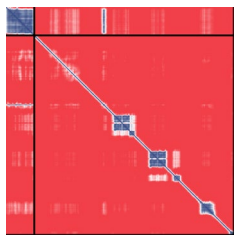
SNTB2 – Q13425



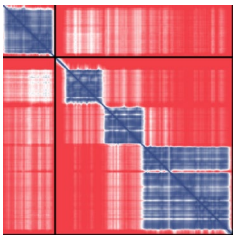
RAB11FIP1 – Q6WKZ4



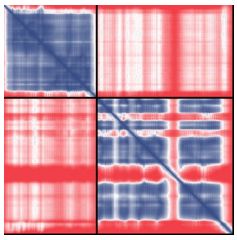
RBM6 – P78332



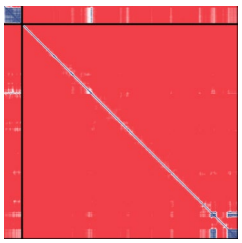
SNX27 – Q96L92



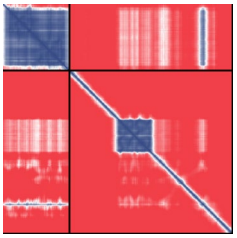
RAB20 – Q9NX57



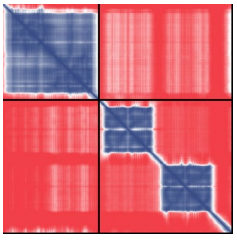
RAI1 – Q7Z5J4



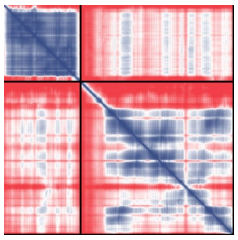
RBFOX2 – O43251



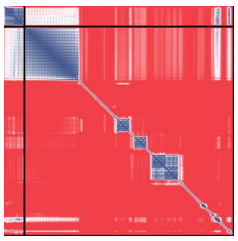
SRSF9 – Q13242



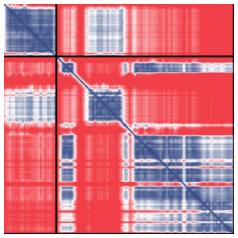
RCN2 – Q14257



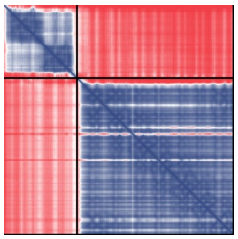
SCRIB – Q14160



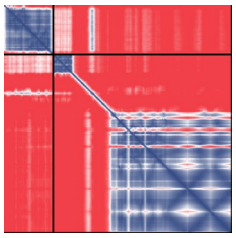
SNTB1 – Q13884



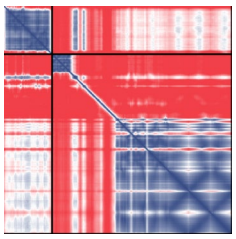
THOC6 – Q86W42



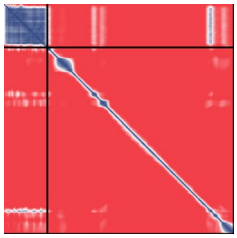
SNX33 – Q8WV41



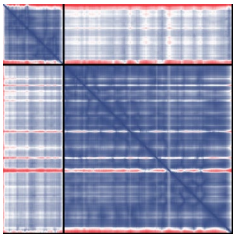
SNX9 – Q9Y5X1



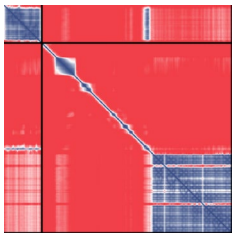
SPTY2D1 – Q68D10



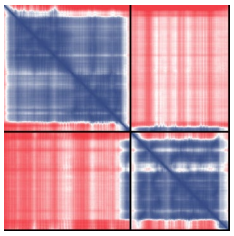
TUBB – P07437



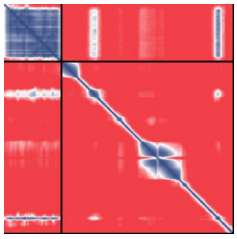
STRN3 – Q13033



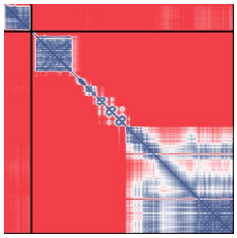
TAX1BP3 – O14907



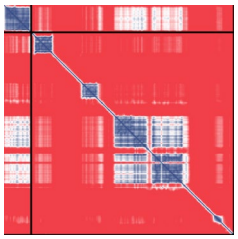
TFEB – P19484



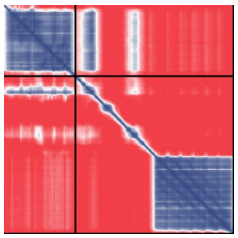
THBS1 – P07996



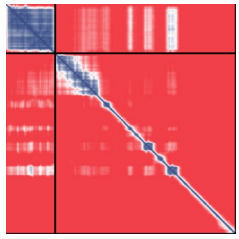
TJP2 – Q9UDY2



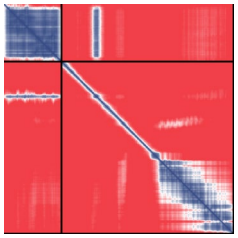
TMOD3 – Q9NYL9



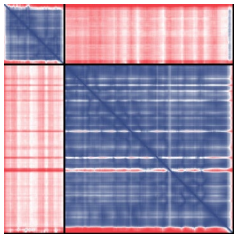
TRAFD1 – O14545



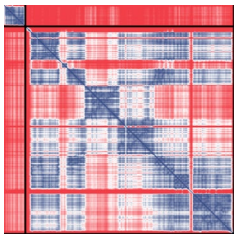
TRIP6 – Q15654



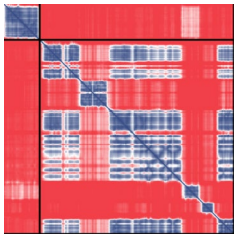
TUBB4B – P68371



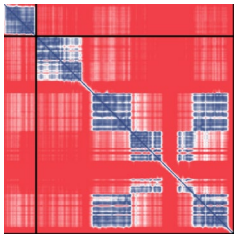
UGGT1 – Q9NYU2



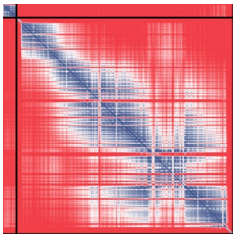
USP13 – Q92995



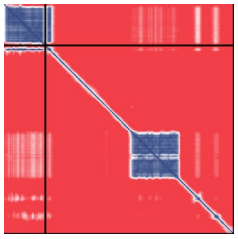
USP15 – Q9Y4E8



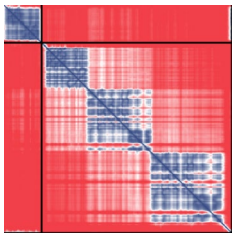
USP9X – Q93008



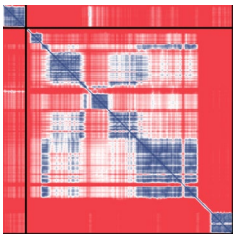
YTHDC1 – Q96MU7



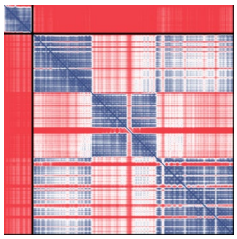
VCP – P55072



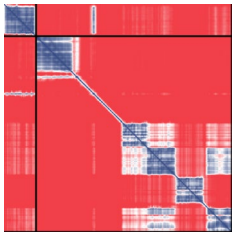
YTHDC2 – Q9H6S0



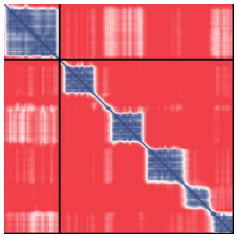
WDR6 – Q9NNW5



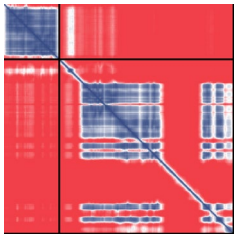
ZC3H7A – Q8IWR0



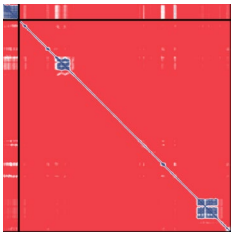
XIAP – P98170



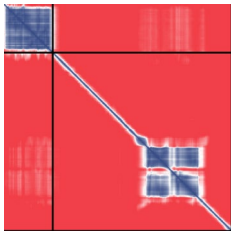
ZC3HC1 - Q86WB0



YLPM1 – P49750



ZNF326 - Q5BKZ1



Digital data provided in shared drive:

-MS data sets and affinity calculations for nHU experiments

^nHU 16E6-MBP bait vs. HaCaT and HeLa cell lines

^nHU C-terminal 16E6 PBM peptide bait vs. HaCaT and HeLa cell lines

^nHU 16E6* splice variant peptide bait vs. HaCaT and HeLa cell lines

-HU data from LxxLLome peptide panel

-BLI data from selected LxxLLome peptides

-Quantifications of nHU-WB and cellular experiments

-WB membranes from all nHU and cellular experiments

Discussion and Future Perspectives

Discussion and Future Perspectives

Discussion

Until this point several independent detailed discussions of the many facets of this project have been presented within the confines of manuscripts embedded within this dissertation— the advantages of quantitative interactomic approaches, how the intricate interactomic profile of E6 contributes to carcinogenesis, and the interpretation of the 16E6 affinity interactome. Here we will revisit these topics within the framework of the entire thesis project, integrating the concepts and further dissecting the results of this study to assess them for their applicability across technical and biological domains.

Studying host-pathogen interactions with quantitative interactomics

The virally encoded HPV oncoprotein E6 has garnered extensive interest in the context of interactomic studies as it not only serves as an excellent model system for understanding the relationship between interactomes and phenotypes, but it is also of high clinical importance due to its implications in the development of cervical cancers. Extensive interactomic studies of E6 has erected a solid foundation of validated interactions that are key drivers of pathogenesis, such as p53 degradation mediated by the 16E6-p53-E6AP complex and the targeting of PDZ containing partners like SCRIB, MAGIs, and SNX27. However, these studies are subject to the same limitations of all qualitative interactomic studies (pg. 13; pg. 16) and therefore overlap among entire E6 interactomes reported, aside from this subset of interactions, is low. Ambiguity surrounds many reported novel interactions of E6 as their detection can vary from method to method depending on the sensitivity of the technique being used and there are not functional studies readily available to confirm their presence in vivo. This led to the assumption that likely many of these are weak interactions, but this does not mean that they should be discounted considering weak interactions play crucial roles in many cellular processes[118] and likely exert binding regulatory effects in the crowed environment of the cellular milieu[119]. The capacity of nHU to detect both weak and transient interactions helped justify the addition of yet another E6 study to the repertoire.

Upon analyzing the 100 16E6-binders identified through nHU assays, our suspicions were confirmed through the identification of several weak interactions (Fig. 1B, pg. 63; Table 1, pg. 64). Among these APOBEC3B, a cytidine deaminase, has been extensively implicated in both HPV infection[120,121] and carcinogenesis[122], the ubiquitin ligase XIAP has been linked to HPV mediated regulation of apoptosis and has been implored as a prognostic indicator of cervical cancer[123], and the DNA mismatch repair protein MSH6 indicates increased risk for cervical cancer when the gene carries germ-line mutations[124] (Fig. 2C, pg. 69). However, to our surprise, we also discovered novel 16E6-binders, such as PTPN14 and USP13, which possessed affinities comparable to that of interactions between 16E6 and UBE3A or SNX27, which are typically regarded as strong interactions of E6. While both proteins have

established roles in HPV pathologies, with PTPN14 as a degradation target of the E7 oncoprotein and reports of USP13 upregulation specific to HPV (+) cervical cancers, neither have been previously identified by E6 interactomic studies. A potential reason for this could be that the interactions are transient and therefore not captured by most interactomic methods. In this case, however, one could expect that more recent proximity labelling studies of 16E6 would have reported these interactions[125]. One possibility is that this could be due to proteomic differences between the cell lines used for experimentation. This is a commonly shared blind spot of all interactomic methods. The number of detectable interactions is dictated by the number of preys present in the sample and when using cell-based methods this means the proteome determines the possible interactome. While this is seen by some as an advantage leading to the detection of only biologically relevant interactions, it creates problems in continuity in that the proteomes, even of the same cell lines, can vary from laboratory to laboratory thus contributing to low overlap among studies. Additionally, as mentioned we can only see as far as our MS can detect. It's possible that interactors from both the BioID study and our nHU study appear below detection thresholds of the other. This illustrates the difficulty of relating findings across experimental methods when there is no common unit of measurement. Finally, it is worth noting that the nHU assay completely disrupts cellular compartmentalization in aims of obtaining all possible physical interactions that can occur within a cell, while BioID maintains this compartmental integrity to rule out any non-biologically relevant interactions. While this could be another possible explanation for the discrepancies multiple studies have shown E6 localization to be some combination of cytoplasmic and nuclear meaning it is likely found relatively ubiquitously throughout the cell[126].

Key findings from mapping the quantitative 16E6 host-affinity interactome

Analysis of the 16E6 nHU results involved implementation of a new approach for identifying key interactions. As opposed to turning to functional enrichment analysis, such as GO or Pathways (KEGG), we implored a mathematical modeling approach to detect shifts in binary protein complex formation in response to proteomic variation. This represents a novel data informed molecular fishing approach to identifying key interactions from interactomic experiments. In previous sections I discussed how in the context of HPV infection complexomic analysis reflects the critical pathological event of p53 degradation during infection and how viewing shifts across different stages of infection can reveal temporally critical interactions or interactions that are consistently pertinent to the maintenance of the viral lifecycle (pg. 77). Here, I would like to elaborate on the 16E6-binders PTPN14 and USP13 that were identified through complexomic calculations as potential critical targets during pre- and post-transformation stages respectively.

