UNIVERSITÉ DE STRASBOURG

THESE

présente par

Igor ORLOV

le 19 Mai 2010

pour obtenir le titre de

Docteur de l'Université de Strasbourg

Discipline: Sciences de vivant

Etude structurale du complexe des récepteurs nucléaires RXR/VDR liés à l'ADN

Commission d'examen:

Mr. Eric Westhof	**Rapporteur interne**
Mr. William Bourguet	**Rapporteur externe**
Mr. Hong Zhou	**Rapporteur externe**
Mr. Dino Moras	**Directeur de thèse**
Mr. Bruno Klaholz	**Directeur de thèse**

Table of contents

Acknowledgements 3
Abbreviations 4
Résume de la thèse 5
Summary 11
1. Introduction 15
1.1 Nuclear receptors 15
1.1.1 The nuclear receptors superfamily 15
1.1.2 The general model of nuclear receptor action . . . 18
1.1.3 Domain structure of nuclear receptors 19
1.1.4 Hormone response elements 23
1.1.5 Agonism, antagonism and its structural basis . . . 24
1.2 Cryo-electron microscopy and 3D reconstruction 26
1.2.1 Electron optics and transmission electron microscopy . 27
1.2.2 Negative staining specimen preparation 30
1.2.3 Ice-embedded specimen preparation 32
1.2.4 3D reconstruction framework 33
1.2.5 Contrast transfer function and CTF correction . . . 35
1.2.6 Alignment 37
1.2.7 Angular orientation of individual particles . . . 41
1.2.8 3D reconstruction 45
1.2.9 Iterative Refinement 47
1.2.10 Evaluation of the 3D reconstruction 48
1.2.11 Multivariate statistical analyses 49
2. Aims of the project 52
3. Methods 53
3.1 The purification of complexes 55
3.2 Initial model 56
3.3 200 kV structure 59
3.4 100 kV test structure 60
3.5 100 kV test structure with a 35mer DNA 61
3.6 3D densities map interpretation 62
4. Results 63
4.1 3D reconstruction software 63
4.2 RXR/VDR heterodimer complex with its target DNA . . . 74
5. Conclusions and perspectives 87
6. References 93
Publications 113
Oral presentations 114
Posters 115

Acknowledgements

I would like to thank Dr. Dino Moras and Dr. Bruno Klaholz for giving me the opportunity to undertake such a hard but exciting project. Thanks to Dr. Natacha Rochel and all members of her group for the preparation of the samples. Thanks to Dr. Jean-Claude Thierry and Dr. Patrick Schultz for discussions and moral support.

Thanks to everybody in Bruno Klaholz and Patrick Schultz EM groups.

Thanks to Dr. Elena Orlova for discussions, suggestions and moral support.

Abbreviations

2D	-	two dimensional
3D	-	three dimensional
Å	-	Angstrom
ACF	-	auto correlation function
AF1, 2	-	activation function 1, 2
CCF	-	cross correlation function
Cryo-EM	-	cryo-electron microscopy
CTF	-	contrast transfer function
DACF	-	doubled auto correlation function
DBD	-	DNA binding domain
DNA	-	deoxyribonucleic acid
EM	-	electron microscopy
ER	-	estrogen receptor
FSC	-	fourier shell correlation
LBD	-	ligand binding domain
RNA	-	ribonucleic acid
mRNA	-	messenger ribonucleic acid
MRA	-	multiple reference alignment
MSA	-	multivariate statistical analysis
PPAR	-	peroxisome proliferator activated nuclear receptor
PSF	-	point spread function
RXR	-	retinoid acid nuclear receptor
VDR	-	vitamin D nuclear receptor

Résume de la thèse

Les récepteurs nucléaires sont une classe de protéines qui régulent l'expression de gènes spécifiques de manière à contrôler des processus moléculaires impliqués dans le développement, l'homéostasie et le métabolisme. Ces protéines ont la capacité de se lier directement à l'ADN qui est unique en comparaison avec d'autres classes de molécules réceptrices, et en présence du ligand ils régulent l'expression des gènes adjacents à la région promotrice. La liaison du ligand par les récepteurs nucléaires résulte en un changement conformationnel de la structure du récepteur qui l'active, entraînant une activation de l'expression génique.

L'importance de l'effet médical de stéroïdes et des vitamines dans un organisme humain est bien connue mais les mécanismes moléculaires des l'interactions entre les molécules de ligand, les récepteurs nucléaires et l'ADN doivent encore être révélés. Jusqu'à présent, deux domaines du récepteur nucléaire (domaine de liaison du ligand, LBD, et domaine de liaison à l'ADN, DBD) ont été étudiées par cristallographie aux rayons-X dans différents complexes avec des ligands naturels et pharmacologiques. Les difficultés d'obtenir des cristaux d'un complexe entier et fonctionnel du complexe ADN des récepteurs nucléaires appel à l'utilisation de méthodes complémentaires à la cristallographie aux rayons-X. Nous avons utilisé la cryo-microscopie électronique de particules isolées (cryo-ME) afin d'étudier l'architecture d'un ensemble représentatif des récepteurs nucléaires avec l'ADN.

La difficulté majeure pour l'étude des complexes ADN nucléaire des récepteurs par cryo-microscopie électronique est la taille relativement réduite de

l'ensemble qui est d'environ 100kDa. A titre de comparaison, le ribosome - un objet typique de cryo-ME a une masse moléculaire d'environ 2.3 MDa. La taille du complexe conduit à un contraste extrêmement faible de l'image. Cela signifie non seulement qu'il y a un problème pour identifier les particules individuelles, mais aussi l'intensité du signal reste faible en comparaison avec l'intensité du bruit quantique typique d'un détecteur. Le tampon qui est nécessaire pour maintenir le complexe récepteur nucléaire intact dans des conditions cryo contribue également à un contraste plus faible.

La qualité de la reconstruction 3D dépend du nombre d'images dans un ensemble de données, la précision de l'alignement de l'image, l'attribution des angles pour les images des particules, la distribution angulaire des points de vue des particules, et la symétrie de l'objet. La microscopie électronique nécessite des logiciels numériquement stables et des moyens de calculs efficaces pour pouvoir traiter des grands jeux de données pouvant produire des reconstitutions en 3D à résolution quasi-atomique. La préparation des échantillons utilisés en ME à la fois avec coloration négative et congélation rapide dans de l'éthane liquide permet de fixer les molécules sur le film de carbone ou dans la glace vitreuse (cryo-ME). Les orientations moléculaires sont fixées dans la glace et, par conséquent, une image du complexe est liée à l'orientation de la molécule par rapport au faisceau d'électrons, et peuvent être utilisés pendant la procédure de reconstruction. La préparation cryo-ME des échantillons, en plus de préserver les molécules dans un état hydraté, fonctionnel, dispose d'un avantage à la coloration négative échantillons fournissant la répartition plus égale des orientations. Toutefois, la cryo-préparation ne garantit pas une répartition uniforme des points de vue, parce que les zones hydrophobes de la surface de la molécule peuvent conduire à des orientations préférentielles vers l'interface air-liquide à l'intérieur du film liquide mince qui s'étend sur les trous de

carbone. L'existence d'orientations préférentielles conduit au renforcement de la densité dans ce sens et mène à des distorsions structurales. Par conséquent, une répartition inégale des orientations complexes doit être pris en considération et doit être compensée lors de la réalisation des reconstitutions en 3D.

En tant que contribution à la méthodologie de reconstruction 3D des particules isolées, nous avons également fait un développement de logiciels dans le but d'améliorer la fiabilité, la vitesse et la stabilité numérique de logiciel de reconstruction 3D. Nous avons développé un logiciel de reconstruction 3D basée sur des projections inversées pondérées. En outre, nous avons introduit une nouvelle approche de la compensation des orientations préférable basées sur les diagrammes de Voronoi qui permet de surmonter le problème de la distribution non uniforme de vues moléculaires. La pondération des projections est importante en particulier pour la détermination des structures asymétriques (par exemple les récepteurs nucléaires, des ribosomes ou d'ARN polymérase), tandis qu'il est moins critique pour les cas de haute symétrie, quand l'unité asymétrique devient plus petite.

Nous avons déterminé la structure du complexe VDR-RXR des récepteurs nucléaires sur un fragment d'ADN court (20 paires de base) dans des conditions cryo à une résolution de 10-12Å. Pour la première fois, l'architecture d'un complexe NR complet a pu être déterminée à haute résolution à l'aide de cryo-ME. Nous avons également réussi à établir le protocole de collecte de données de petits complexes protéines-ADN dans des conditions cryo. L'instabilité numérique de l'analyse de corrélation croisée sur de telles données à faible contraste nécessite l'utilisation de l'analyse statistique des images en deux dimensions pour pouvoir faire une reconstruction en trois dimensions, et aussi une analyse statistiques en trois dimensions permettant de révéler une éventuelle hétérogénéité des données.