In addition to their position on the complexomic spectrum PTPN14 and USP13 possess established roles in HPV pathologies suggesting that it is likely these are true interactors of 16E6. The phosphatase PTPN14 is targeted for degradation by the HPV E7 oncoprotein impacting the downstream HIPPO Pathway

which regulates cellular proliferation and in the context of HPV contributes to tumorigenesis[127]. Whereas USP13 is a deubiquitinase implicated in various diseases[128], but is most importantly associated specifically with HPV (+) cervical cancers[129]. This made them interesting candidates for further biochemical and functional characterization to try and establish possible mechanisms for binding and disease. Neither proteins possess PDZ domains or putative LxxLL motifs, but nHU spike-in assays to identify potential interaction interfaces indicated that both proteins likely interact through the 16E6 C-terminal (Fig. 6A, pg. 74). Further experimentation using the 16E6 PBM as bait in nHU experiments demonstrated the loss of both PTN14 and USP13 binding. This could be further affirmation of a direct interaction in that their binding would have been retained if connection was established through a PDZ intermediary, but we can't discount that it's possible that the loss of a secondary interaction site is responsible for the loss of interaction. Further experiments using the 16E6ΔPBM would help towards clarifying this. Additionally, the results of the performed cellular experiments utilizing HEK293T-16E6 cell lines transfected with 16E7 should be revisited. There appears to be a slight stabilizing effect of 16E6 on PTPN14 and 16E6 and 16E7 on USP13 which should be confirmed with statistical analysis. Should these be positive results, they resemble the complex interplay of the E6 and E7 networks and necessitate the establishment of more sophisticated functional models for further study. In general, lack of adequate functional models has presented a challenge in validating discovered 16E6-binder interactions and their study. To accurately understand the role of E6 in HPV mediated cellular rewiring the interconnected interaction network needs to remain intact, and it is often disrupted or dysregulated in our models. It's also possible that many individual interactions will exert negligible effect, but once in the cumulative context of their network their effects can be observed. The functional models we employed were likely too simplistic to address the questions being asked and were not accurate representations as they involved overexpression of 16E6 and E7 which we know does not accurately reflect their typical cellular concentrations. This is a technical expertise outside the scope of myself and our lab, thus marking an excellent starting point for collaborations.

Intricately interwoven interactomes

It is important to remember that the 16E6-host affinity interactome represents one facet of a much greater and more complex network of intricately interwoven interactomes. The dynamic relationship of E6 and E7 has been discussed at length (pg. 33) and represents the interplay of intraviral interactions. However, even if we consider all of the interaction networks of the pathogenic entity, we must remember that it is functioning within an organism-wide system of intercalated interaction networks. To this end it is necessary to take on even broader system wide perspectives considering the role of HPVs in the context of host components such as the microbiome[130,131] and seeking to understand how cross-talk between all these layers leads to transitions from normal cellular functioning to the establishment of

disease. Artificial intelligence (AI) will become a powerful tool in modelling these events and as previously discussed, the data obtained from quantitative interactomic data will be pivotal in these applications (pg. 80).

Future Perspectives

In my mind there are endless possibilities for the continuation of this project, and I personally wish I could be the one to tend to them. These developments fall into two categories: biological study and technical development. I have discussed below some of my ideas with biological projects continuing the theme of HPV driven oncogenesis and technological developments focused on adapting nHU for more extensive uses. It is my hope that my body of work has left a solid foundation upon which others will either continue to build our understanding of HPV oncoprotein interactions or will result in the adaptation of approaches and methodologies for solving other complex problems in systems biology.

Functional analysis on novel 16E6-binders

The most direct continuation of my project would be to perform functional analysis of novel 16E6-binders found in our study, particularly PTPN14 and USP13. A first step towards this would be to validate the interaction interface. This could be done by creating various truncations of the tagged 16E6-binders and expressing them in cell lines. Repeating nHU experiments using different cell lysates expressing the different truncations would allow for identification of the region where the binding interface is located. Additional follow-up experiments such as Alanine scanning can be performed to identify the specific residues involved in contact. With this knowledge the interactions can be further validated in vivo through BiFCs. Once the interaction is verified, cellular models for functional analysis can be established to gain deeper insight into how these interactions underpin HPV pathogenesis.

Continued study of HPV E6 with quantitative interactomics

To further understand how complexomic shifts reflect disease states, further mapping of 16E6 affinity interactome using primary patient samples coupled to complexomic modelling would shed light onto temporal changes in interaction networks during transformation. Furthermore, there are several HPV types, from high and low-risk groups, which induce a broad range of clinical presentations. Mapping the affinity interactome of different E6 proteins across HPV groups would be exceptionally insightful into understanding principles of pathogenesis and the genotype-phenotype link. As previously mentioned, the various types of HPV E6 proteins display distinct binding profiles and comparative analysis across types and risk groups could help illuminate common infection mechanisms which likely resemble critical processes for sustaining the viral lifecycle and differentiate them from pathogenic pathways. There have already been efforts to map interactomes of different HPV types[132-134] while other studies focus on the variable outcomes of

shared HPV E6 interactions across the viral types[135-138]. Conversely this same type of comparative analysis could be used to gain a more nuanced understanding of within a particular HPV type through the analysis of polymorphisms. Polymorphisms in 16E6 represent interesting variability in phenotypic presentation and increased susceptibility to cervical cancer[139,140]. As these polymorphisms are often confined within the borders of a particular geographical location and thus linked to genetic factors their comparative analysis represents an opportunity to gain insight into host-factors that are deterministic in phenotypic presentation[141,142] whereas comparison across viral types will allow us to understand viral factors in pathogenesis. In either case the nHU resembles a highly adaptable and easy to implement technique that produces global affinity interactome maps that can be compared with a high degree of confidence.

Study of HPV oncoproteins with quantitative interactomics

Moving in a lateral direction, mapping of the 16E7 affinity interactome for combinatorial analysis could begin to reveal the true extent of interconnection between the networks of these two proteins during transformation. As mentioned, it will be critical to create a map of the intra-HPV oncoprotein interaction networks to understand how this influences HPV-host interactions. Mapping of the 16E5 affinity interactome would also prove interesting, although it is suspected to have a much less promiscuous interaction profile than that of 16E6 or 16E7 given its putative role as a viroporin.

Evaluating therapeutics with nHU

The successful adaptation of using nHU to test small molecule inhibitors represents a broader application across evaluating therapeutics. Therapeutic peptides in particular have emerged as a popular development and peptides targeting the 16E6 LxxLL binding pocket are in development[143,144]. Given the very flexible format of HU approaches and the cellular pertinence of nHU they represent very powerful and versatile tools in measuring multiple indicators of therapeutic potential at high throughput. I am delighted that this project will continue in the hands of Theo Juncker and that a model system will be established to optimize the various applications of this technique.

Towards development of a single affinity step quantitative interactomic method

Some of the most interesting biological interactions involve the most difficult proteins to study due to their challenging biochemical nature. While many developments in protein expression have been made making it possible to produce and purify these proteins in the lab, it is often only possible in small quantities. An advantage of the nHU is large protein concentrations aren’t necessary and extensive data can be obtained from a singular experiment. That said the process of recombinantly producing and purifying a challenging protein is often undesirable when pull-down formats are widely available. To this end, merging the traditional

format of a pull-down with the quantitative nature of the nHU would create a powerful tool in interactomics. The variability in cellular concentrations represents the largest barrier in this development as we need fixed concentrations of both baits and preys to accurately determine affinities. A steppingstone in this direction would be a single affinity purification step followed directly by nHU with cellular extracts. One possibility for developing this format would be to use strep-tagged baits and Streptactin resin to perform an affinity purification and resin saturation step simultaneously. The Streptactin resin is highly specific and has capacity and dimensions similar to that of Streptavidin resin so adequate bait saturation is possible with little risk of non-specific interactions.

References

1 Kiemer, L. & Cesareni, G. Comparative interactomics: comparing apples and pears? *Trends Biotechnol* **25**, 448-454, doi:10.1016/j.tibtech.2007.08.002 (2007).

2 Berggård, T., Linse, S. & James, P. Methods for the detection and analysis of protein–protein interactions. *Proteomics* **7**, 2833-2842 (2007).

3 Cozzone, A. J. Proteins: Fundamental chemical properties. *Encyclopedia of Life Sciences. No month listed-2002. John Wiley & Sons Ltd*, 1-10 (2002).

4 Cusick, M. E., Klitgord, N., Vidal, M. & Hill, D. E. Interactome: gateway into systems biology. *Human Molecular Genetics* **14**, R171-R181, doi:10.1093/hmg/ddi335 (2005).

5 Bludau, I. & Aebersold, R. Proteomic and interactomic insights into the molecular basis of cell functional diversity. *Nature Reviews Molecular Cell Biology* **21**, 327-340, doi:10.1038/s41580-020-0231-2 (2020).

6 Luck, K., Sheynkman, G. M., Zhang, I. & Vidal, M. Proteome-scale human interactomics. *Trends in biochemical sciences* **42**, 342-354 (2017).

7 Marsh, J. A. & Teichmann, S. A. Structure, dynamics, assembly, and evolution of protein complexes. *Annual review of biochemistry* **84**, 551-575 (2015).

8 Levy, E. D. & Pereira-Leal, J. B. Evolution and dynamics of protein interactions and networks. *Current opinion in structural biology* **18**, 349-357 (2008).

9 Zhan, X., Yan, C., Zhang, X., Lei, J. & Shi, Y. Structures of the human pre-catalytic spliceosome and its precursor spliceosome. *Cell Research* **28**, 1129-1140, doi:10.1038/s41422-018-0094-7 (2018).

10 Hein, M. Y. *et al.* A human interactome in three quantitative dimensions organized by stoichiometries and abundances. *Cell* **163**, 712-723 (2015).

11 Huttlin, E. L. *et al.* The BioPlex network: a systematic exploration of the human interactome. *Cell* **162**, 425-440 (2015).

12 Rolland, T. *et al.* A proteome-scale map of the human interactome network. *Cell* **159**, 1212-1226 (2014).

13 Wan, C. *et al.* Panorama of ancient metazoan macromolecular complexes. *Nature* **525**, 339-344 (2015).

14 Luck, K. *et al.* A reference map of the human binary protein interactome. *Nature* **580**, 402-408 (2020).

15 Yu, H. *et al.* High-quality binary protein interaction map of the yeast interactome network. *Science* **322**, 104-110 (2008).

16 Li, S. *et al.* A map of the interactome network of the metazoan C. elegans. *Science* **303**, 540-543 (2004).

17 Hink, M. A., Bisseling, T. & Visser, A. J. Imaging protein-protein interactions in living cells. *Plant molecular biology* **50**, 871-883 (2002).

18 Kerppola, T. K. Bimolecular fluorescence complementation (BiFC) analysis as a probe of protein interactions in living cells. *Annu. Rev. Biophys.* **37**, 465-487 (2008).

19 Brückner, A., Polge, C., Lentze, N., Auerbach, D. & Schlattner, U. Yeast Two-Hybrid, a Powerful Tool for Systems Biology. *International Journal of Molecular Sciences* **10**, 2763-2788 (2009).

20 Keegan, L., Gill, G. & Ptashne, M. Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. *Science* **231**, 699-704 (1986).

21 Fields, S. & Song, O.-k. A novel genetic system to detect protein–protein interactions. *Nature* **340**, 245-246 (1989).
22 Serebriiskii, I., Peri, S. & Golemis, E. Two-Hybrid Protein–Protein Interactions. (2013).
23 Isono, E. & Schwechheimer, C. Co-immunoprecipitation and protein blots. *Plant Developmental Biology: Methods and Protocols*, 377-387 (2010).
24 Louche, A., Salcedo, S. P. & Bigot, S. Protein–protein interactions: pull-down assays. *Bacterial Protein secretion systems: methods and protocols*, 247-255 (2017).
25 Roux, K. J., Kim, D. I., Burke, B. & May, D. G. BioID: a screen for protein-protein interactions. *Current protocols in protein science* **91**, 19.23. 11-19.23. 15 (2018).
26 Nguyen, T. M. T., Kim, J., Doan, T. T., Lee, M.-W. & Lee, M. APEX proximity labeling as a versatile tool for biological research. *Biochemistry* **59**, 260-269 (2019).
27 Qin, W., Cho, K. F., Cavanagh, P. E. & Ting, A. Y. Deciphering molecular interactions by proximity labeling. *Nature methods* **18**, 133-143 (2021).
28 Zhu, H. & Snyder, M. Protein chip technology. *Current Opinion in Chemical Biology* **7**, 55-63, doi:https://doi.org/10.1016/S1367-5931(02)00005-4 (2003).
29 Templin, M. F. *et al.* Protein microarray technology. *TRENDS in Biotechnology* **20**, 160-166 (2002).
30 Haab, B. Advances in protein microarray technology for protein expression and interaction profiling. *Current opinion in drug discovery & development* **4**, 116-123 (2001).
31 Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature* **422**, 198-207 (2003).
32 Liu, X., Abad, L., Chatterjee, L., Cristea, I. M. & Varjosalo, M. Mapping protein–protein interactions by mass spectrometry. *Mass Spectrometry Reviews* (2024).
33 Yu, C. & Huang, L. Cross-linking mass spectrometry (XL-MS): An emerging technology for interactomics and structural biology. *Analytical chemistry* **90**, 144 (2018).
34 O'Reilly, F. J. & Rappsilber, J. Cross-linking mass spectrometry: methods and applications in structural, molecular and systems biology. *Nature structural & molecular biology* **25**, 1000-1008 (2018).
35 Lee, K. & O'Reilly, F. J. Cross-linking mass spectrometry for mapping protein complex topologies in situ. *Essays in Biochemistry* **67**, 215-228 (2023).
36 Herfurth, M., Müller, F., Søgaard-Andersen, L. & Glatter, T. A miniTurbo-based proximity labeling protocol to identify conditional protein interactomes in vivo in Myxococcus xanthus. *STAR Protocols* **4**, 102657, doi:https://doi.org/10.1016/j.xpro.2023.102657 (2023).
37 Merkley, E. D., Cort, J. R. & Adkins, J. N. Cross-linking and mass spectrometry methodologies to facilitate structural biology: finding a path through the maze. *J Struct Funct Genomics* **14**, 77-90, doi:10.1007/s10969-013-9160-z (2013).
38 Goll, J. & Uetz, P. The elusive yeast interactome. *Genome biology* **7**, 1-6 (2006).
39 Bell, A. W. *et al.* A HUPO test sample study reveals common problems in mass spectrometry–based proteomics. *Nature methods* **6**, 423-430 (2009).
40 Turck, C. W. *et al.* The association of biomolecular resource facilities proteomics research group 2006 study: relative protein quantitation. *Molecular & Cellular Proteomics* **6**, 1291-1298 (2007).

41 Varjosalo, M. *et al.* Interlaboratory reproducibility of large-scale human protein-complex analysis by standardized AP-MS. *Nature methods* **10**, 307-314 (2013).
42 Listgarten, J. & Emili, A. Statistical and Computational Methods for Comparative Proteomic Profiling Using Liquid Chromatography-Tandem Mass Spectrometry. *Molecular & Cellular Proteomics* **4**, 419-434, doi:10.1074/mcp.R500005-MCP200 (2005).
43 Ramani, A. K., Bunescu, R. C., Mooney, R. J. & Marcotte, E. M. Consolidating the set of known human protein-protein interactions in preparation for large-scale mapping of the human interactome. *Genome Biology* **6**, R40, doi:10.1186/gb-2005-6-5-r40 (2005).
44 Hart, G. T., Ramani, A. K. & Marcotte, E. M. How complete are current yeast and human protein-interaction networks? *Genome biology* **7**, 1-9 (2006).
45 Taipale, M. *et al.* Quantitative analysis of HSP90-client interactions reveals principles of substrate recognition. *Cell* **150**, 987-1001 (2012).
46 Wippel, H. H., Chavez, J. D., Tang, X. & Bruce, J. E. Quantitative interactome analysis with chemical cross-linking and mass spectrometry. *Current opinion in chemical biology* **66**, 102076 (2022).
47 Chavez, J. D. *et al.* Quantitative interactome analysis reveals a chemoresistant edgotype. *Nature communications* **6**, 7928 (2015).
48 Charbonnier, S., Zanier, K., Masson, M. & Travé, G. Capturing protein–protein complexes at equilibrium: the holdup comparative chromatographic retention assay. *Protein expression and purification* **50**, 89-101 (2006).
49 Vincentelli, R. *et al.* Quantifying domain-ligand affinities and specificities by high-throughput holdup assay. *Nature Methods* **12**, 787-793, doi:10.1038/nmeth.3438 (2015).
50 Gogl, G. *et al.* Quantitative fragmentomics allow affinity mapping of interactomes. *Nature Communications* **13**, 5472, doi:10.1038/s41467-022-33018-0 (2022).
51 Zambo, B., Morlet, B., Negroni, L., Trave, G. & Gogl, G. Native holdup (nHU) to measure binding affinities from cell extracts. *bioRxiv*, 2022.2006.2022.497144, doi:10.1101/2022.06.22.497144 (2022).
52 Delalande, F. *et al.* The Holdup Multiplex, an assay for high-throughput measurement of protein-ligand affinity constants using a mass-spectrometry readout. *BioRxiv*, 2022.2012. 2008.519103 (2022).
53 Charbonnier, S., Gallego, O. & Gavin, A.-C. The social network of a cell: recent advances in interactome mapping. *Biotechnology annual review* **14**, 1-28 (2008).
54 Vincentelli, R. *et al.* High-throughput protein expression screening and purification in Escherichia coli. *Methods* **55**, 65-72 (2011).
55 Beadle, G. W. & Tatum, E. L. Genetic Control of Biochemical Reactions in Neurospora. *Proc Natl Acad Sci U S A* **27**, 499-506, doi:10.1073/pnas.27.11.499 (1941).
56 International Human Genome Sequencing, C. Finishing the euchromatic sequence of the human genome. *Nature* **431**, 931-945, doi:10.1038/nature03001 (2004).
57 Carter, H., Hofree, M. & Ideker, T. Genotype to phenotype via network analysis. *Current opinion in genetics & development* **23**, 611-621 (2013).

58 Parikshak, N. N., Gandal, M. J. & Geschwind, D. H. Systems biology and gene networks in neurodevelopmental and neurodegenerative disorders. *Nature Reviews Genetics* **16**, 441-458 (2015).
59 Cui, Y. *et al.* CRISP-view: a database of functional genetic screens spanning multiple phenotypes. *Nucleic acids research* **49**, D848-D854 (2021).
60 Hartwell, L. H., Hopfield, J. J., Leibler, S. & Murray, A. W. From molecular to modular cell biology. *Nature* **402**, C47-C52 (1999).
61 Rual, J.-F. *et al.* Towards a proteome-scale map of the human protein–protein interaction network. *Nature* **437**, 1173-1178, doi:10.1038/nature04209 (2005).
62 Goh, K.-I. *et al.* The human disease network. *Proceedings of the National Academy of Sciences* **104**, 8685-8690, doi:doi:10.1073/pnas.0701361104 (2007).
63 Vidal, M., Cusick, Michael E. & Barabási, A.-L. Interactome Networks and Human Disease. *Cell* **144**, 986-998, doi:10.1016/j.cell.2011.02.016 (2011).
64 Vidal, M. Interactome modeling. *FEBS Letters* **579**, 1834-1838, doi:https://doi.org/10.1016/j.febslet.2005.02.030 (2005).
65 Via, A., Uyar, B., Brun, C. & Zanzoni, A. How pathogens use linear motifs to perturb host cell networks. *Trends in biochemical sciences* **40**, 36-48 (2015).
66 Jean Beltran, P. M., Federspiel, J. D., Sheng, X. & Cristea, I. M. Proteomics and integrative omic approaches for understanding host–pathogen interactions and infectious diseases. *Molecular systems biology* **13**, 922 (2017).
67 Dyer, M. D., Murali, T. & Sobral, B. W. The landscape of human proteins interacting with viruses and other pathogens. *PLoS Pathog* **4**, e32 (2008).
68 Coyaud, E. *et al.* Global interactomics uncovers extensive organellar targeting by Zika virus. *Molecular & Cellular Proteomics* **17**, 2242-2255 (2018).
69 Caillet-Saguy, C. *et al.* Host PDZ-containing proteins targeted by SARS-CoV-2. *The FEBS Journal* **288**, 5148-5162 (2021).
70 Khadka, S. *et al.* A physical interaction network of dengue virus and human proteins. *Molecular & Cellular Proteomics* **10** (2011).
71 Schweppe, D. K. *et al.* Host-microbe protein interactions during bacterial infection. *Chemistry & biology* **22**, 1521-1530 (2015).
72 Chen, B. *et al.* Large-Scale Identification of Bacteria–Host Crosstalk by Affinity Chromatography: Capturing the Interactions of Streptococcus suis Proteins with Host Cells. *Journal of Proteome Research* **10**, 5163-5174 (2011).
73 Li, X.-M., Huang, S. & Li, X. D. Photo-ANA enables profiling of host–bacteria protein interactions during infection. *Nature Chemical Biology* **19**, 614-623 (2023).
74 Culibrk, L., Croft, C. A. & Tebbutt, S. J. Systems biology approaches for host–fungal interactions: An expanding multi-omics frontier. *OMICS: A Journal of Integrative Biology* **20**, 127-138 (2016).
75 Ball, B., Bermas, A., Carruthers-Lay, D. & Geddes-McAlister, J. Mass spectrometry-based proteomics of fungal pathogenesis, host–fungal interactions, and antifungal development. *Journal of Fungi* **5**, 52 (2019).
76 James, K. & Olson, P. D. The tapeworm interactome: inferring confidence scored protein-protein interactions from the proteome of Hymenolepis microstoma. *BMC genomics* **21**, 1-22 (2020).

77 LaCount, D. J. *et al.* A protein interaction network of the malaria parasite Plasmodium falciparum. *Nature* **438**, 103-107 (2005).
78 Gazestani, V. H. *et al.* A Protein Complex Map of Trypanosoma brucei. *PLOS Neglected Tropical Diseases* **10**, e0004533, doi:10.1371/journal.pntd.0004533 (2016).
79 Mammas, I. N., Spandidos, D. A. & Sourvinos, G. Genomic diversity of human papillomaviruses (HPV) and clinical implications: an overview in adulthood and childhood. *Infection, Genetics and Evolution* **21**, 220-226 (2014).
80 Neveu, G. *et al.* Comparative analysis of virus–host interactomes with a mammalian high-throughput protein complementation assay based on Gaussia princeps luciferase. *Methods* **58**, 349-359 (2012).
81 Graham, Sheila V. The human papillomavirus replication cycle, and its links to cancer progression: a comprehensive review. *Clinical Science* **131**, 2201-2221, doi:10.1042/CS20160786 (2017).
82 *PaVE The papillomavirus episteme. Available from: pave.niaid.nih.gov*.
83 Faridi, R., Zahra, A., Khan, K. & Idrees, M. Oncogenic potential of Human Papillomavirus (HPV) and its relation with cervical cancer. *Virology Journal* **8**, 1-8 (2011).
84 McLaughlin-Drubin, M. E. & Münger, K. Oncogenic activities of human papillomaviruses. *Virus research* **143**, 195-208 (2009).
85 Burd, E. M. Human papillomavirus and cervical cancer. *Clin Microbiol Rev* **16**, 1-17, doi:10.1128/cmr.16.1.1-17.2003 (2003).
86 Scarth, J. A., Patterson, M. R., Morgan, E. L. & Macdonald, A. The human papillomavirus oncoproteins: a review of the host pathways targeted on the road to transformation. *J Gen Virol* **102**, doi:10.1099/jgv.0.001540 (2021).
87 Martinez-Zapien, D. *et al.* Structure of the E6/E6AP/p53 complex required for HPV-mediated degradation of p53. *Nature* **529**, 541-545, doi:10.1038/nature16481 (2016).
88 Beckett, D., Kovaleva, E. & Schatz, P. J. A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Sci* **8**, 921-929, doi:10.1110/ps.8.4.921 (1999).
89 Sidi, A. O. *et al.* Strategies for bacterial expression of protein-peptide complexes: application to solubilization of papillomavirus E6. *Protein Expr Purif* **80**, 8-16, doi:10.1016/j.pep.2011.06.013 (2011).
90 Kay, B. K., Thai, S. & Volgina, V. V. in *High Throughput Protein Expression and Purification: Methods and Protocols* (ed Sharon A. Doyle) 185-198 (Humana Press, 2009).
91 Li, Y. & Sousa, R. Expression and purification of E. coli BirA biotin ligase for in vitro biotinylation. *Protein Expression and Purification* **82**, 162-167, doi:https://doi.org/10.1016/j.pep.2011.12.008 (2012).
92 Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nature Methods* **9**, 676-682, doi:10.1038/nmeth.2019 (2012).
93 Wilkins, M. R. *et al.* Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol* **112**, 531-552, doi:10.1385/1-59259-584-7:531 (1999).
94 Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical*

Biochemistry **72**, 248-254, doi:https://doi.org/10.1016/0003-2697(76)90527-3 (1976).
95 Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods* **13**, 731-740, doi:10.1038/nmeth.3901 (2016).
96 Edwards, R. J., Davey, N. E. & Shields, D. C. SLiMFinder: A Probabilistic Method for Identifying Over-Represented, Convergently Evolved, Short Linear Motifs in Proteins. *PLOS ONE* **2**, e967, doi:10.1371/journal.pone.0000967 (2007).
97 Davey, N. E., Haslam, N. J., Shields, D. C. & Edwards, R. J. SLiMFinder: a web server to find novel, significantly over-represented, short protein motifs. *Nucleic Acids Research* **38**, W534-W539, doi:10.1093/nar/gkq440 (2010).
98 Krystkowiak, I. & Davey, N. E. SLiMSearch: a framework for proteome-wide discovery and annotation of functional modules in intrinsically disordered regions. *Nucleic Acids Research* **45**, W464-W469, doi:10.1093/nar/gkx238 (2017).
99 Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589, doi:10.1038/s41586-021-03819-2 (2021).
100 Varadi, M. *et al.* AlphaFold Protein Structure Database in 2024: providing structure coverage for over 214 million protein sequences. *Nucleic Acids Res* **52**, D368-d375, doi:10.1093/nar/gkad1011 (2024).
101 Schmid, E. W. & Walter, J. C. Predictomes: A classifier-curated database of AlphaFold-modeled protein-protein interactions. *bioRxiv*, doi:10.1101/2024.04.09.588596 (2024).
102 Milo, R., Jorgensen, P., Moran, U., Weber, G. & Springer, M. BioNumbers—the database of key numbers in molecular and cell biology. *Nucleic acids research* **38**, D750-D753 (2010).
103 Crish, J. F. *et al.* Suprabasal expression of the human papillomavirus type 16 oncoproteins in mouse epidermis alters expression of cell cycle regulatory proteins. *Carcinogenesis* **21**, 1031-1037 (2000).
104 Dürst, M., Glitz, D., Schneider, A. & zur Hausen, H. Human papillomavirus type 16 (HPV 16) gene expression and DNA replication in cervical neoplasia: analysis by in situ hybridization. *Virology* **189**, 132-140 (1992).
105 Kathleena, W. Too many cooks in the kitchen: HPV driven carcinogenesis–the result of collaboration or competition? *Tumour Virus Research*, 200311 (2024).
106 Tytgat, H. L. *et al.* Endogenous biotin-binding proteins: an overlooked factor causing false positives in streptavidin-based protein detection. *Microb Biotechnol* **8**, 164-168, doi:10.1111/1751-7915.12150 (2015).
107 Zámbó, B. *et al.* Cellular expression and function of naturally occurring variants of the human ABCG2 multidrug transporter. *Cell Mol Life Sci* **77**, 365-378, doi:10.1007/s00018-019-03186-2 (2020).
108 Lin, Z. *et al.* Decoding WW domain tandem-mediated target recognitions in tissue growth and cell polarity. *eLife* **8**, e49439, doi:10.7554/eLife.49439 (2019).
109 Glaunsinger, B. A., Lee, S. S., Thomas, M., Banks, L. & Javier, R. Interactions of the PDZ-protein MAGI-1 with adenovirus E4-ORF1 and high-risk papillomavirus E6 oncoproteins. *Oncogene* **19**, 5270-5280 (2000).