Afin de répondre à la localisation des domaines des récepteurs nucléaires individuelles plus en détail et de fournir une interprétation complète nous avons recalé la structure cristalline connue des LBDs et de l'ADN/DBD dans la carte cryo-ME. Pour cela, nous avons construit un modèle d'RXR / VDR LBD sur la base de la structure cristalline des RXR / RAR LBD dans laquelle la partie RAR a été remplacé par la structure cristalline du VDR LBD, ce qui prend en compte la conservation élevée des l'homo- et hétéro-dimères LDB. Par ailleurs, aucune structure de l'ADN-lié au DBD structure fonctionnelles RXR / VDR est disponible pour un élément DR3 avec RXR et VDR DBDs sur les extrémités 5' et 3' de l'ADN respectivement. Nous avons donc construit un modèle du complexe DBD RXR / VDR / DR3 dérivé de la structure cristalline du DBD RXR / TR / DR4 et d'une partie d'un complexe VDR- 5'-VDR / RXR-3' lié à l'ADN. Le recalage de ce modèle et de celui de RXR / VDR LBD dans la carte de densité cryo-ME a permis d'annoter sans ambiguïté RXR et VDR, tout en attribuant la main (chiralité) du complexe.

L'ajustement détermine l'orientation 5'-3' de l'ADN à cause de la précision des densités de DBDs, et il est compatible avec la contrainte de la connectivité entre les LBDs et DBDs de RXR et VDR, respectivement. La carte indique clairement que le DBDs est positionné à cote de l'ADN, avec une partie sur le côté face de la LBDs et l'ADN et non pas en face d'eux. Le recalage permet aussi de définir la polarité des complexes à l'égard de l'ADN, en montrant que le point d'ancrage principal à l'ADN se produit sur le côté 5 'à travers la région charnière et le DBDs plutôt que par exemple sur le centre de l'élément de réponse, avec le RXR DBD en effet sur l'extrémité 5'-comme on peut être prévue pour un élément DR3.

La structure du complexe révèle que hétéro dimère LBD est une unité assez rigide et stable au sein du complexe même en présence de DBDs et l'ADN. Cela est compatible avec la conservation de séquence élevé de l'interface LBD et la stabilité

des hétéro dimères LBD qui a facilité la détermination des structures cristallines de plusieurs (RXR / RAR, RXR / PPAR, RXR / LXR) mais pas encore pour RXR / VDR. La structure actuelle suggère que les RXR / VDR LBD adopte la même disposition que RXR / RAR ou RXR / PPAR LBDs. Notamment, la carte indique la position des hélices centrales de l'interface LBD comprenant hélices H9 et H10/11 de VDR et RXR, et un élément de forte densité est imputable à l'interface de contact entre l'hélice LDB H4 (VDR) et l'hélice H7 (RXR). L'orientation du dimère LBD est approximativement perpendiculaire à l'ADN, avec seulement une inclinaison faible (le pseudo double axe de symétrie par hétéro dimère LBD est incliné ~ 20 º vers le côté en aval de l'ADN, et hétéro dimère est pivoté de ~ 5 º sur le côté VDR). En présence d'un ligand agoniste comme dans le complexe de l'espèce, la trans-hélice H12 d'activation est connue pour fermer la cavité ou se lie le ligand et adopter une conformation active requise pour le recrutement de co-activateurs comme la co-activateurs des récepteurs stéroïdiens (SRC). Notamment, la structure du complexe RXR / VDR avec l'ADN montre que hélice H12 de transactivation et le site de liaison du co-activateur VDR pointe vers la direction 3 'et side-on à l'ADN (en face de la DBDs), suggérant que les co-activateur agit en aval de l'élément de réponse NR. Ceci est cohérent avec l'idée que les récepteurs nucléaires auxquels l'ADN est lié maintiennent l'helice H12 de trans-activation accessible pour le recrutement de la chromatine, la modification de co-régulateur de protéines qui agissent sur la région nucléosome assortis de l'ADN en aval.

L'étude déplace ainsi la limite de faisabilité généralement acceptée de 200-300kDa en montrant que la cryo-microscopie électronique est capable de l'étude des complexes moléculaire autour 100kDa, à condition qu'un affinement de la structure soit fait. Il convient également de noter qu'une symétrie interne serait bénéfique pour le traitement des données mais pas indispensable comme le montre le présent

ouvrage. Des études de ce type pourraient à l'avenir être facilitées par des caméras CCD plus sensibles et, éventuellement, par des filtres d'énergie et de correcteurs d'aberration. Le présent ouvrage montre la voie à des études de structure d'un ensemble d'objets macromoléculaires relativement petits y compris celles de protéine / acide nucléique complexes qui représentent la majorité des objectifs de la biologie structurale actuelle, mais ont été difficiles à cristalliser ou jugée trop petite pour la méthode classique de cryo-EM. Ainsi, l'étude des complexes NR avec différents types d'éléments de réponse, comme des répétitions inversées, renversés et directe peut être envisagée afin de charactériser lers architectures, spécifiques.

Les résultats de ce travail de thèse ont été soumis au journal Nature (Igor Orlov, Natacha Rochel, Dino Moras, Bruno P. Klaholz, Cryo-electron microscopy structure of the 100kDa full nuclear receptor RXR/VDR heterodimer complex with its target DNA) et un manuscrit sur le logiciel de reconstruction 3D est en préparation.

Summary

Nuclear receptors are a class of proteins that regulate the expression of specific genes thereby controlling the molecular events during development, homeostasis and metabolism. These proteins have the ability to directly bind to DNA which is unique in comparison with other classes of receptor molecules and regulate the expression of adjacent genes. The binding of a ligand to a nuclear receptor results in conformational changes in the receptor structure which in turn activates the receptor resulting in up-regulation of gene expression. It should be mentioned, that recent studies show that some nuclear receptors can regulate the gene expression without binding a ligand.

The importance of the medical effect of steroids and vitamins on the human organism is well known but the molecular mechanisms of the interaction between ligand molecules, nuclear receptors and the DNA are still to be revealed. Up to present date the two core domains of the nuclear receptor (ligand-binding domain, LBD and DNA-binding domain, DBD) have been studied by X-ray crystallography in different complexes with natural and pharmacological ligands. The difficulties of obtaining crystals of a fully functional nuclear-receptor DNA complex compel to the usage of methods complementary to the X-ray crystallography. We have used single particle cryo-electron microscopy (cryo-EM) to investigate the architecture of a representative nuclear receptor complex with DNA.

The major difficulty for studying the nuclear receptor DNA complexes by cryo-electron microscopy is the rather small size of the complex that is approximately 100kDa. For comparison, the ribosome – a typical object for cryo-EM has a molecular weight of approximately 2-3 MDa. The small size of the complex leads to extremely

low image contrast. This means that it is difficult to identify individual particles. The salt concentration that is needed to keep the nuclear receptor complex intact under cryo-conditions also contributes to a lower contrast.

The quality of a 3D reconstruction essentially depends on the number of images in a data set, the accuracy of the image alignment, the precision of the assignment of angles to the particle images, the angular distribution of the particle views, and the symmetry of the object. Electron microscopy analysis requires numerically stable and computationally effective methods such that large data sets can produce 3D reconstructions at high resolution. Sample preparation used in EM both with negative stain and plunge-freezing in liquid ethane leads to the fixation of molecules on the carbon film or within vitreous ice (cryo-EM). Molecular positions are fixed in ice and, therefore, an image of the complex is linked to the orientation of the molecule with respect to the electron beam, and can be used during the reconstruction procedure. The cryo-EM sample preparation, in addition to preserving the molecules in a hydrated, functional state, has the advantage in comparison with negatively stained samples of providing a more even distribution of orientations. However, the cryo-preparation does not ensure an even distribution of views, because hydrophobic areas on the molecule surface may lead to preferential orientations towards the air-liquid interface within the thin liquid film that spans the holes in holey carbon. The existence of preferential orientations leads to a strengthening of densities in this direction and leads to distortions of structural features. Therefore, an uneven distribution of the complex orientations should be taken into account and compensated when performing 3D reconstructions.

As a contribution to the methodology of single particle 3D reconstruction we also did a software development in order to improve the reliability, speed and numerical stability of single particle 3D reconstruction software. We have developed

a 3D reconstruction software based on weighted-back projections. Furthermore, we have introduced a new approach for the compensation of preferable orientations based on Voronoi diagrams that overcomes the problem of non-uniform distribution of molecular views. Weighting of projections is important in particular for the determination of asymmetrical structures (e.g. nuclear-receptors, ribosome or RNA polymerase), while it is less critical for high-symmetry cases when the asymmetric unit becomes rather small.

The main result of this PhD study is the determination of the structure of the VDR-RXR nuclear receptor on a short DNA fragment (20 base pairs). For the first time, the architecture of a full NR complex has been determined using single-particle cryo-EM under cryo-conditions at a resolution of 10-12Å. We established a protocol for data collection of small protein-DNA complexes under cryo conditions. The numerical instability of cross-correlation analysis on such a low contrast dataset required the usage of the statistical analysis of two-dimensional images in order to improve the signal to noise ratio, and also a three-dimensional statistical analysis to reveal possible heterogeneity of the data.

The structure of the complex reveals that the LBD heterodimer is a rather rigid and stable unit within the complex also in the presence of DBDs and DNA. This is consistent with the high sequence conservation of the LBD interface and the stability of LBD heterodimers The orientation of the LBD dimer is approximately perpendicular to the DNA, with only a small tilt. In the presence of an agonist ligand as in the complex here, the trans-activation helix H12 is known to close the occupied ligand-binding pocket and adopt an active conformation required for the recruitment of co-activators such as steroid receptor co-activators (SRCs). Notably, the structure of the RXR/VDR DNA complex shows that transactivation helix H12 and the co-activator binding site of VDR point towards the 3’ direction and side-on to the DNA (opposite to

the DBDs), suggesting that the co-activator acts downstream from the NR response element. This is consistent with the idea that DNA-bound nuclear receptors maintain the trans-activation helix H12 area accessible for recruiting chromatin-modifying co-regulator proteins that act on the nucleosome-bound region of the downstream DNA.

The results of this PhD work have been submitted to Nature (Igor Orlov, Natacha Rochel, Dino Moras, Bruno P. Klaholz, Cryo-electron microscopy structure of the 100kDa full nuclear receptor RXR/VDR heterodimer complex with its target DNA) and a further manuscript on the 3D reconstruction software is in preparation.

1. Introduction

1.1 Nuclear receptors

Gene expression is regulated via proteins called transcription factors and nuclear receptors that bind to specific regulatory sequences in the DNA and modulating the transcription of genes into messenger RNA. They can activate or repress target genes by binding directly to DNA response elements as homo or hetero-dimers or by binding to other classes of DNA-bound transcription factors. These proteins are often present in the cell in an inactive form and may regulate gene expression in a ligand dependent manner forming complexes with molecules that appear to serve as coactivators or corepressors, causing local modification of chromatin structure in order to regulate expression of their target genes. However recent studies show the ability of NRs to function in a ligand-independent manner. Working in concert with other proteins, nuclear receptors regulate the expression of specific genes thereby controlling the development, homeostasis, and metabolism of the organism. Several members of nuclear receptor family are directly associated with human malignancies including breast cancer, prostate cancer and leukaemia. The pathogenesis of each of these diseases is underpinned by the activities of a member of the superfamily; estrogen receptor-a (ERa) in breast cancer, androgen receptor (AR) in prostate cancer, and retinoic acid receptor a (RARa) in acute promyelocytic leukaemia (Leonard et al. 2007).

1.1.1 The nuclear receptors superfamily

The superfamily of nuclear receptors contains more than 200 different proteins. Evolutionary analysis of the receptors has led to a subdivision in six different subfamilies (Laudet 1997, Aranda & Pascual 2001). This classification is homology based. One large family is formed by thyroid hormone receptors (TRs), retinoic acid receptors (Mangelsdorf & Evans, 1995), vitamin D receptors (VDRs) and peroxisome proliferator-activated receptors (PPARs) as well as different orphan receptors. Ligands for some of orphan receptors have been recently identified. The second subfamily contains the retinoid X receptors (RXRs) together with chicken ovalbumin upstream stimulators (COUPs), hepatocyte nuclear factor 4 (HNF4), testis receptors (TR2) and receptors involved in eye development (TLX and PNR). RXRs bind 9-cis-retinoic acid and play an important role in nuclear receptor signaling, as they are partners for different receptors that bind as heterodimers to DNA. Ligands for other receptors have not been identified, whereas long-chain fatty acid acyl-CoA thioesters may be endogenous ligands for HNF4. The third family is formed by the steroid receptors and the highly related orphan receptors estrogen-related receptors (ERRs). The fourth, fifth, and sixth subfamilies contain the orphan receptors NGFI-B, FTZ-1/SF-1, and GCNF, respectively (Giguere 1999).

Recent functional studies allowed classifying nuclear receptors according to their functional mechanisms of action into three major classes. Ligand binding to type one nuclear receptors (Figure 1.1A) in the cytoplasm results in the dissociation of heat shock proteins, homo-dimerization and translocation into the nucleus with subsequent binding to specific hormone response element. Typical members of type one are steroid- and thyroid-hormone receptors and oestrogen receptors (Linja et al., 2004). The receptors of a second class are located in the nucleus regardless of the ligand binding status and bind to DNA as hetero-dimers (Figure 1.1B). In the absence of ligand, type two nuclear receptors are often complexed with corepressor proteins.

That means that type two nuclear receptors are capable to function in a ligand-independent manner. Type tree nuclear receptor class is similar to type one family but primarily binds to DNA as monomers. Type two nuclear receptors include for example the retinoic acid receptor, retinoid X receptor and thyroid hormone receptor (Klinge et al., 1997). The third class of nuclear receptors also is similar to type one receptors in that both classes bind to DNA as homodimers. However, type three nuclear receptors, in contrast to type one, bind to direct repeat instead of inverted repeat HREs. These receptors are orphan receptors (Wortham et al. 2007).

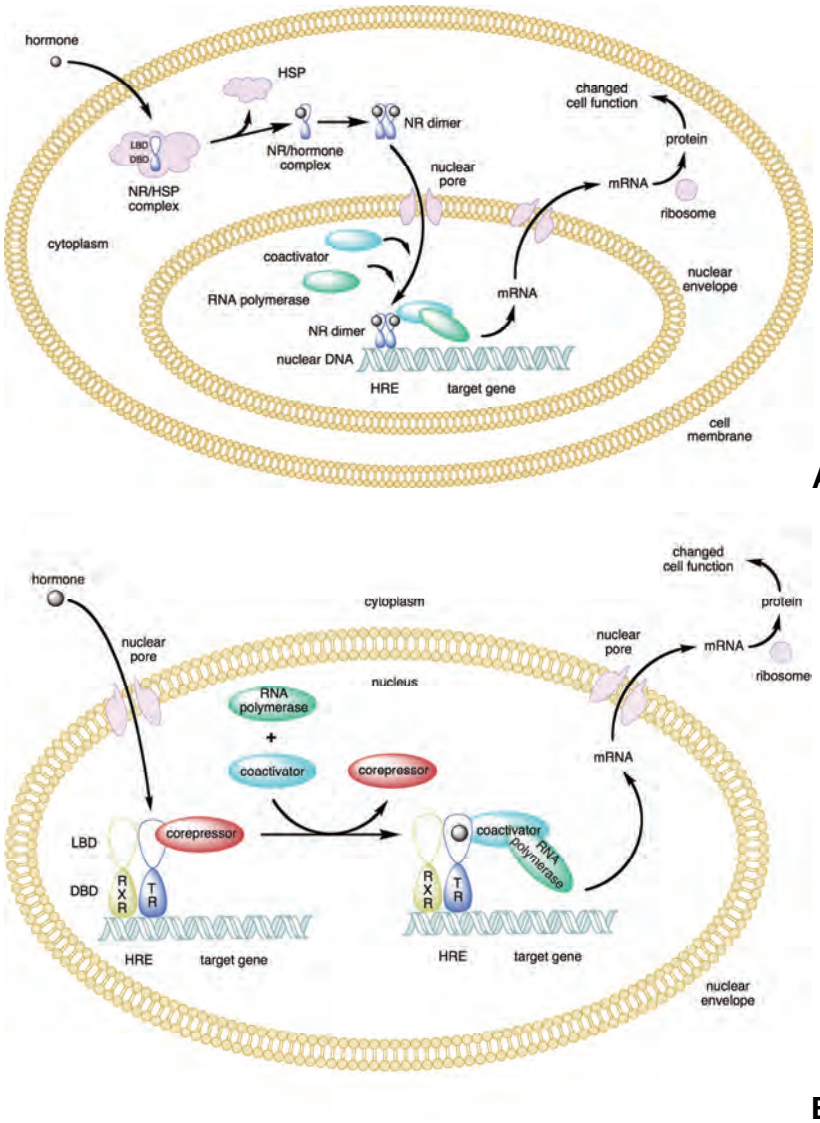


Figure 1.1 **A.** The action mechanism of a class I nuclear receptor. **B.** The action mechanism of a class II nuclear receptor.

1.1.2 The general model of nuclear receptor action

The repression or silencing state that is characterized by the absence of ligand (Figure 1.2A). The NR is bound to the response element and recruit co-repressor that associate with Sin3 (Alland et al. 1997; Hassing et al. 1997; Heinsel et al. 1997; Kadosh and Struhl 1997; Latherty et al. 1997; Nagy et al. 1997; Zhang et al. 1997; Parzin and Kadonaga 1997). The Sin3 in turn binds to the histone deacetylase (HDAC). The HDAC in this repression complex is assumed to provoke the histone deacetylation and as a result chromatin condensation. That processus blocks the assembly of the transcription initiation complex at the specific site.