110 White, E. A., Münger, K. & Howley, P. M. High-risk human papillomavirus E7 proteins target PTPN14 for degradation. *MBio* **7**, 10.1128/mbio. 01530-01516 (2016).
111 Hawley-Nelson, P., Vousden, K. H., Hubbert, N. L., Lowy, D. R. & Schiller, J. T. HPV16 E6 and E7 proteins cooperate to immortalize human foreskin keratinocytes. *EMBO J* **8**, 3905-3910 (1989).
112 Lim, J. *et al.* Evidence for direct interaction between the oncogenic proteins E6 and E7 of high-risk human papillomavirus (HPV). *Journal of Biological Chemistry* **299** (2023).
113 Beutner, K. R. & Ferenczy, A. Therapeutic approaches to genital warts. *The American journal of medicine* **102**, 28-37 (1997).
114 Rietz, A. *et al.* Molecular Probing of the HPV-16 E6 Protein Alpha Helix Binding Groove with Small Molecule Inhibitors. *PLOS ONE* **11**, e0149845, doi:10.1371/journal.pone.0149845 (2016).
115 Baleja, J. D. *et al.* Identification of inhibitors to papillomavirus type 16 E6 protein based on three-dimensional structures of interacting proteins. *Antiviral research* **72**, 49-59 (2006).
116 Klein, C. & Vassilev, L. Targeting the p53–MDM2 interaction to treat cancer. *British journal of cancer* **91**, 1415-1419 (2004).
117 Shen, H. & G Maki, C. Pharmacologic activation of p53 by small-molecule MDM2 antagonists. *Current pharmaceutical design* **17**, 560-568 (2011).
118 Vaynberg, J. & Qin, J. Weak protein–protein interactions as probed by NMR spectroscopy. *TRENDS in Biotechnology* **24**, 22-27 (2006).
119 Gong, Z. *et al.* Preferential Regulation of Transient Protein–Protein Interaction by the Macromolecular Crowders. *The Journal of Physical Chemistry B* **126**, 4840-4848 (2022).
120 Warren, C. J., Westrich, J. A., Doorslaer, K. V. & Pyeon, D. Roles of APOBEC3A and APOBEC3B in human papillomavirus infection and disease progression. *Viruses* **9**, 233 (2017).
121 Vieira, V. C. *et al.* Human papillomavirus E6 triggers upregulation of the antiviral and cancer genomic DNA deaminase APOBEC3B. *MBio* **5**, 10.1128/mbio. 02234-02214 (2014).
122 Revathidevi, S., Murugan, A. K., Nakaoka, H., Inoue, I. & Munirajan, A. K. APOBEC: A molecular driver in cervical cancer pathogenesis. *Cancer letters* **496**, 104-116 (2021).
123 Mitsuishi, T. *et al.* Combined analysis of cell growth and apoptosis-regulating proteins in HPVs associated anogenital tumors. *BMC cancer* **10**, 1-10 (2010).
124 Antill, Y. C. *et al.* Lynch syndrome and cervical cancer. *Int J Cancer* **137**, 2757-2761, doi:10.1002/ijc.29641 (2015).
125 Leiendecker, L. *et al.* Human papillomavirus 42 drives digital papillary adenocarcinoma and elicits a germ cell–like program conserved in HPV-positive cancers. *Cancer Discovery* **13**, 70-84 (2023).
126 Rapp, L. & Chen, J. J. The papillomavirus E6 proteins. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer* **1378**, F1-F19 (1998).
127 Olmedo-Nieva, L., Muñoz-Bello, J. O., Manzo-Merino, J. & Lizano, M. New insights in Hippo signalling alteration in human papillomavirus-related cancers. *Cellular Signalling* **76**, 109815, doi:https://doi.org/10.1016/j.cellsig.2020.109815 (2020).

128 Tao, Y., Xu, X., Shen, R., Miao, X. & He, S. Roles of ubiquitin-specific protease 13 in normal physiology and tumors. *Oncology Letters* **27**, 58 (2023).
129 Morgan, E. L. *et al.* The deubiquitinase (DUB) USP13 promotes Mcl-1 stabilisation in cervical cancer. *Oncogene* **40**, 2112-2129, doi:10.1038/s41388-021-01679-8 (2021).
130 Lebeau, A. *et al.* HPV infection alters vaginal microbiome through down-regulating host mucosal innate peptides used by Lactobacilli as amino acid sources. *Nature Communications* **13**, 1076 (2022).
131 Dominguez-Lafage, C. *et al.* The cutaneous virome: from virology to personalized medicine. *Virologie* **27**, 333-354 (2023).
132 Eckhardt, M. *et al.* Multiple Routes to Oncogenesis Are Promoted by the Human Papillomavirus-Host Protein Network. *Cancer Discov* **8**, 1474-1489, doi:10.1158/2159-8290.Cd-17-1018 (2018).
133 Rozenblatt-Rosen, O. *et al.* Interpreting cancer genomes using systematic host network perturbations by tumour virus proteins. *Nature* **487**, 491-495 (2012).
134 White, E. A. *et al.* Comprehensive Analysis of Host Cellular Interactions with Human Papillomavirus E6 Proteins Identifies New E6 Binding Partners and Reflects Viral Diversity. *J Virol* **86**, 13174-13186, doi:doi:10.1128/jvi.02172-12 (2012).
135 Howie, H. L. *et al.* Beta-HPV 5 and 8 E6 promote p300 degradation by blocking AKT/p300 association. *PLoS Pathog* **7**, e1002211 (2011).
136 Muench, P. *et al.* Cutaneous papillomavirus E6 proteins must interact with p300 and block p53-mediated apoptosis for cellular immortalization and tumorigenesis. *Cancer research* **70**, 6913-6924 (2010).
137 Mesplède, T. *et al.* p53 degradation activity, expression, and subcellular localization of E6 proteins from 29 human papillomavirus genotypes. *J Virol* **86**, 94-107 (2012).
138 Fu, L. *et al.* Degradation of p53 by human Alphapapillomavirus E6 proteins shows a stronger correlation with phylogeny than oncogenicity. *PloS one* **5**, e12816 (2010).
139 Zehbe, I., Voglino, G., Delius, H., Wilander, E. & Tommasino, M. Risk of cervical cancer and geographical variations of human papillomavirus 16 E6 polymorphisms. *The Lancet* **352**, 1441-1442 (1998).
140 Zhao, J. *et al.* Phylogeny and polymorphism in the E6 and E7 of human papillomavirus: alpha-9 (HPV16, 31, 33, 52, 58), alpha-5 (HPV51), alpha-6 (HPV53, 66), alpha-7 (HPV18, 39, 59, 68) and alpha-10 (HPV6, 44) in women from Shanghai. *Infectious agents and cancer* **14**, 1-11 (2019).
141 Ermini, L., Wilson, I., Goodship, T. & Sheerin, N. Complement polymorphisms: geographical distribution and relevance to disease. *Immunobiology* **217**, 265-271 (2012).
142 Zehbe, I. *et al.* Human papillomavirus 16 E6 polymorphisms in cervical lesions from different European populations and their correlation with human leukocyte antigen class II haplotypes. *International journal of cancer* **94**, 711-716 (2001).
143 Bruzzoni-Giovanelli, H. *et al.* Interfering peptides targeting protein–protein interactions: the next generation of drugs? *Drug Discovery Today* **23**, 272-285 (2018).

144 Ye, X. *et al*. Discovery of reactive peptide inhibitors of human papillomavirus oncoprotein E6. *Chemical Science* **14**, 12484-12497 (2023).

Contributions

Kathleen Weimer (KW): Conceptualization of the project, experimental design, experimentation and data acquisition, data analysis, writing and revision of manuscripts, and funding acquisition.

Described below are specific contributions made by GG to this project—

-KW acquired funding through the IMCBio PhD program to begin a project in the lab of GT. KW performed eight-point nHU experiments of 16E6 using HeLa and HaCaT cell lines for analysis by nHU-WB. Aside from initial WBs performed by BZ, all WBs and their quantifications were performed by KW. Analysis of binding curves and data representation of titration experiments were performed by KW. All nHU spike-in experiments were performed and analyzed by KW. nHU experiments using the C-terminal PBM of 16E6 and E6* peptide as bait were performed by KW and mass spec data was processed by KW. Cellular extracts and cellular assays were performed and analyzed by KW. BLI experiments were performed by KW at the IGBMC Biophysics Platform. AlphaFold and Predictomes analysis was performed by KW. Writing and revision of manuscripts was carried out by KW.

Gilles Trave (GT): Conceptualization of the project, oversight of the project, revision of manuscripts, and funding acquisition.

Gergo Gogl (GG): Conceptualization of the project, oversight of the project, experimental design, experimentation, data analysis, review of data analysis

Described below are specific contributions made by GG to this project—

-GG performed the first single point and eight-point nHU experiments to train KW in technical aspects of the assay. The mass spec data produced from the single point experiments were used as the results for measuring the 16E6 affinity interactome with nHU. The data from these experiments was processed and figures representing the data were produced by GG with KW. GG assisted in the design of all nHU spike-in experiments. Library design of LxxLL peptides used for HU assays was performed by GG with KW following initial panels proposed by KW. Peptides were dissolved by KW, GG, and BZ. HU assays were performed by GG and KW, HU measurements were taken by GG, processing of data was performed by GG and KW, and GG produced figures representing the HU data. Data from BLI experiments was processed by GG and figures created by GG. GG reviewed all data processing and representations performed by KW.

Boglarka Zambo (BZ): Experimental design and experimentation

Described below are specific contributions made by GG to this project—

-BZ performed the first preparation of total cell extracts (HeLa and HaCaT) used in the first nHU experiments performed by GG to train KW in technical aspects of the process. BZ performed the first three Western Blots in the nHU-WB analysis to

train KW in technical aspects. BZ provided engineered HEK293T cell lines stably expressing 16E6 for KW to use in transfection studies. BZ assisted KW in the design of all cellular assays.

Jia Wen (JW): Provide UBE3A used in nHU spike-in experiments.

IGBMC Proteomics Platform: The IGBMC Proteomics Platform conducted all mass spec experiments and performed initial data processing in Perseus software.

List of publications

Kathleen Weimer, Boglarka Zambo, & Gergo Gogl (2023). Molecules interact. But how strong and how much?. BioEssays.

Kathleen Weimer (2024) Too many cooks in the kitchen: HPV driven carcinogenesis – the result of collaboration or competition?. Tumour Virus Research. (*in review*)

Kathleen Weimer & Gilles Trave (2024) Modelling HPV 16E6-host protein complex formation with quantitative interactomics. (*manuscript in preparation*)

List of oral presentations

21 October 2024 Oral presentation at the Department of Pathology and Laboratory Medicine, Tumor Biology Program, Children's Hospital of Philadelphia, Pennsylvania, USA

16-21 July 2024 Oral presentation at DNA Tumour Virus Meeting (ICGEB) Trieste, Italy

23 November 2023 Poster presentation at SPB session IGBMC Strasbourg, France

27-29 September 2023 Oral presentation at AFIPP Scientific Days Fort-Mahon-Plage, France

-Recipient of AFIPP Scientific Days Student Presenter Award

8-12 July 2023 Poster presentation at FEBS 47th Annual Conference Tours, France

20-21 April 2023 Flash-talk at University of Strasbourg Doctoral Student Days

Posters appear on the successive pages in the following order:

8-12 July 2023 Poster presentation at FEBS 47th Annual Conference Tours, France

23 November 2023 Poster presentation at SPB session IGBMC Strasbourg, France

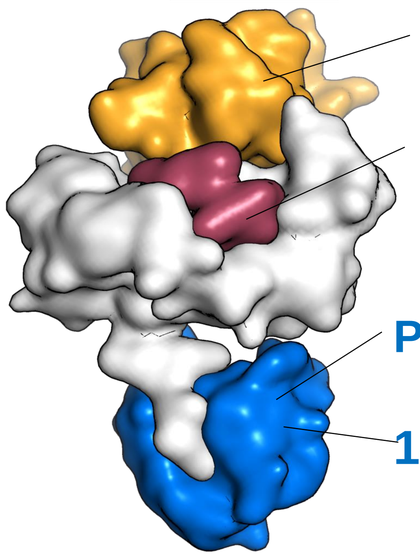
The HPV16 E6 affinity interactome

Weimer K., Zambo B., Trave G.*, Gogl G.*
IGBMC, INSERM U1258/CNRS UMR 7104/Universite de Strasbourg, Illkirch, France

Pathogenic hijacking by the promiscuous HPV16 oncoprotein E6

Up to **99%** of cervical cancer cases are caused by a group of "High-Risk" Human Papilloma Viruses (HR-HPVs). Of these **>50%** of diagnosed cervical cancers are reported as **HPV16(+)**. The **HPV16 oncoprotein E6** has been identified as a driver of HPV related pathologies interacting with diverse cellular targets.

HPV E6



E6 binds to diverse cellular targets:

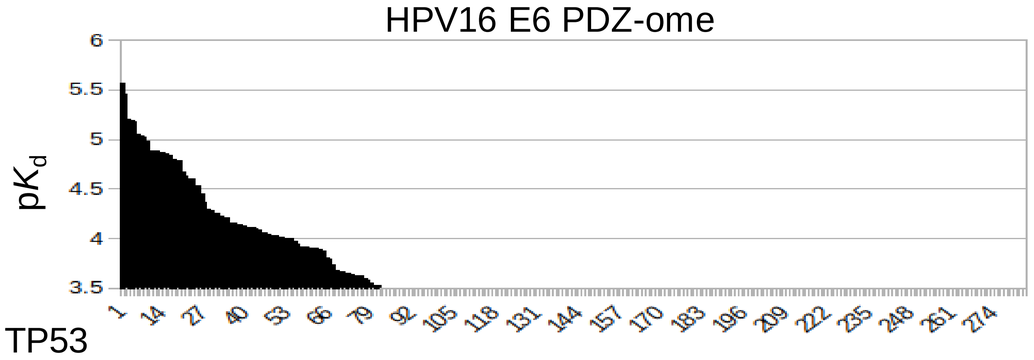
Core domain of **TP53** *(Martinez-Zapien et al. Nature 2016)*
Recruited by UBE3A-bound E6

LxxLL motif containing proteins such as UBE3A or PXN (*Zanier et al., Science 2013*)
Ubiquitin ligase UBE3A promotes proteasomal degradation of TP53

PDZ domain containing proteins such as SNX27 or SCRIB *(Vincentelli et al., Nat Meth. 2015)*

14-3-3 protein upon phosphorylation *(Gogl et al., Nat Comm. 2021)*
PDZ and 14-3-3 proteins mediate cellular localization upon binding