The activation state that is the response to the binding of the ligand (Figure 1.2B). Response element bound nuclear receptor with a ligand recruits the co-activator, which in turn bind to hystone acetyl transferases (Bannister and Kouzarides 1996; Ogryzko et al. 1996; Yang et al. 1996; Wang et al 1997). Hromatin is “opening

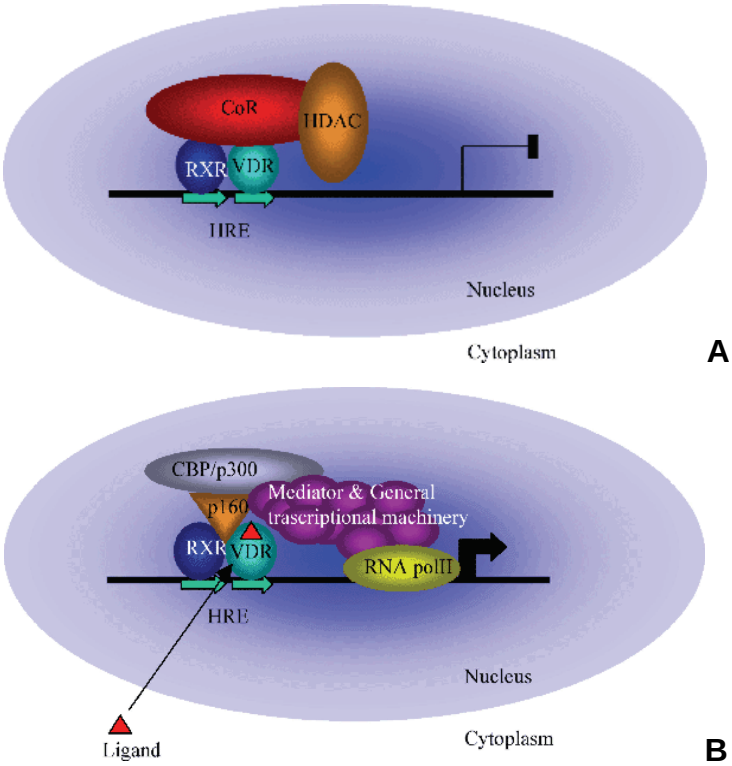


Figure 1.2 **A.** Silencing or transrepression mechanism. **B.** Transactivation mechanism.

up" due to histone acetylation. After that the formation of a trancription initiation complex becomes possible.

1.1.3 Domain structure of nuclear receptors

Most nuclear receptors have a common modular domain structure (Figure 1.3) consisting of ligand independent activation domain (A/B domain), a DNA-binding domain (DBD), hinge region and a ligand-binding domain (LBD) (Aranda 2001, Glass 2001). Some receptors also a COOH terminal region (F) of unknown function. The DBD is highly conserved domain that mediates binding of nuclear-receptor monomers, homodimers and heterodimers to specific sequences of DNA called hormone response elements. Most response elements contain two or more closely spaced core recognition motifs, each of which comes into contact with a single DBD. Hinge region is a flexible domain connecting the DBD with the LBD. The ligand-binding domain is moderately conserved in sequence but highly conserved in structure. The main transcriptional role of the LBD is to mediate interactions with co-

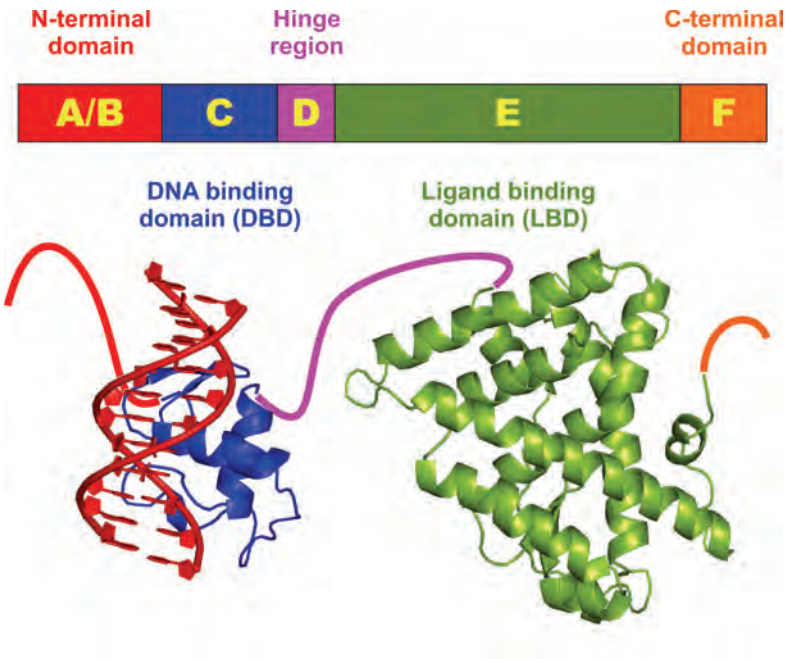


Figure 1.3 Domain structure of nuclear receptors.

activator or co-repressor proteins in a ligand-dependent manner. In addition the LDB plays role in several nuclear-receptor functions such as homodimerization or heterodimerization (Mangelsdorf et al., 1995; Moras & Gronemeyer, 1998), subcellular localization, and ligand-dependent transcriptional activation/repression.

The A/B region is the most variable both in size and sequence and in many cases contains an AF-1 domain. The A/B domain shows promoter- and cell-specific activity, suggesting that it is likely to contribute to the specificity of action among receptor isoforms and that it could interact with cell type-specific factors. On the other hand, the modulatory domain is the target for phosphorylation mediated by different signaling pathways, and this modification can significantly affect transcriptional activity (Shao and Lazar 1999). There are several reports indicating that RARs and other receptors can be phosphorylated by cyclin-dependent kinases and that this phosphorylation is important for ligand-dependent and -independent transactivation (Rochette-Egly et al 1992; Rochette-Egly et al 1997; Taneja R 1997).

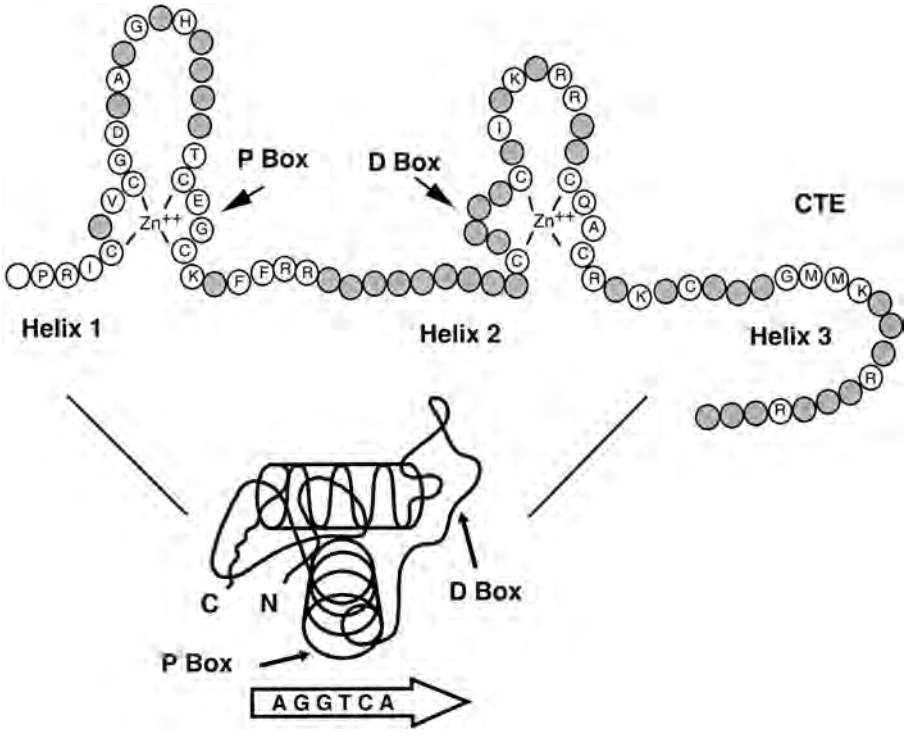


Figure 1.4 The DNA binding domain of nuclear receptors. (Aranda and Pascual, 2001)

The DBD is the most conserved domain of the nuclear receptor (Figure 1.4). It has the ability to recognize specific target sequences and activate genes. The DBD contains nine cysteines, as well as other residues that are conserved across the nuclear receptor superfamily and provides the high-affinity DNA binding. The DBD domain contains also two zinc-binding motifs that span approximately 60-70 amino acids (Gronemeyer & Moras, 1995, Arada & Pascual, 2001). In each zinc-binding motif for cysteines form a tetrahedron around a single zinc ion. Amino acids required for discrimination of core DNA recognition motif are present at the base of the first finger in a region known as “P box”, and in another region known as “D box” which is situated near the base of the second zinc finger. The core DBD contains two α-helices. The first one begins at the third conserved cysteine residue and is known as recognition helix. It binds the major groove of the DNA making contact with specific bases. The second alpha helix spans the COOH terminus of the second zinc-binding motif and forms the right angle with the recognition helix (Lee et al 1993). Many structures of DBDs are done by X-ray crystallography: GR (Luisi et al., 1991), ERα (Schwabe et al., 1993), RXRα/TRβ (Rastinejad et al., 1995), RXR/RAR (Rastinejad et al., 2000).

The D domain or the hinge region is not well conserved among the different receptors and allows rotation of the DBD.

The first structure of the LBD (homodimer of RXR) was done in by X-ray crystallography (Bourguet et al. 1995), followed by many others LBD structures with and without ligands: RARγ (Renaud et al, 1995; Klaholz et al., 1998), TRα (Wagner et al., 1995), ERα (Brzozowski et al., 1997), ERβ (Pike et al, 1999), PR (Williams & Singler, 1998), VDR (Rochel et al., 2000). The crtystallographic studies demonstrate that overall structures of different receptors are quite similar. The LBDs are formed by 12 conserved α-helical regions canonically numbered from H1 to H12. A β turn is

situated between H5 and H6. The LBDS is typically folded into tree-layered antiparallel helical sandwich (Figure 1.5). A central core layer containing three helices is packed between two outer layers and the overall structure has a cavity, the ligand-binding pocket, which accommodates the ligand. The ligand binding pocket is mainly hydrophobic.

However, there are some variations in the canonical structure of the LBD – the PPARγ, RARγ and VDR, for example. The PPARγ is contains an extra α-helix named H2'. No helix H2 was found in RARγ. VDR receptor contains a poorly structured insertion between helices H1 and H3 for which no functional role had been defined. Additionally the size of the ligand-binding pocket varies among the different receptor, being for example very large for PPARγ, which allows binding of very different ligands.A comparison of apo- and holo-LBD structures reveals a common mechanism by which the ligand dependant activation occurs (Renaud et al, 1995;

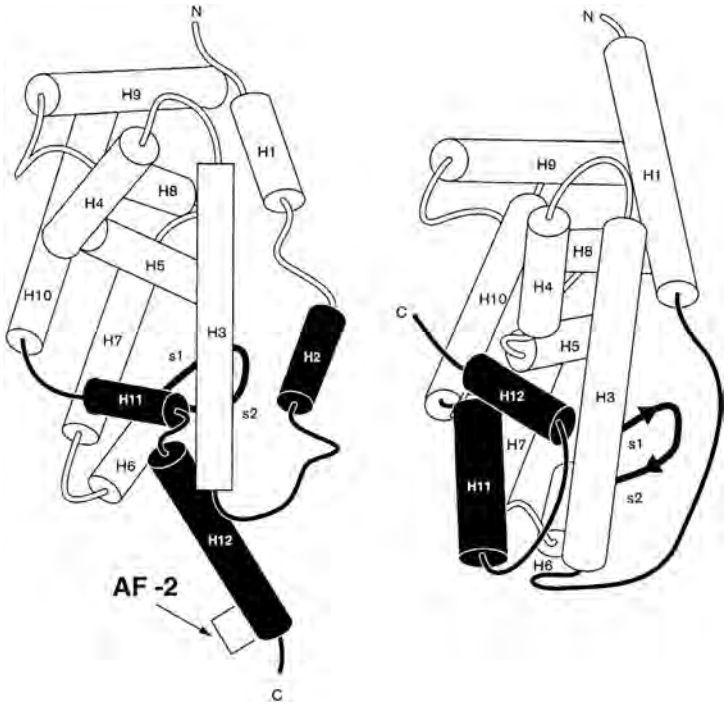


Figure 1.5 Schematic drawing of the nuclear receptor ligand binding domain. On the left the crystal structure of RXR LBD is shown. On the right, the ligand-bound LBD of RARγ is shown. (Aranda & Pascual, 2001)

Arada & Pascual, 2001, Egéa et al. 2000). Upon the ligand binding the helix H11 is repositioned in the continuity of the helix H10. The helix H12 rotates and seals the ligand-binding cavity and further stabilizes the ligand binding by contributing additional ligand-protein interactions. Helix H12 is a crucial component of the NR LBDs, because its ligand-induced re-positioning in the holo NR contributes in a critical manner to the surface recognized by the coactivators and thereby generates a transcriptional active AF-2 domain.

1.1.4 Hormone response elements

Nuclear receptors regulate transcription by binding to specific DNA sequences known as hormone response elements (Evans, 1988). The response element is part of the promoter of a gene. These elements are located in regulatory sequences normally present in the 5'-flanking region of the target gene. The analysis of a large number of naturally occurring as well as synthetic HREs revealed that a sequence of 6 base pairs constitutes the core recognition motif. Although some monomeric receptors can bind to a single hexameric motif, most receptors bind as homo- or heterodimers to HREs composed typically of two core hexameric motifs. For dimeric HREs, the half-sites can be configured as palindromes, inverted palindromes, or direct repeats (Glass, 1994; Mangelsdorf et al., 1995; Mangelsdorf & Evans, 1995). Structures of DBDs homo- and hetero-dimers show different arrangements of DBDs on a DNA. On palindromic sequences the DBDs are positioned “back-to-back”, while on direct repeat sequences the orientation is “head-to-tail”. That is phenomena of the so-called “polarity”.

For a hexameric sequence two consensus motifs have been identified: the sequence AGAACA is preferentially recognized by steroid receptors, whereas AGG/TTCA serves as recognition motif for the remaining receptors of the

superfamily. Steroid hormone receptors typically bind to palindromes of the AGAACA sequence separated by three nucleotides, with the exception of the ERs that recognize the consensus AGGTCA motif with the same configuration. In contrast to steroid receptors that almost exclusively recognize palindromic elements, nonsteroidal receptors can bind to HREs with different configurations. Some of these response elements are capable of mediating transcriptional responses to more than one ligand. This is the case of the palindromic element AGGTCATGACCT that confers regulation by both thyroid hormones and retinoic acid. Analysis of natural and synthetic HREs has shown that the most potent HREs for nonsteroid receptors are configured as DRs. Analysis of variably spaced DRs suggested that the length of the spacer region was an important determinant of the specificity of hormonal responses. Thus DRs separated by 3, 4, and 5 bp (i.e., DR3, DR4, and DR5) mediate preferential regulation by vitamin D, thyroid hormone, and retinoic acid, respectively (Naar et al., 1991, Umesono et al. 1991). The subsequent demonstration that DR1 serves as the preferred HRE for the RXR or for the PPAR and that RARs can also activate transcription through a DR2, expanded the model from a 3-to-5 rule to a 1-to-5 rule (Mangelsdorf & Evans, 1995). Furthermore, a DR0 sequence can also act as a receptor binding site, and widely spaced DRs can act as promiscuous response elements for different nonsteroid receptors and even for ERs (Kato et al., 1995).

1.1.5 Agonism, antagonism and its structural basis

Nuclear receptor ligands may display dramatically diverse effects ranging in a spectrum from agonism via antagonism to inverse agonism. The stimulation of gene expression by the ligand is referred to as an agonist response (such as the hormones estradiol and testosterone). The agonistic effects of endogenous hormones can also be mimicked by certain synthetic ligands, for example, the glucocorticoid receptor

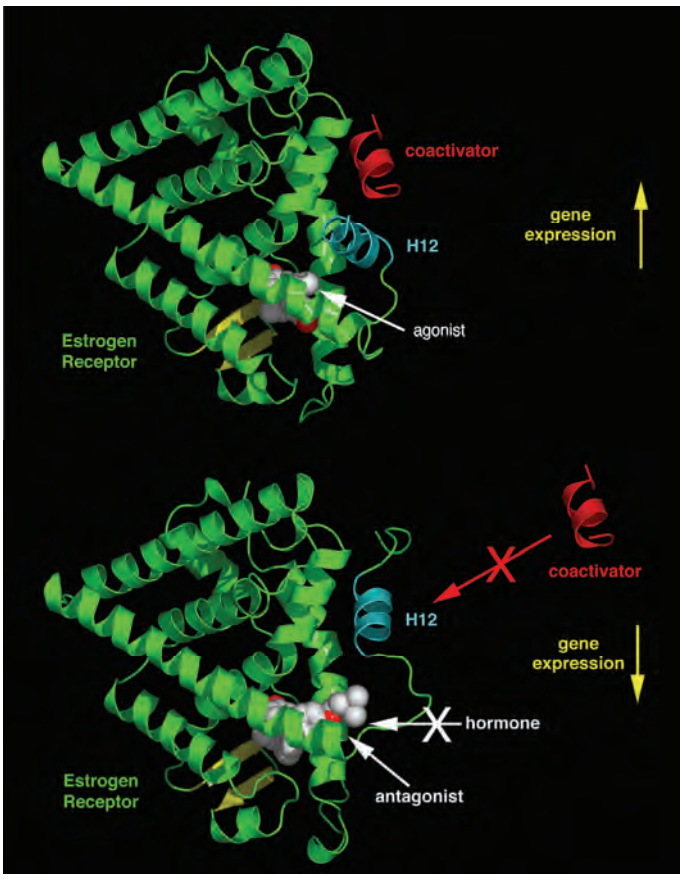


Figure 1.6 Stuctural basis for the mechanism of nuclear receptor agonist and antagonist action. (Brzozowski et al., 1997)

anti-inflammatory drug dexamethasone. Agonist ligands work by inducing a conformation of the receptor which favors coactivator binding. Other synthetic nuclear receptor ligands have no apparent effect on gene transcription in the absence of endogenous ligand. However they block the effect of agonist through competitive binding to the same binding site in the nuclear receptor. These ligands are referred to as antagonists. An example of antagonistic nuclear receptor drug is mifepristone which binds to the glucocorticoid and progesterone receptors and therefore blocks the activity of the endogenous hormones cortisol and progesterone respectively. Antagonist ligands work by inducing a conformation of the receptor which prevents coactivator and promotes corepressor binding. Finally, some nuclear receptors promote a low level of gene transcription in the absence of agonists (also referred to as basal or constitutive activity) (Busch et al., 2004). Synthetic ligands

which reduce this base level of activity in nuclear receptors are known as inverse agonists.