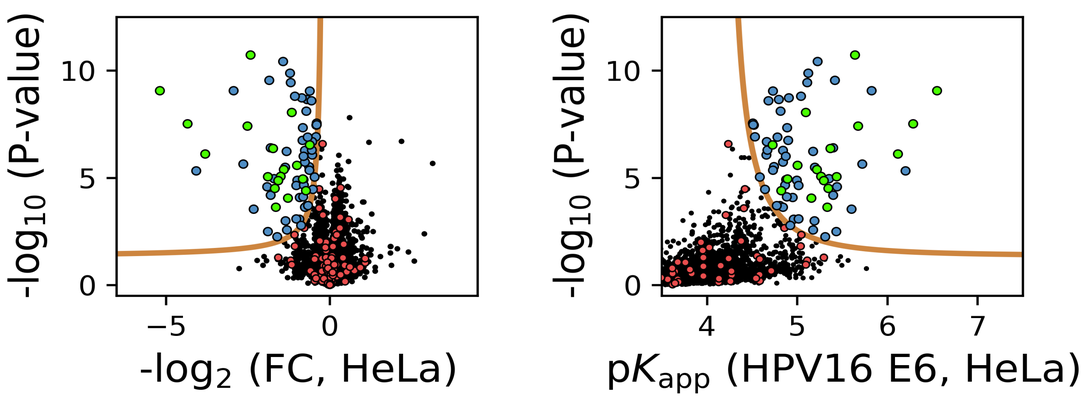
HPV16 E6 PDZ-ome

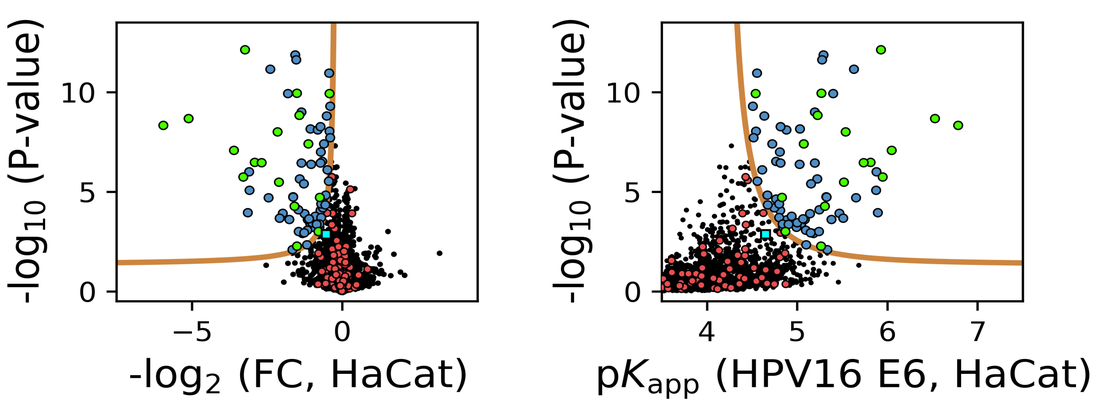


84 out of the **266** PDZ domains show detectable binding to the HPV16 E6 C-terminal tail (*Gogl et al., Nat Comm. 2022*). However, fragmentomic studies only provide a partial view. Proximal to phenotype, interactomes sit at a critical junction of the genotype-phenotype link making complete interactomes essential for understanding pathogenesis. ***Here, for the first time we measure the affinity interactome of full length HPV16 E6.***

Characterizing the full-length affinity interactome of HPV16 E6

Using native holdup (nHU) (*Zambo et al., Sci Adv. 2022*) we measured the affinity interactome of full length HPV16 E6 against full length proteins from cell extracts

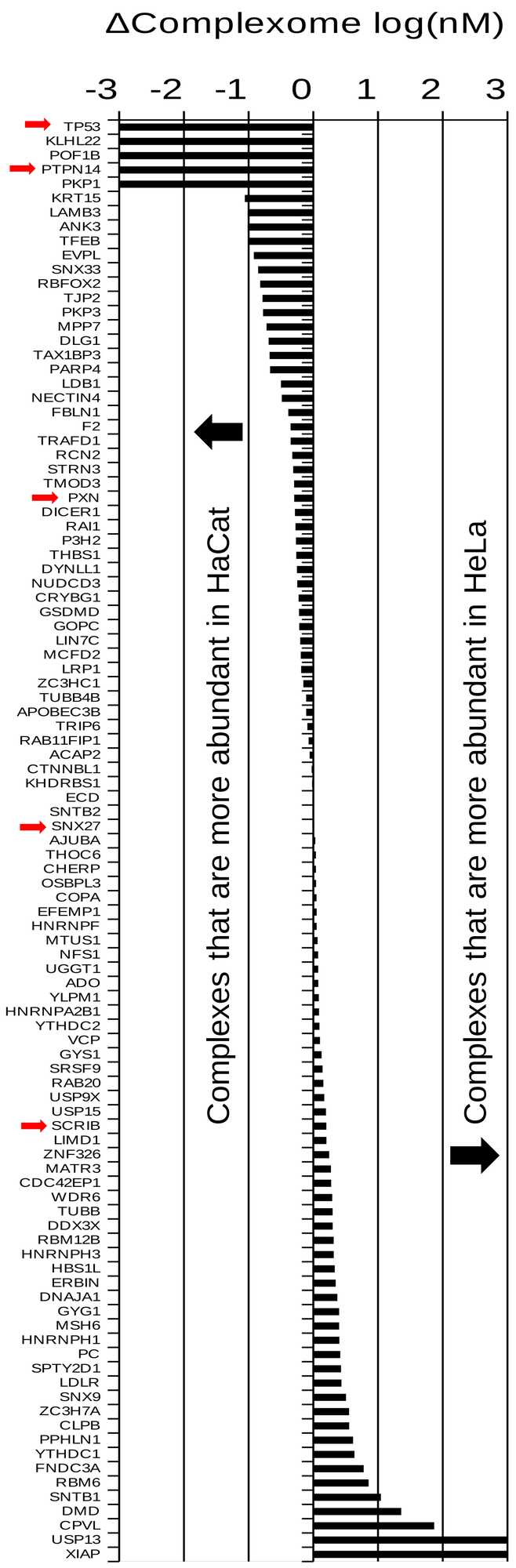




= Newly described partners
= Previously described partners
= Previously described partners below binding threshold
= TP53
= Significance Threshold

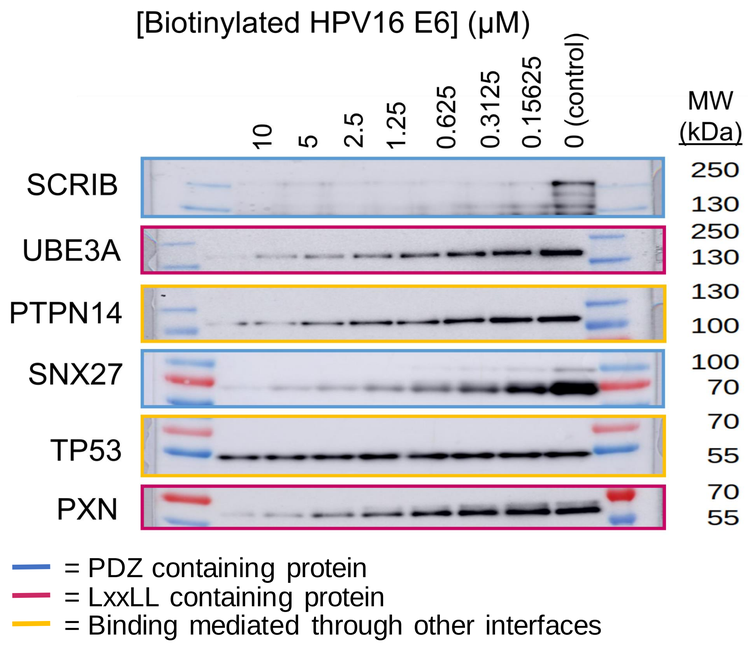
$pK_{app} = -\log_{10} K_{app}$
K_{app} = apparent dissociation constant

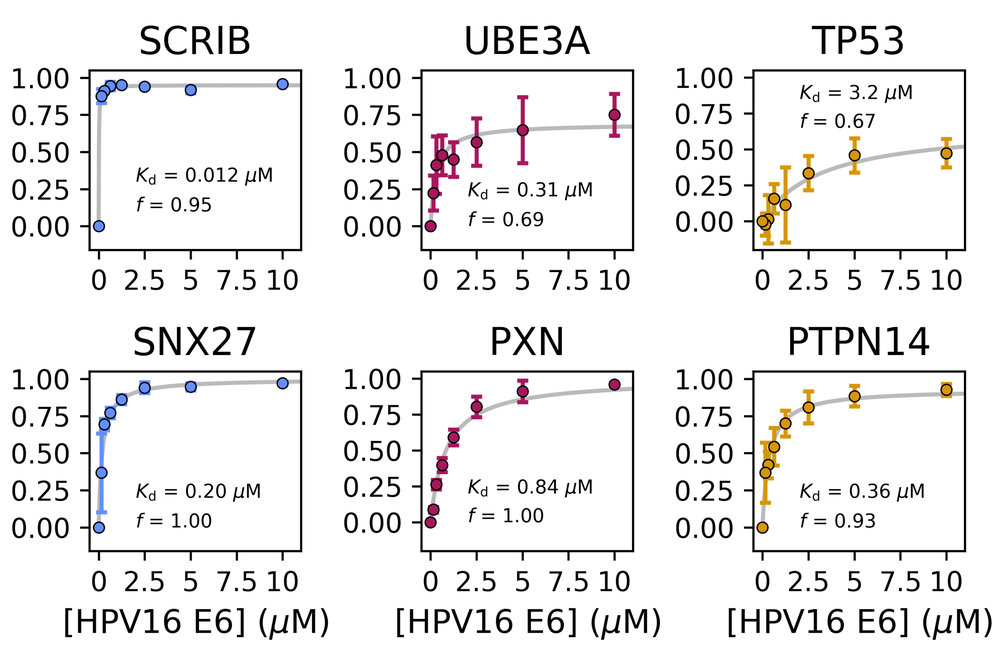
HPV16 E6 Complexome



Approximately **4500** affinities were measured and **100** partners identified from HPV18(+) HeLa and HPV(-) keratinocyte HaCat cell lines.

nHU titration experiment for selected partners, analyzed by Western Blots





All selected partners follow a simple binding mechanism regardless of the mediating interface (PDZ-PBM, LxxLL binding pocket, etc.)

Measured affinities and proteomes can be utilized to estimate the amounts of complexes within a cell, or the **complexome**

Using the **complexome** of HeLa and HaCat to identify critical targets of HPV16 E6

Complexomic changes indicate many previously uncharacterized targets of HPV16 E6.

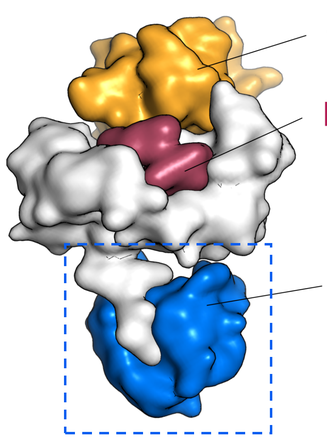
Investigating HPV16 E6 interactions with native holdup

Weimer K., Zambo B., Trave G.*, Gogl G.*
IGBMC, INSERM U1258/CNRS UMR 7104/Universite de Strasbourg, Illkirch, France

Pathogenic hijacking by the promiscuous HPV16 E6 oncoprotein

Up to **99%** of cervical cancer cases are caused by a group of "High-Risk" Human Papilloma Viruses (HR-HPVs). Of these **>50%** of diagnosed cervical cancers are reported as **HPV16(+)**. The **HPV16 E6 oncoprotein** has been identified as a driver of HPV related pathologies, interacting with diverse cellular targets.

HPV E6

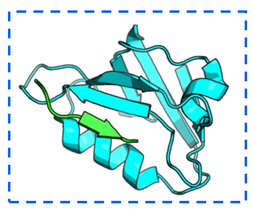


E6 binds to diverse cellular targets:

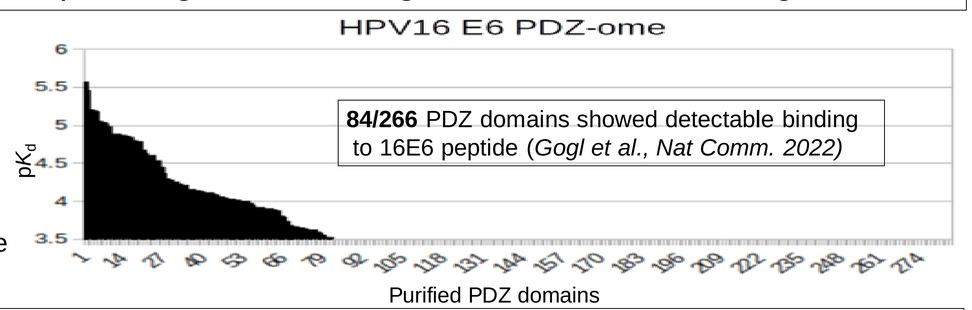
Core domain of **TP53** *(Martinez-Zapien et al. Nature 2016)*
Recruited by UBE3A-bound E6

LxxLL motif containing proteins such as UBE3A or PXN (*Zanier et al., Science 2013*)
Ubiquitin ligase UBE3A promotes proteasomal degradation of TP53

PDZ domain containing proteins such as SNX27 or SCRIB *(Vincentelli et al., Nat Meth. 2015)*



MAGI1_2 PDZ - **16E6** peptide (PDB: **6TWQ**)

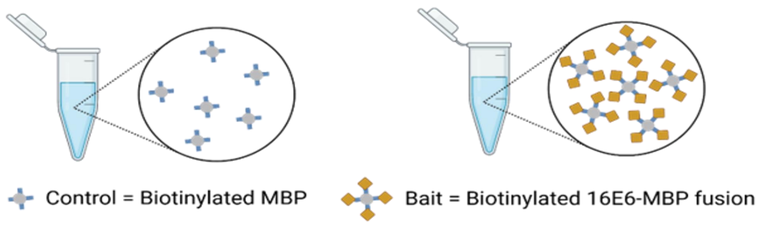


Measuring binding affinities allows us to relativize interactions in new ways as exemplified by the affinity ranking profile of the PDZ-ome above. However, this fragmentomic study only provides a partial view of the interactions. Proximal to phenotype, interactomes sit at a critical junction of the genotype-phenotype link making complete interactomes essential for understanding pathogenesis. ***Here for the first time we measure the full length HPV16 E6 affinity interactome.***

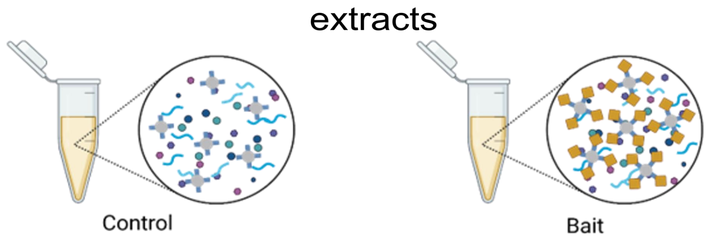
Characterizing the full-length affinity interactome of HPV16 E6

Native Holdup (nHU)

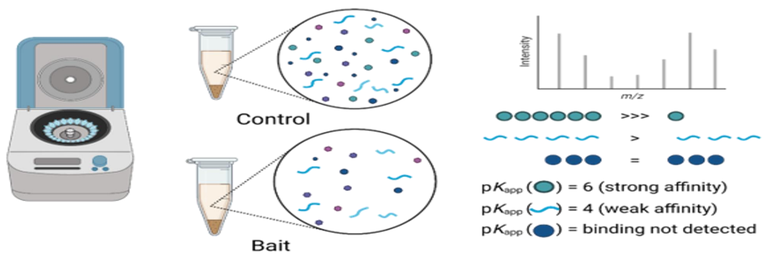
Immobilization of bait and control on streptavidin resin at fixed concentration



Incubation of bait and control resins with cellular extracts

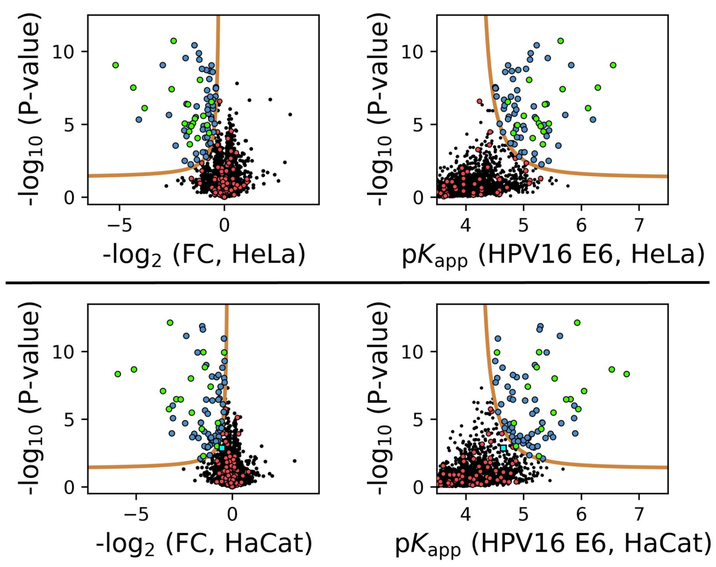


Separation of bound and unbound fractions followed by analysis of depleted preys

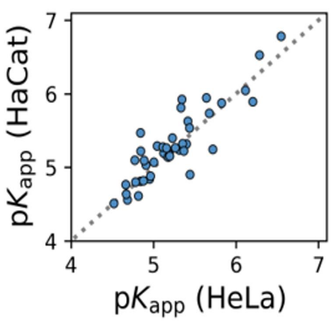


Using native holdup (nHU) (*Zambo et al., Sci Adv. 2022*) we measured the affinity interactome of full length HPV16 E6 against full length partners in a pseudo-native state at high throughput and gained further mechanistic insights.