When an agonist is bound to a nuclear receptor, the C-terminal alpha helix H12 of the LDB is positioned to allow the binding of a coactivator protein to the surface of the LBD. Antagonists occupy the same ligand binding cavity of the nuclear receptor. However antagonist ligands in addition have a side chain extension which sterically displaces H12 to occupy roughly the same position in space as coactivators bind. Hence coactivator binding to the LBD is blocked (Figure 1.6). (Shiau et al., 1998; Brzozowski et al., 1997)

1.2 Cryo-electron microscopy and 3D reconstruction

The knowledge of molecular interactions is the key for understanding of the cellular function. Molecules involved in dynamic association compose complexes that are often both too large and too flexible for structural studies by X-ray crystallography. Moreover, a prerequisite for such structure determination is the availability of good crystals. The large molecular complexes often hard to crystallise, or they may form crystals only after flexible parts have been removed. Conversely, successful crystallisation may interfere with study of macromolecular interactions because crystal packing may favour particulars models of interaction while disfavouring or entirely prohibiting others. NMR spectroscopy allows the study of molecular structure and dynamics in solution but can currently be applied only to systems with relatively small mass.

Cryo-electron microscopy was developed over the last three decades and typically is used in structural analysis of single molecules or of the macromolecular assemblies (single-particle analysis) with typical molecular masses ranging from several hundred kilodaltons (Akiba et al., 1996; Kato et al. 2009) to megadaltons

(Chiu, 1993; Jiang et al. 2008, Yu et al. 2008). This technique allows observing biomolecules in their natural environment, being frozen in solution so quickly, that water does not form crystals but has an amorphous state similar to glass (Dubochet et al. 1982). Molecules will be fixed in vitrified ice in an arbitrary orientation. Using images of many molecules in all possible orientation one can restore three-dimensional structure of the molecule.

1.2.1 Electron optics and transmission electron microscopy

The electron microscope and the light microscope have very similar constructions. Both systems require small powerful light source, have condenser lenses to make a parallel beam, have objective lenses, the main part of each microscope, which create a magnified image on the detector plane. The main

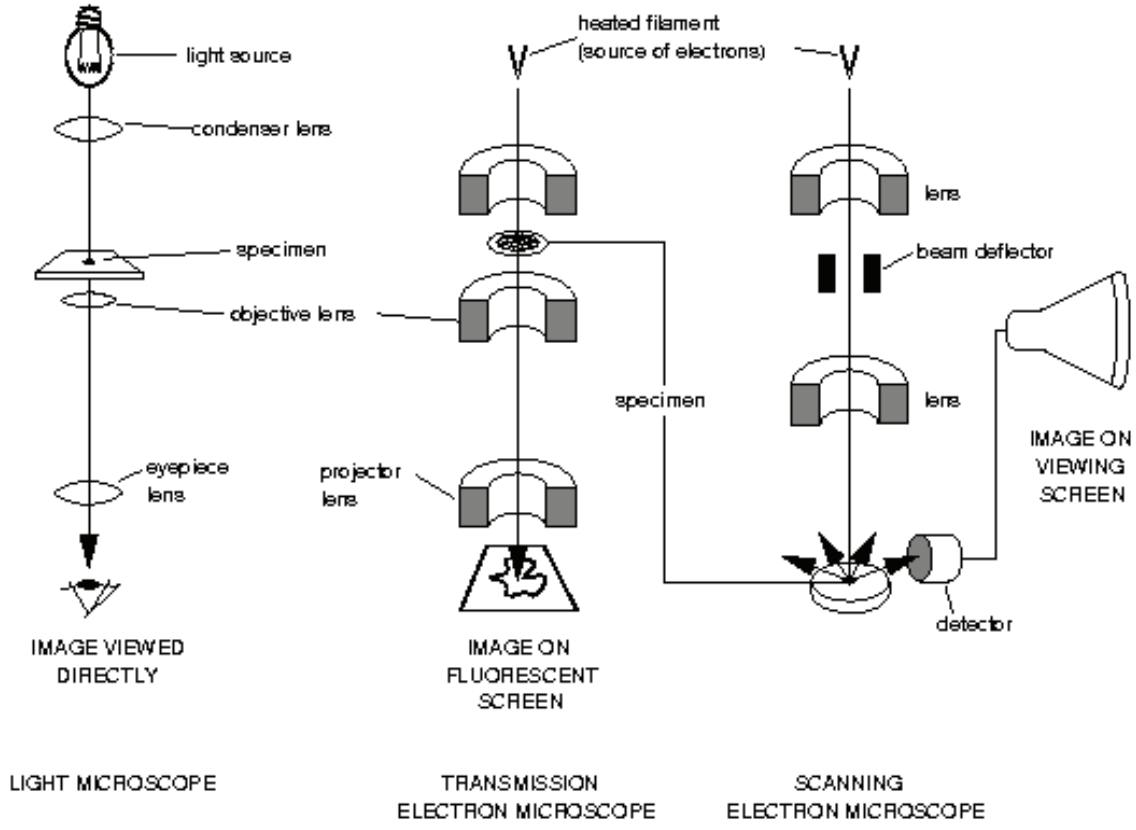


Figure 1.7 Light microscope, transmission electron microscope and scanning electron microscope.

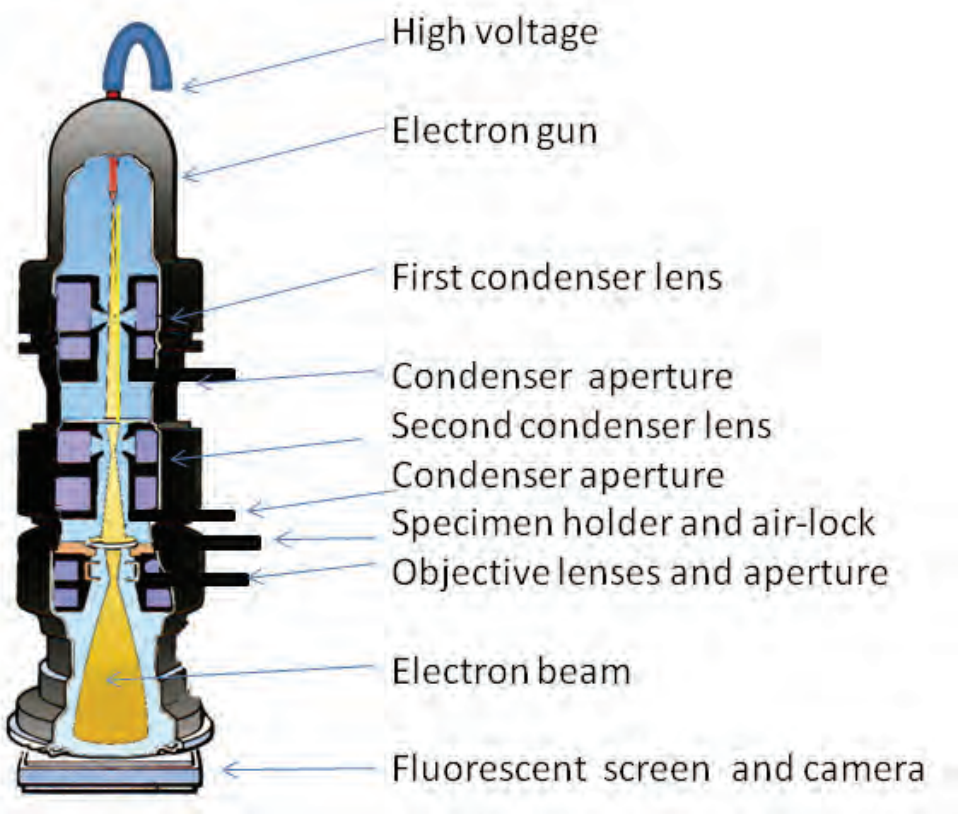


Figure 1.8 Basic scheme of a transmission electron microscope (TEM).

difference of EM microscope and light microscope are the working wavelength and mechanisms of image formation. Essential components of a transmission electron microscope are: the electron gun, condenser lens system, optical enlarging system and the detection system (Figure 1.7 and Figure 1.8).

Two other sources of electrons, that are brighter, have become popular in recent years: the lanthanum hexaboride (LaB6) gun and the field emission gun (FEG). In the former, a lanthanum hexaboride crystal when heated gives off up to 10x more electrons than the tungsten filament heated to the same temperature. Electrons in the gun are accelerated across a potential difference of the order of 100/300 kilo-volts between the cathode (at high negative potential) and anode (at ground potential). Beam shaping and control of emission is affected by properties of the shield and acceleration by the anode.

The function of the condenser lens system is to focus the electron beam emerging from the electron gun onto the specimen to permit optimal illuminating conditions for visualizing and recording the image. Nowadays a double lens condenser system is widely used. It adds considerable flexibility to the illuminating system by allowing a wider range of intensities with a given gun adjustment and making it possible to reduce the area of the object which is irradiated. Condenser apertures in each lens limit the amount of electrons striking the specimen (protecting it from excessive irradiation) and limit the number of X-rays generated from electrons hitting parts of the microscope column.

The electron beam passes through the specimen and enters the optical enlarging system. The enlarging system of the electron microscope consists of an objective lens followed by one or more projector lenses. The objective determines resolution and contrast in the image, and all subsequent lenses bring the final image to a convenient magnification for observation and recording.

The function of the objective aperture is to intercept electrons, which have been scattered by the specimen through excessively large angles. The aperture may be positioned either in the gap between the two pole pieces (restricting the field of observation on the object to an area about the size of the aperture) or, more commonly, in the back focal plane of the objective lens located a short distance behind the pole piece gap. In this position the field is not restricted and widely scattered electrons are still prevented from reaching the image plane. Contamination effects are reduced in this position since only scattered electrons strike the periphery of the aperture opening.

The electron optical image is projected onto the fluorescent screen where the kinetic energy of the electrons is transformed into light energy through fluorescence. The fluorescent screen consists of a surface coated with a layer of activated zinc

sulfide crystals. The fluorescent screen is needed to enable the operator to view the image for selection of an appropriate area. The electron image is usually made into a permanent record on a recording device mounted either above or below the fluorescent screen. For image registration photographic emulsion or a CCD camera are used. For a text and review on this subject see, for example, Slayter 1993, Frank 1996, Frank 2006, Hawkes and Valdrè 1990 and Wade 1992.

Radiation damage limits advantages of transmission electron microscopy data regardless of the initial quality of the specimen. Radiation damage rather than the instrumental resolving power is the main limiting factor in obtaining high-resolution images of biological molecules.

Specimen damage is primarily caused by ionisation resulting from the inelastic interaction of electrons with the orbital electrons of the organic material. This, in turn, leads to rearrangement of chemical bonds, formation of free radicals and diffusion of the fragments. Whereas elastic interactions produce low image contrast but no damage, inelastic interactions can produce permanent changes. The most important structural change that occur in biological molecules is molecular fragmentation leading to mass loss as the material liquefies or vaporizes, molecular cross-linking in which molecules join together so that molecular weight rises and the whole specimen eventually becomes one molecule, a three-dimensional network which is amorphous and insoluble. Most biological specimens tolerate an exposure of no more than ~40-80 electrons per $Å^2$, though the highest resolution features of the specimen are already affected at electron exposures as low as ~20 electrons per $Å^2$ (Glaser, 1971).

1.2.2 Negative staining specimen preparation

The amount of scattering, which occurs at any particular specimen point is dependent on its density and overall thickness and is relatively independent of the

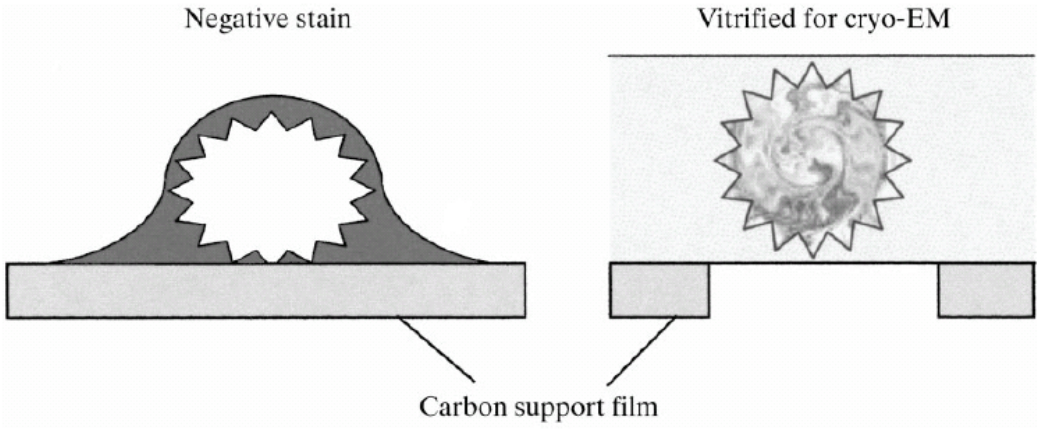


Figure 1.8 Negative stain sample (left panel) and cryo-sample (right panel). Negative stain can flatten particles as it is shown on the figure. (Saibil 2000)

atomic number, chemical composition, or other specimen properties. With increasing density and thickness the probability of scattering increases. Since the level of contrast is determined by the average atomic number of the specimen and biological specimens consist mainly of carbon, nitrogen, oxygen, and hydrogen atoms, weak contrast is a limiting problem in the imaging of biological specimens. Inherent contrast may be increased by the preferential addition of materials of high atomic number during specimen preparation. The main purpose of negative-staining is to surround or embed the biological object in a suitable electron dense material which provides high contrast and good preservation (Figure 1.8).

Hall (1955) was the first to accidentally demonstrate the negative staining effect in a study in which particles were being positively-stained with phosphotungstic acid. Imperfectly washed particles were surrounded and embedded in the dried reagent and, instead of appearing dark on a light background, they were seen light on a dark background. Huxley (1958) independently noticed the same effect with cell sections. Brenner et al. (1959) also observed the same phenomenon with T2 bacteriophage and were the first to call it "negative staining". The negative staining method has been widely used to obtain images of macromolecules with high

contrast. First 3D reconstructions were performed using images of negatively stained ribosomes (Frank et al., 1988; Radermacher, 1988; Radermacher et al., 1987).

Negative stains are chemicals containing heavy atoms Ur or W. An aqueous suspension is mixed with 1 or 2% uranyl acetate and applied to a carbon-coated copper grid. The excess liquid is blotted away and the suspension is allowed to dry. The staining and blotting could be done several times to achieve the more uniform staining over the grid plane. Although, to some extent, the stain goes to the aqueous channels, the structural information in the images basically limited to shape of the molecule as it appears in projection. The method is simple and provides high contrast, but the native structure may be distorted by dehydration. In addition no interior density variations of the molecule are visualised. Some samples can be destroyed by reaction with negative stain. For a review of negative-stain techniques see for example Harris & Horne, 1991 and Hayat, 2000.

1.2.3 Ice-embedded specimen preparation

The modern alternative to the use of dry negative stain as an embedding medium is the preparation of rapidly frozen hydrated biological specimens (Dubochet et al. 1982; Dubochet et al. 1988) (Figure 1.8 and Figure 1.9). Here the sample is applied to a holey carbon film. After removing the excess of the fluid by blotting, a thin layer of solution on an electron microscopic grid is obtained. The grid is then plunged into liquid propane or ethane. Liquid ethane rather than liquid nitrogen is used as the cryogen because liquid ethane has a higher heat capacity. The cryo-fixation in liquid ethane is more reproducible. The aqueous suspension converts to a vitreous phase directly, bypassing the crystalline ice. Thus the frozen grid will have a thin aqueous layer with biological macromolecules embedded in an almost natural hydrated environment. After freezing, grids can be kept for a long time at liquid-

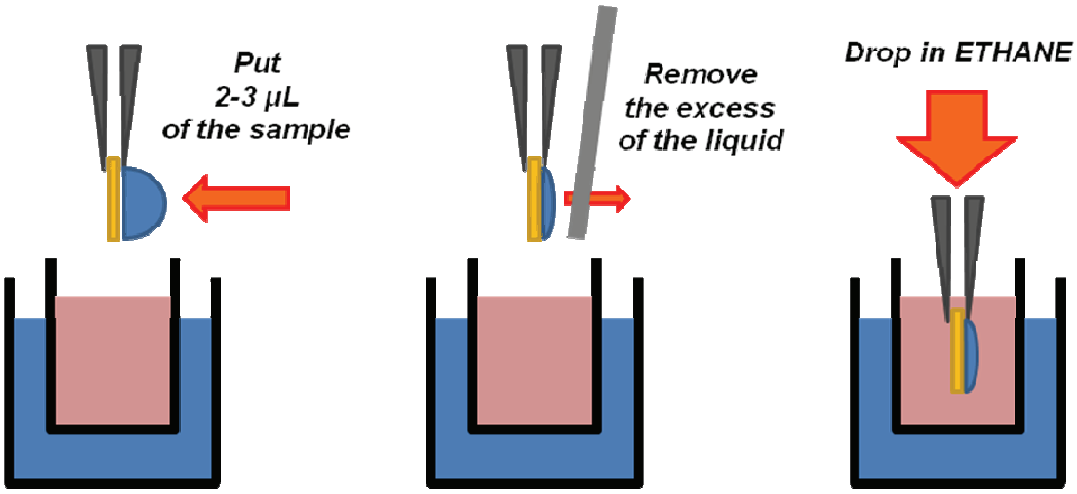


Figure 1.9 Three main steps of cryo sample preparation. Blue container is the container with liquid nitrogen. Pink container is the container with liquid ethane.

nitrogen temperatures. These structurally preserved frozen-hydrated samples are then imaged in an electron microscope at liquid-nitrogen temperatures (Dubochet et al. 1988). These samples have very low amplitude contrast and therefore correction of such images is getting very substantial.

Structural analysis with details smaller than 10Å is possible with the frozen specimens. However, such objects are highly sensitive to electron irradiation. In order to reduce radiation damage, images have to be taken at low illumination doses. In addition, it was found that low temperature reduces radiation damage (Luther et al., 1988). Specimen stability is significantly increased at liquid-nitrogen temperatures and is even better at liquid-helium temperatures (Lamvik, 1991; Stark et al., 1996). Thus, electron microscopes equipped with cryo specimen holders having liquid nitrogen or helium operating at accelerating voltages of 100-300 kV can provide images for high-resolution 3D analysis.