Measuring the HPV16 E6 affinity interactome with nHU-MS

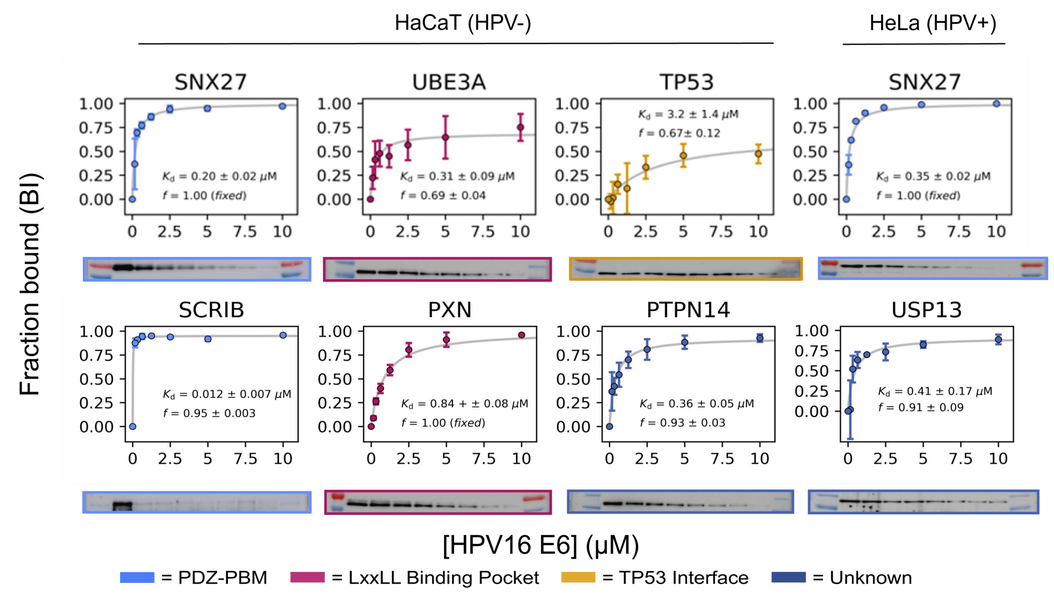


Approximately **4500** affinities were measured and **100** partners identified from HPV18(+) HeLa and HPV(-) keratinocyte HaCat cell lines.



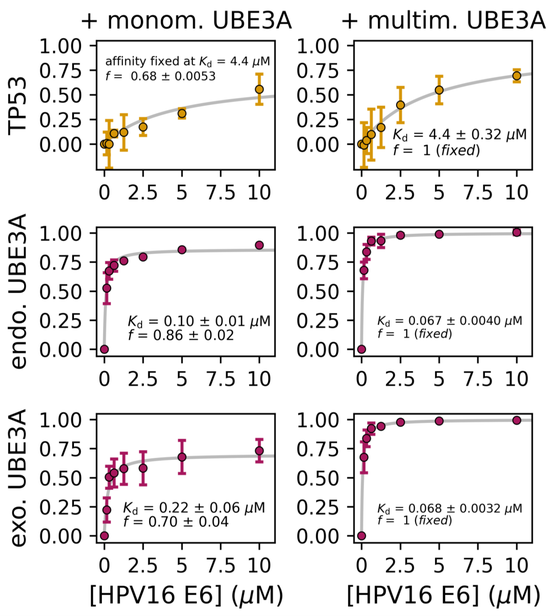
Agreement between affinities of both cell lines demonstrates they are an intrinsic property of the system.

Mechanistic insights with nHU-WB titration



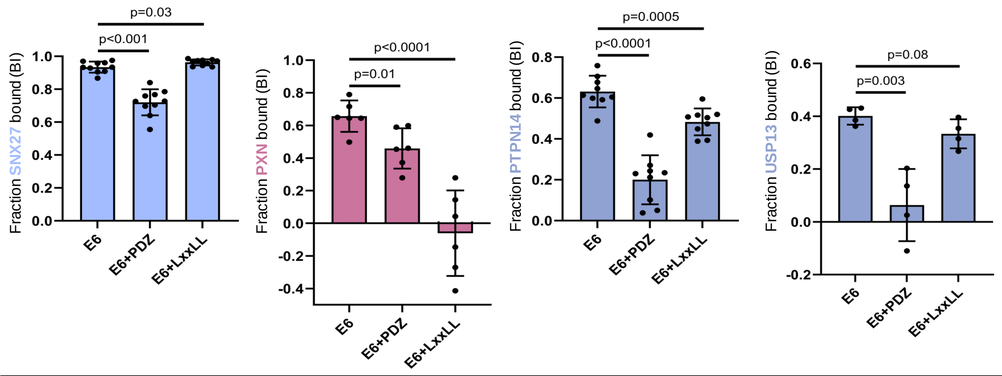
All selected partners follow a simple binding mechanism regardless of the mediating interface (PDZ-PBM, LxxLL binding pocket, etc.)

Investigating UBE3A binding activity with nHU spike in



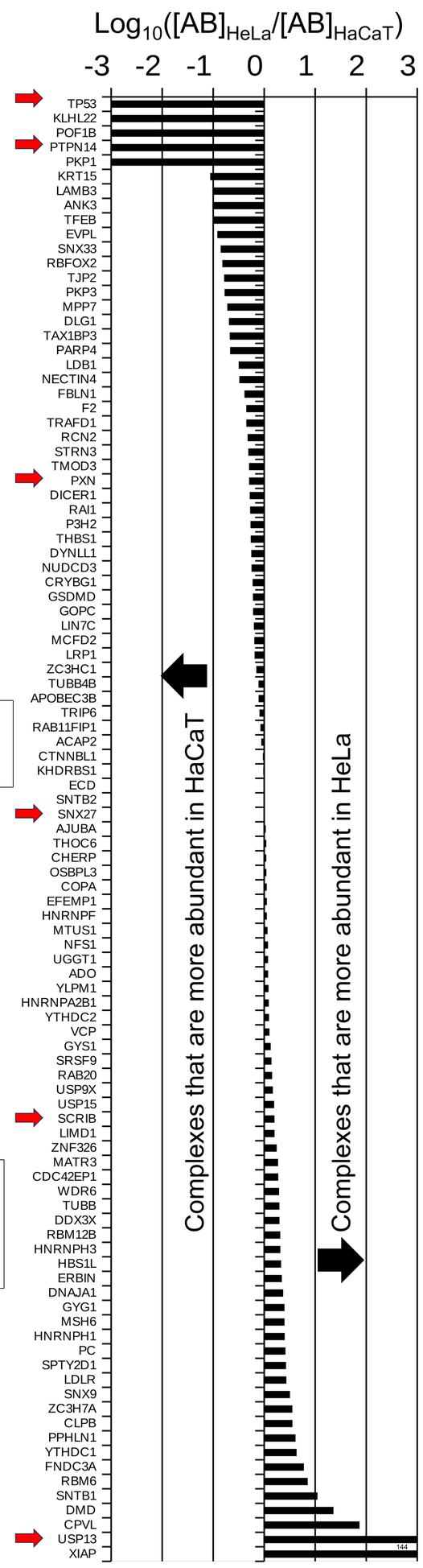
UBE3A binding activity is recovered when cellular extracts are supplemented with recombinantly produced multimeric UBE3A. It appears that most cellular UBE3A is monomeric.

Identifying binding interfaces with nHU spike in



Interaction between HPV16 E6 and novel partners PTPN14 and USP13 are disrupted when cellular extracts are supplemented with purified PDZ domain indicating interactions are mediated by the C-terminus of E6.

HPV16 E6 ΔComplexome



Measured affinities and proteomes can be utilized to estimate the **"complexome"**, i.e. the amounts of complexes within a cell.

Using the complexome of HeLa and HaCat to identify critical targets of HPV16 E6

Kathleen Weimer ED414 Equipe : TRAVE
Numéro d'étudiant : 22120266

Résumé de la thèse : Cartographie de l'interactome de la protéine hôte HPV16 E6 à l'aide de méthodes d'interactomique quantitative pour étudier les propriétés biophysiques de la liaison et modéliser la formation de complexes cellulaires.

Introduction :

Les papillomavirus humains (HPV) constituent l'infection sexuellement transmissible la plus répandue. Alors que les présentations cliniques des infections à HPV peuvent aller de l'asymptomatique subclinique aux cancers anogénitaux, un sous-ensemble de HPV connus sous le nom de HPV à haut risque (HR-HPV) sont spécifiquement associés aux cancers, causant jusqu'à 99% de tous les cas de cancer du col de l'utérus. Parmi ces cas, plus de 50 % sont signalés comme étant liés à une infection par le HPV16.[1]. Il est bien établi que l'oncoprotéine E6 du HPV joue un rôle central dans la pathogenèse du cancer du col de l'utérus et l'E6 du HPV16 est considérée comme l'E6 prototypique. Étant donné la grande diversité génotypique et phénotypique des HPV, ceux-ci sont souvent considérés comme des systèmes modèles pour les études interactomiques.[2]. En outre, la capacité de l'E6 à reconnecter les réseaux cellulaires facilitant la progression de la maladie, en conjonction avec son importance clinique, a suscité plusieurs études interactomiques. Ici, nous visons à caractériser l'interactome quantitatif de l'E6 prototypique du HPV16 en utilisant le Holdup natif (nHU).

Méthodologie et conception expérimentale :

Interactomique quantitative avec le nHU

L'interactome d'affinité de 16E6 a été mesuré en utilisant le nHU couplé à l'analyse par spectrométrie de masse (nHU-MS) et le Western Blot (nHU-WB).[3]. Une 16E6 recombinante fusionnée à la MBP a été produite et biotinylée dans E. coli (BL21), puis purifiée par purification d'affinité à l'amylose, suivie de l'élimination des agrégats solubles par ultra-centrifugation pendant une nuit et d'une résolution finale par chromatographie d'exclusion de taille. La 16E6-MBP biotinylée purifiée a été utilisée comme appât dans des essais nHU contre des proies de deux lignées cellulaires, HeLa (HPV+) et HaCaT (HPV-). Les essais nHU ont été réalisés conformément au protocole et aux principes décrits dans Zambo et al. en saturant la résine de streptavidine à une concentration d'appât (C_{bait}) de 42 µM et en utilisant de la résine saturée de MBP comme contrôle de non-liaison. Les mesures en un seul point ont été analysées à l'aide de nHU-Mass Spec. Les séries de titrage en huit points du 16E6-MBP ont été analysées à l'aide de nHU-WB.

Compléter les interactomes complets par des détails moléculaires obtenus par fragmentomique

Pour mieux comprendre les interfaces de liaison du 16E6 qui médient les interactions, des approches fragmentomiques ont été utilisées pour étudier les contributions des interfaces individuelles. Une bibliothèque de peptides naïfs contenant 86 motifs LxxLL biotinylés a été conçue et utilisée dans des essais HU[4] afin d'affiner le motif consensuel facilitant la liaison à la poche de liaison de 16E6 LxxLL. Les résultats des essais HU ont fait l'objet d'une validation croisée à l'aide de la BLI. Pour comprendre les contributions apportées par le PBM C-terminal du 16E6, un peptide biotinylé correspondant à la région a été utilisé comme appât dans des expériences nHU contre des extraits de HaCaT et de HeLa.

Détermination des mécanismes de liaison à l'aide d'essais nHU spike-in

Pour mieux comprendre les mécanismes de liaison qui sous-tendent l'interaction 16E6-UBE3A bien caractérisée et l'interaction tertiaire subséquente avec TP53, un essai de compétition nHU ou "spike-in" a été mis au point. En ajoutant aux extraits HaCaT utilisés pour le nHU un excès d'UBE3A recombinante dans différents états oligomériques, nous avons cherché à évaluer l'impact sur l'activité de liaison de l'UBE3A et le recrutement de TP53. Suite à la mise en œuvre réussie du spike-in nHU, nous avons adapté l'essai pour évaluer les mécanismes de liaison des nouveaux liants 16E6 identifiés dans cette étude. Les extraits cellulaires ont été complétés par un excès de peptide LxxLL ou de domaine PDZ purifié pour séquestrer la poche de liaison de LxxLL et le PBM C-terminal, respectivement. Nous avons ensuite contrôlé l'inhibition de la liaison aux nouveaux partenaires à l'aide de nHU-WB.

Kathleen Weimer ED414 Equipe : TRAVE
Numéro d'étudiant : 22120266

Étude de l'impact cellulaire du ciblage du PTPN14 de l'hôte par le 16E6

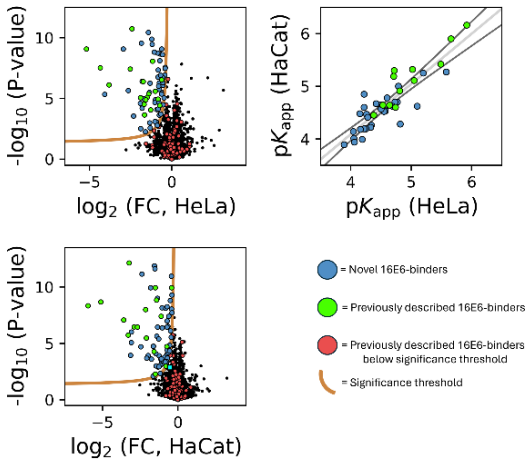
Pour distinguer le rôle indépendant de 16E6 ciblant la PTPN14 de l'hôte d'un rôle composé potentiel avec le HPV E7, les protéines ont été surexprimées dans des cellules HEK293T. Une lignée de cellules HEK293T surexprimant 16E6 a été transfectée transitoirement avec E7 et l'impact sur PTPN14 a été évalué en comparant l'effet de l'expression de 16E6 seule à celui de l'expression de 16E6 et E7 et à celui d'un contrôle (HEK293T). Pour évaluer les mécanismes de liaison potentiels médiant l'interaction entre 16E6 et PTPN14, un peptide contenant un consensus PY de PTPN14 a été synthétisé et utilisé dans une expérience nHU spike-in pour perturber les interactions avec les protéines contenant le domaine WW qui pourraient agir comme un intermédiaire facilitant la liaison 16E6-PTPN14. Les effets inhibiteurs du peptide sur la liaison de PTPN14 ont été contrôlés par nHU-WB.

Évaluation de la mesure dans laquelle le 16E6 modifie le système Ubiquitin-Proteosome de l'hôte

Pour évaluer l'impact cellulaire de l'interaction de 16E6 avec des partenaires jouant un rôle dans le système ubiquitine-protéosome de l'hôte, les E3 ubiquitine ligases, UBE3A et XIAP, et la dé-ubiquintinase, USP13, ont été désactivées, par siRNA, dans les cellules HeLa et HEK293T. Les niveaux de protéines de XIAP, USP13, UBE3A et TP53 ont été contrôlés.

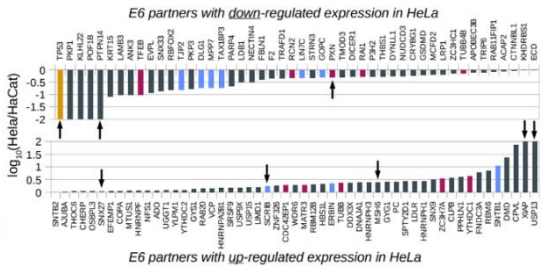
Résultats :

Cartographie de l'interactome d'affinité du 16E6 avec le nHU



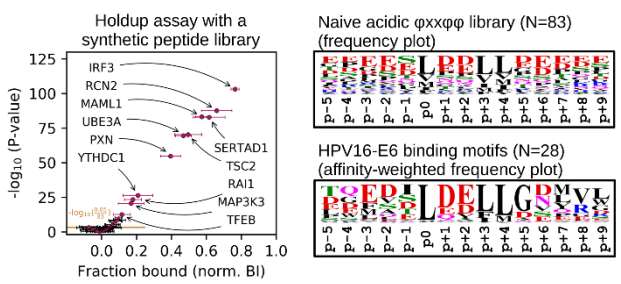
Au total, environ 4500 affinités apparentes (pK_{app}) ont été mesurées entre les deux lignées cellulaires et environ 100 partenaires ont été identifiés comme liant le 16E6 à l'aide du nHU. Un grand nombre des 16E6-binders identifiés avaient déjà été décrits dans d'autres études, apparaissant à la fois au-dessus (vert) et au-dessous (rouge) de notre seuil de signification (or). Nous avons caractérisé les liants restants comme nouveaux (vert) étant donné qu'ils n'ont pas été décrits précédemment. La bonne concordance entre les pK_{apps} des deux lignées cellulaires démontre que les affinités sont des propriétés intrinsèques des systèmes cellulaires.

Prédire le complexome 16E6



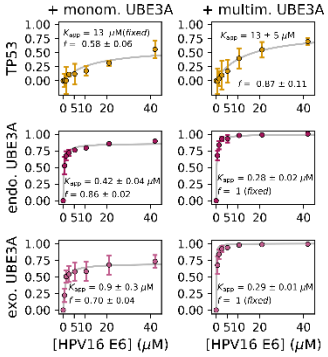
En combinant le pK_{app} mesuré avec le protéome quantifié, tous deux obtenus par Mass Spec, nous avons pu prédire la probabilité de formation d'un complexe binaire avec 16E6, ou le *complexome.*

La fragmentomique révèle des mécanismes de liaison du 16E6



Sur les 86 peptides LxxLL utilisés pour les essais HU, 28 ont été identifiés au-dessus du seuil de liaison. Les logos de séquence utilisant la pondération de l'affinité mettent en évidence les résidus clés qui pourraient contribuer aux interactions avec le 16E6.

Aborder les activités partielles de l'UBE3A avec le nHU spike-in



Les extraits cellulaires supplémentés avec l'UBE3A-MBP oligomérique ont montré une récupération de l'activité de liaison de l'UBE3A et le recrutement de TP53, mais l'activité de TP53 n'a jamais été atteinte à 100 %. Comparé au nHU spike-in avec l'UBE3A-MBP monomérique qui ressemble à la courbe de liaison et à l'activité partielle de l'UBE3A endogène, il semble que la forme monomérique de l'UBE3A soit la plus disponible au niveau cellulaire, bien que la forme oligomérique démontre une activité accrue.

Identification des interfaces d'interaction du 16E6 avec le nHU spike-in

En utilisant SNX27 et PXN comme étalons internes pour la liaison de PDZ et LxxLL à 16E6, respectivement, le pic nHU dans les essais a montré que les interfaces de 16E6 étaient séquestrées et que la liaison à ces interfaces était inhibée. Le contrôle de l'effet inhibiteur des nouveaux partenaires PTPN14 et USP13 a montré que l'interaction avec 16E6 était probablement médiée par la région C-terminale. Cependant, la nHU utilisant un peptide biotinylé correspondant au PBM C-terminal de 16E6 comme appât a entraîné une perte de la liaison de 16E6 à PTPN14 et USP13, ce qui suggère que l'interface médiatrice est un nouveau site C-terminal ou qu'il doit y avoir une interface secondaire nécessaire pour faciliter l'interaction à travers le PBM avec ces partenaires.

Évaluation de l'impact de l'expression des gènes E6 et E7 sur le gène PTPN14

La transfection transitoire de l'E7 a été validée dans des cellules HEK293T exprimant l'E6. L'impact de l'expression de l'E6 et des expressions de l'E6 et de l'E7 sur la PTPN14 a été contrôlé, mais dans les deux cas, aucun changement significatif n'a été observé. L'expérience de spike-in utilisant le peptide PY de PTPN14 n'a pas entraîné d'inhibition de la liaison 16E6-PTPN14.

Inhibition des composants du système ubiquitine-protéosome ciblés par le 16E6

L'inhibition de XIAP, UBE3A et USP13 à l'aide de siRNA a été validée dans des cellules HEK293T et des cellules HeLa exprimant l'E6 après 48 heures. Comme prévu, l'impact de l'abaissement de l'UBE3A a entraîné un rétablissement des niveaux d'expression de TP53. Cependant, l'inhibition de XIAP et USP13 n'a pas eu d'effet significatif sur l'expression de TP53 et l'inhibition de l'une de nos cibles n'a pas eu d'effet significatif sur l'expression des autres cibles.

Conclusions :

En mesurant les affinités protéiques, une propriété intrinsèque des systèmes cellulaires, nous disposons d'une métrique standard qui peut être utilisée pour comparer les données entre des méthodes orthogonales et créer une continuité entre les ensembles de données. En outre, la détermination de la force des interactions nous donne des indications biophysiques et biochimiques qui nous permettent de contextualiser les interactomes. Contrairement à d'autres méthodes d'interactomique à haut débit, le nHU perturbe la compartimentation, ce qui signifie que toute interaction physique possible se produira en dépit de sa pertinence biologique. Cela nous permet de conserver d'importantes propriétés biophysiques du système mesuré, tandis que des méthodes telles que les calculs complexomiques peuvent être utilisées pour analyser les données à l'aide d'une lentille biologique. En combinant les profils d'affinité avec la disponibilité cellulaire des partenaires, selon le protéome quantifié, nous pouvons prédire la probabilité de formation de complexes binaires et identifier les cibles critiques potentielles de 16E6. En utilisant cette méthode de pêche moléculaire informée par les données, nous avons pu caractériser plus en détail les fixateurs établis du 16E6 ainsi que tous les nouveaux. Les essais de compétition nHU ont démontré notre capacité à étudier les détails moléculaires, les activités de liaison dans le cas de

l'UBE3A et les interfaces de liaison dans le cas de l'USP13 et de la XIAP. Enfin, nous avons pu commencer à concevoir des essais fonctionnels pour explorer davantage les impacts biologiques des interactions identifiées dans cette étude.

Contributions :

Cette étude a été conçue par Gergo Gogl (GG) et Gilles Trave (GT). L'expérience nHU initiale qui a permis de quantifier l'interactome d'affinité de 16E6 a été réalisée par GG et les données obtenues ont été analysées par GG et Kathleen Weimer (KW). Les expériences Mass Spec et le traitement initial des données dans Perseus ont été réalisés par la plateforme Mass Spec de l'IGBMC. Toutes les expériences et analyses nHU ultérieures ont été conçues par KW avec les conseils de GG et exécutées par KW. Les peptides utilisés pour les essais nHU ont été synthétisés par la plate-forme peptidique de l'IGBMC. Les peptides utilisés dans la bibliothèque de peptides ont été synthétisés par Soren Ostergaard (SO). Les premiers Western Blots ont été réalisés par Boglarka Zambo (BZ) et tous les Western Blots suivants ont été réalisés par KW. Les essais cellulaires ont été conçus par BZ et les cellules HEK293T exprimant l'E6 utilisées dans les expériences ont été fournies par BZ. Les essais cellulaires ont été exécutés et analysés par KW. GT a assuré le financement de cette étude.

Références :

1 Dunne, E. F. & Park, I. U. HPV and HPV-Associated Diseases. *Infectious Disease Clinics of North America* **27**, 765-778, doi:https://doi.org/10.1016/j.idc.2013.09.001 (2013).
2 Neveu, G. *et al*. Comparative analysis of virus–host interactomes with a mammalian high-throughput protein complementation assay based on Gaussia princeps luciferase. *Methods* **58**, 349-359, doi:https://doi.org/10.1016/j.ymeth.2012.07.029 (2012).
3 Zambo, B., Morlet, B., Negroni, L., Trave, G. & Gogl, G. Native holdup (nHU) to measure binding affinities from cell extracts. *bioRxiv*, 2022.2006.2022.497144, doi:10.1101/2022.06.22.497144 (2022).
4 Gogl, G. *et al*. Quantitative fragmentomics allow affinity mapping of interactomes. *Nature Communications* **13**, 5472, doi:10.1038/s41467-022-33018-0 (2022).

Publications :

Les molécules interagissent. Mais quelle est la force et l'ampleur de cette interaction ? (2023). Kathleen Weimer, Boglarka Zambo, & Gergo Gogl. *BioEssays*. DOI : 10.1002/bies.202300007

En préparation : Modélisation de la formation du complexe protéique hôte-pathogène de l'oncoprotéine E6 du HPV16, à l'aide de méthodes d'interactionomique quantitative. Kathleen Weimer, Gergo Golgl, & Gilles Trave.

Communications orales et par affiches :

L'interactome d'affinité E6 du HPV16. Kathleen Weimer, Boglarka Zambo, Gergo Gogl, Gilles Trave. DNA Tumour Virus Meeting (ICGEB) 16-21 juillet 2024 Trieste, Italie (présentation orale)

L'interactome d'affinité du HPV16 E6. Kathleen Weimer, Boglarka Zambo, Gergo Gogl, Gilles Trave. Retraite du département BSI (IGBMC) 4-6 février 2024 Mont Sainte-Odile, France (présentation orale)

L'interactome d'affinité du HPV16 E6. Kathleen Weimer, Boglarka Zambo, Gergo Gogl, Gilles Trave. Journées scientifiques de l'AFIPP 27-29 septembre 2023 Fort-Mahon-Plage, France (présentation orale)

L'interactome d'affinité du HPV16 E6. Kathleen Weimer, Boglarka Zambo, Gergo Gogl, Gilles Trave. FEBS 47th Annual Conference 8-12 July 2023 Tours, France (poster presentation)

L'interactome d'affinité HPV16 E6 : une étude de cas démontrant la puissance de l'interactomique quantitative. Kathleen Weimer, Boglarka Zambo, Gergo Gogl, Gilles Trave. Journées de l'école doctorale 2023 de l'Université de Strasbourg 20-21 avril 2023 Strasbourg, France (flash talk)

Kathleen Weimer ED414 Equipe: TRAVE
Student number: 22120266

Thesis Abstract: Mapping the HPV16 E6-host protein interactome using quantitative interactomic methods to investigate biophysical properties of binding and model cellular complex formation

Introduction:

Human Papilloma Viruses (HPVs) are the most widespread sexually transmitted infection. While clinical presentations of HPV infections can range from subclinical asymptomatic to anogenital cancers, a sub-set of HPVs known as High-Risk HPVs (HR-HPVs) are specifically associated with cancers causing up to 99% of all cervical cancer cases. Of these cases >50% are reported as related to HPV16 infection[1]. It's been well established that the HPV oncoprotein E6 has a central role in cervical cancer pathogenesis and HPV16 E6 is regarded as the prototypical E6. Given the high genotypic and phenotypic diversity of HPVs, they are often considered model systems for interactomic studies[2]. Furthermore, the ability of E6 to rewire cellular networks facilitating disease progression in conjunction with its clinical significance has prompted several interactomic investigations. Here we aim to characterize quantitative interactome of the prototypical HPV16 E6 using native Holdup (nHU).

Methodology and Experimental Design:

Quantitative Interactomics with nHU

The affinity interactome of 16E6 was measured using nHU coupled to Mass Spec analysis (nHU-MS) and Western Blot (nHU-WB)[3]. A recombinant 16E6 fused to MBP was produced and biotinylated in E. coli (BL21), then purified with amylose affinity purification followed by the removal of soluble aggregates via overnight ultra-centrifugation and final resolution with size exclusion chromatography. The purified biotinylated 16E6-MBP was used as bait in nHU assays against preys of two cell lines, HeLa (HPV+) and HaCaT (HPV-). The nHU assays were performed according to the protocol and principles described in Zambo et al., saturating the streptavidin resin to a bait concentration (C_{bait}) of 42 μM and using resin saturated with MBP as a non-binding control. Single point measurements were analysed using nHU-Mass Spec. Eight-point titration series of the 16E6-MBP were analysed using nHU-WB.

Complementing full-length interactomes with molecular details obtained by fragmentomics

To further understand the 16E6 binding interfaces mediating interactions, fragmentomic approaches were used to investigate the contributions of the individual interfaces. A naive peptide library containing 86 biotinylated LxxLL motifs was designed and used in HU assays[4] to further refine the consensus motif facilitating binding to the 16E6 LxxLL binding pocket. The results of HU were cross-validated using BLI. To understand contributions made by the 16E6 C-terminal PBM, a biotinylated peptide corresponding to the region was used as bait in nHU experiments against HaCaT and HeLa extracts.

Determining binding mechanisms with nHU spike-in assays

To further understand the binding mechanisms underpinning the well characterized 16E6-UBE3A interaction and subsequent tertiary interaction with TP53, a nHU competition assay or "spike-in" was developed. By supplementing the HaCaT extracts used for nHU with an excess of recombinant UBE3A in different oligomeric states we sought to evaluate the impact on UBE3A binding activity and recruitment of TP53. Upon this successful implementation of the nHU spike-in we adapted the assay further to evaluate the binding mechanisms of novel 16E6-binders identified in this study. Cellular extracts were supplemented with an excess of LxxLL peptide or purified PDZ domain to sequester the LxxLL binding pocket and C-terminal PBM, respectively. We then monitored the inhibition of binding to novel partners using nHU-WB.

Investigating the cellular impact of 16E6 targeting of host PTPN14

To distinguish an independent role for 16E6 targeting host PTPN14 from a potential compound role with HPV E7, the proteins were overexpressed in HEK293T cells. An engineered HEK293T cell line overexpressing 16E6 was transiently transfected with E7 and the impact on PTPN14 was assessed by comparing the effect of 16E6 expression alone to that of 16E6 and E7 expression and a control (HEK293T). To evaluate potential binding mechanisms mediating interaction

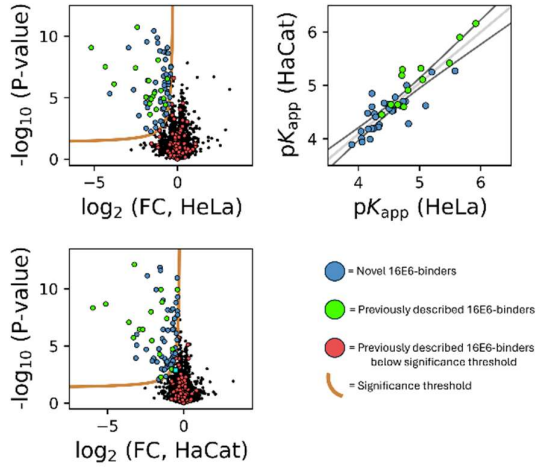
between 16E6 and PTPN14, a peptide containing a PY consensus from PTPN14 was synthesized and used in a nHU spike-in experiment to disrupt interactions with WW domain containing proteins that could act as an intermediary facilitating 16E6-PTPN14 binding. Inhibitory effects of the peptide on PTPN14 binding were monitored by nHU-WB.

Evaluating the extent to which 16E6 modifies the host Ubiquitin-Proteosome-System

To evaluate the cellular impact of 16E6 interaction with partners possessing roles in the host ubiquitin-proteosome-system, E3 ubiquitin ligases, UBE3A and XIAP, and de-ubiquintinase, USP13, were knocked down, via siRNA, in HeLa and HEK293T cells. Following knockdown, the protein levels of XIAP, USP13, UBE3A, and TP53 were monitored.

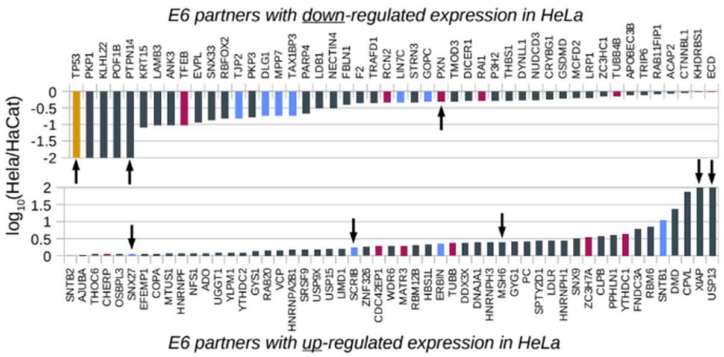
Results:

Mapping the 16E6 affinity interactome with nHU



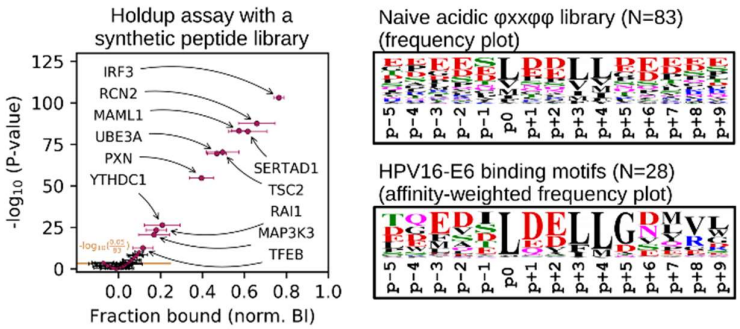
In total, approximately 4500 apparent affinities (pK_{app}) were measured between the two cell lines and about 100 partners were identified as 16E6-binders using nHU. Of the identified 16E6-binders many were previously described in other studies, appearing both above (green) and below (red) our significance threshold (gold). We have characterized the remaining binders as novel (green) considering they have not been previously described. Good agreement between the pK_{apps} of both cell lines demonstrates that affinities are intrinsic properties of cellular systems.

Predicting the 16E6 complexome



By combining the measured pK_{app} with the quantified proteome, both obtained by Mass Spec, we were able to predict the probability of binary complex formation with 16E6, or the *complexome.*

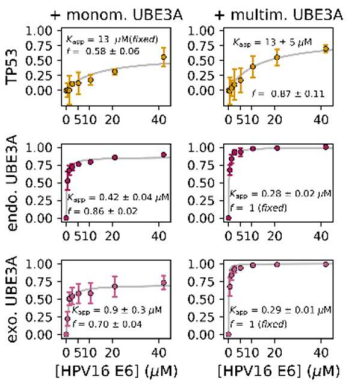
Fragmentomics reveal mechanistic insights to 16E6 binding



Of the 86 LxxLL peptides used for HU assays, 28 were identified above the binding threshold. Sequence logos using affinity weighting highlight key residues that could contribute to interactions with 16E6.

Addressing UBE3A partial activities with nHU spike-in

Cellular extracts supplemented with oligomeric UBE3A-MBP demonstrated recovery of UBE3A binding activity and recruitment of TP53, however 100% activity was never reached for TP53. Compared to the nHU spike-in with monomeric UBE3A-MBP which resembles the binding curve and partial activity of endogenous UBE3A, it appears that the monomeric form of UBE3A is the most cellularly available despite the oligomeric form demonstrating increased activity.



Identifying 16E6 interaction interfaces with nHU spike-in

Using SNX27 and PXN as internal standards for PDZ and LxxLL binding to 16E6, respectively, the nHU spike in assays showed that 16E6 interfaces were sequestered and binding to these interfaces were inhibited. Monitoring of the inhibitory effect of novel partners PTPN14 and USP13 showed interaction with 16E6 was likely mediated through the C-terminal region. However, nHU using a biotinylated peptide corresponding to the C-terminal PBM of 16E6 as bait resulted in a loss of 16E6 binding to PTPN14 and USP13 suggesting that the mediating interface is a novel C-terminal site or there must be a secondary interface necessary to facilitate interaction through the PBM with these partners.

Evaluating the impact of E6 and E7 expression on PTPN14

Transient transfection of E7 was validated in E6 expressing HEK293T cells. The impact of E6 expression and E6 and E7 expression on PTPN14 was monitored, however in either case there was no significant change observed. Spike-in experiment using the PTPN14 PY peptide did not result in inhibition of 16E6-PTPN14 binding.

Knock down of ubiquitin-proteosome-system components targeted by 16E6

Knock downs with siRNA of XIAP, UBE3A, and USP13 were validated in E6 expressing HEK293T cells and HeLa cells after 48 hours. As expected, the impact of knocking down UBE3A resulted in a recovery of TP53 expression levels. However, the knockdown of XIAP and USP13 had no significant effect on TP53 expression and the silencing of any one of our targets had no significant effect on the expression of the other targets.

Conclusions:

In measuring protein affinities, an intrinsic property of cellular systems, we are provided with a standard metric that can be used to compare data across orthogonal methods and create continuity among datasets. Additionally, determining the strength of interactions gives us biophysical and biochemical insights that allow us to contextualize interactomes. Unlike other high throughput interactomic methods, nHU disrupts compartmentalization meaning any possible physical interaction will occur despite its biological relevance. This allows us to retain important biophysical properties of the system being measured, while methods such as complexomic calculations can be used to analyze the data using a biological lens. In combining affinity profiles with the cellular availability of partners, according to the quantified proteome, we can predict the likelihood of binary complex formation and identify potential critical targets of 16E6. Using this method of data informed molecular fishing we were able to further characterize established 16E6-binders as well all novel ones. The nHU competition assays demonstrated our ability to investigate molecular details, binding activities in the case of UBE3A and binding interfaces in the case of

USP13 and XIAP. Finally, we were able to begin designing functional assays to further explore the biological impacts of interactions identified in this study.

Contributions:

This study was conceptualized by Gergo Gogl (GG) and Gilles Trave (GT). The initial nHU experiment resulting in the quantified affinity interactome of 16E6 was performed by GG and the resulting data was analyzed by GG and Kathleen Weimer (KW). Mass Spec experiments and initial data processing in Perseus were performed by the IGBMC Mass Spec Platform. All subsequent nHU experiments and analysis were designed by KW with guidance from GG and executed by KW. Peptides used for nHU assays were synthesized by the IGBMC peptide platform. Peptides used in the peptide library were synthesized by Soren Ostergaard (SO). Initial Western Blots were performed by Boglarka Zambo (BZ) with all subsequent Western Blots performed by KW. Cellular assays were designed by BZ and E6 expressing HEK293T cells used in experiments were provided by BZ. Cellular assays were executed and analyzed by KW. GT secured funding for this study.

References:

1 Dunne, E. F. & Park, I. U. HPV and HPV-Associated Diseases. *Infectious Disease Clinics of North America* **27**, 765-778, doi:https://doi.org/10.1016/j.idc.2013.09.001 (2013).
2 Neveu, G. *et al*. Comparative analysis of virus–host interactomes with a mammalian high-throughput protein complementation assay based on Gaussia princeps luciferase. *Methods* **58**, 349-359, doi:https://doi.org/10.1016/j.ymeth.2012.07.029 (2012).
3 Zambo, B., Morlet, B., Negroni, L., Trave, G. & Gogl, G. Native holdup (nHU) to measure binding affinities from cell extracts. *bioRxiv*, 2022.2006.2022.497144, doi:10.1101/2022.06.22.497144 (2022).
4 Gogl, G. *et al*. Quantitative fragmentomics allow affinity mapping of interactomes. *Nature Communications* **13**, 5472, doi:10.1038/s41467-022-33018-0 (2022).

Publications:

Molecules interact. But how strong and how much? (2023). Kathleen Weimer, Boglarka Zambo, & Gergo Gogl. *BioEssays*. DOI: 10.1002/bies.202300007

In preparation: Modelling host-pathogen protein complex formation of the HPV16 oncoprotein, E6, using quantitative interactomic methods. Kathleen Weimer, Gergo Golgl, & Gilles Trave.

Oral and Poster Communications:

The HPV16 E6 affinity interactome. Kathleen Weimer, Boglarka Zambo, Gergo Gogl, Gilles Trave. DNA Tumour Virus Meeting (ICGEB) 16-21 July 2024 Trieste, Italy (oral presentation)

The HPV16 E6 affinity interactome. Kathleen Weimer, Boglarka Zambo, Gergo Gogl, Gilles Trave. BSI Department Retreat (IGBMC) 4-6 February 2024 Mont Sainte-Odile, France (oral presentation)

The HPV16 E6 affinity interactome. Kathleen Weimer, Boglarka Zambo, Gergo Gogl, Gilles Trave. AFIPP Scientific Days 27-29 September 2023 Fort-Mahon-Plage, France (oral presentation)

The HPV16 E6 affinity interactome. Kathleen Weimer, Boglarka Zambo, Gergo Gogl, Gilles Trave. FEBS 47th Annual Conference 8-12 July 2023 Tours, France (poster presentation)

The HPV16 E6 affinity interactome: a case study demonstrating the power of quantitative interactomics. Kathleen Weimer, Boglarka Zambo, Gergo Gogl, Gilles Trave. University of Strasbourg Doctoral School Days 2023 20-21 April 2023 Strasbourg, France (flash talk)

Acknowledgements

I would like to extend a thank you to everyone that supported me throughout this journey. I especially would like to express gratitude to Gilles, who welcomed me into his lab allowing me to escape a difficult situation. A quality that I highly admire about Gilles is that he will endlessly support a person in the pursuit of science.

This project would have not been possible without the training and guidance provided by Gergo and Bogi. I appreciate all the time and dedication these two have given to the project. When they agreed to take on this project they displayed a great amount of faith in me and everything they taught me will continue to serve me for my entire career. I can't thank them enough for this.

I'd like to thank everyone in the Trave team. It's been a wonderful and supportive environment to work in. At some point I have received help, ideas, or feedback from every member of the team that has helped me grow as a researcher. A special thank you goes my team members Wen, Theo, and Alex who have offered me an extra set of hands when I needed them. It brings me great joy to know the inhibitor project I started with Alex will continue with Theo. I can't think of a person more capable or up for the task.

I must also mention my father and mother, James and Diana Weimer who have always pushed me to follow my aspirations. Without their constant love and support I know this would not have been possible without them.

I have made a few exceptional friends along the way—Victoria, Paula, and Lionel I am so glad that this opportunity has led me to you and that we can share in this together. Then there are the friends that have always been by my side even if we are separated by an ocean or more. Ellen, Kevin, Hannah, and Hannah thank you for continuing to share in this life with me, you have become my favorite people to celebrate and commiserate with.

The list goes on as I am an exceptionally lucky person to know so many wonderful people. My final thank you goes to an incredible person that has more faith in me than I deserve. In the world of science, I constantly push myself to take on new perspectives and find new ways of looking at problems. Well, this person has completely changed my perspective on life. Thank you Erwan for all that you have done. I can always count on you to believe in me even if I don't believe in myself.