1.2.4 Three-Dimensional Reconstruction framework

The process of 3D reconstruction is the reconstruction of the electron density map from images of individual particles. In general, one doing the 3D reconstruction

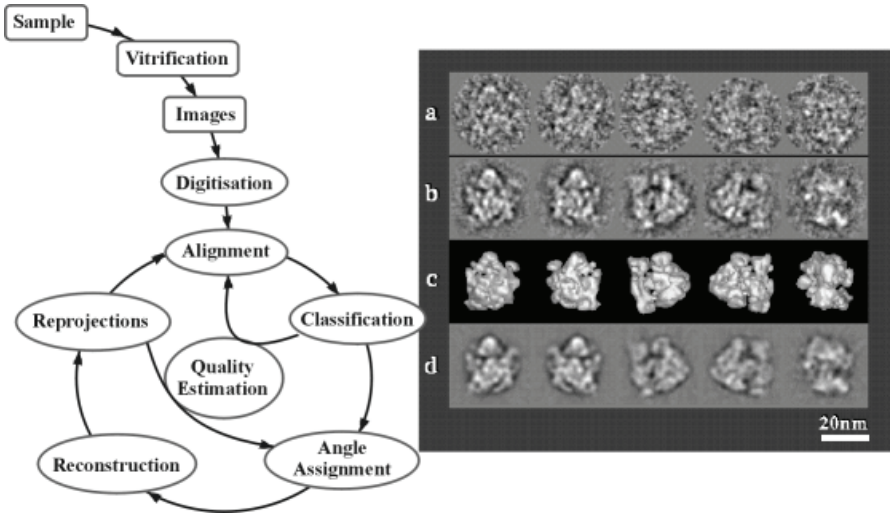


Figure 1.10 The variants of the 3D reconstruction framework (Stark et. al. 2000).

assumes that all individual images belong to the same object and therefore could be interpreted as individual projections from that single object. Prior to 3D reconstruction individual particle images should be extracted from microphotographs and the correction of the transfer function should be done. The main framework of the 3D reconstruction is contains the alignment of individual particle images, Euler angles assignment or relative orientations determination and the computation of tree-dimensional map of electron densities. In a classical approach, all three steps are

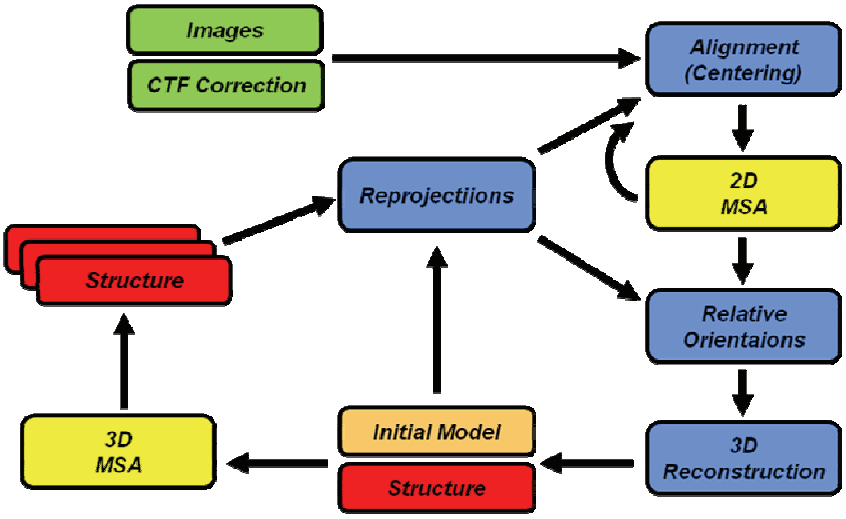


Figure 1.11 Another variant of the scheme shown on Figure 1.10. This variant takes into account the potential heterogeneity of the object. Green color – preparations steps, blue color – main steps, red color – electron density maps, yellow color – 2D and 3D multivariate statistical analises.

done consecutively and the results of every single step are strongly related with the results of the previous one. The resultant 3D density map is usually called “the initial model” and has relatively low resolution. This “initial model” could be further refined using iterative techniques (Figure 1.10 and Figure 1.11) (Ruprecht and Nield, 2001; Serysheva et al., 1995).

1.2.5 Contrast transfer function and CTF correction

The contrast transfer function (CTF) of the electron microscope specifies the relative contrast of features in the image for all spatial frequency ranges and depends both on the spherical aberration and the level of focus of the objective lens, therefore CTF can be used to estimate numerically the quality of EM microphotographs (Erickson and Klug, 1970). By correcting for the CTF it is possible to enhance the signal to noise ratio of images and thus, improving the reliability of alignment and averaging procedures, increase the resolution of the final 3D model

Without correction for the effects of the contrast transfer function the

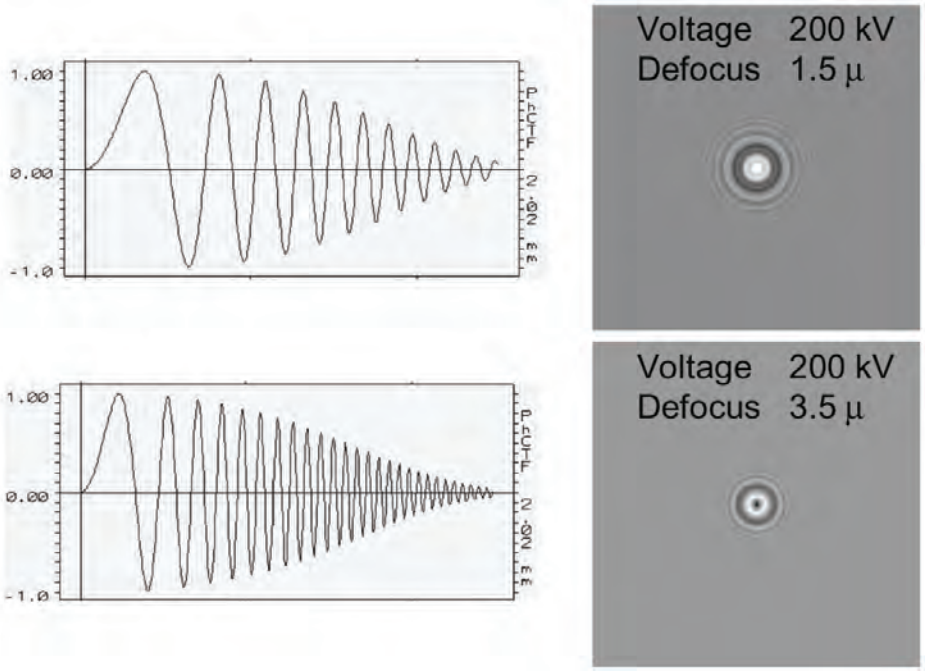


Figure 1.12 Contrast transfer functions (CTF) and point spread functions (PSF).

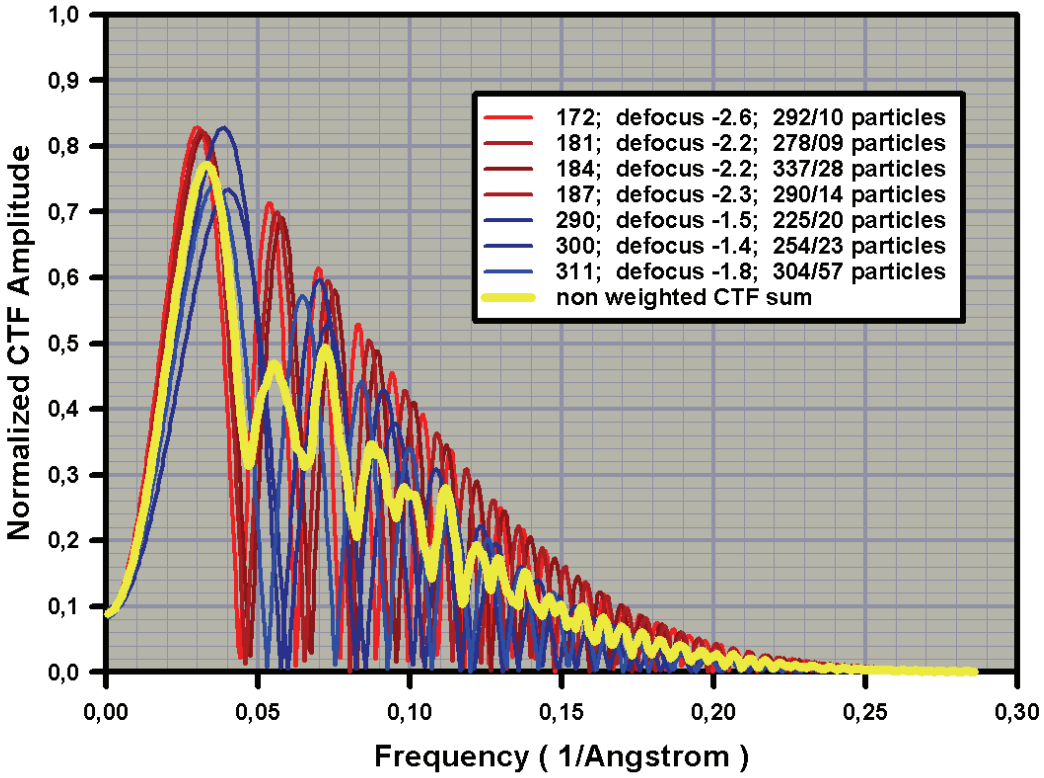


Figure 1.13 The CTF correction principle. After CTF correction all images will have zones with close to zero amplitude of the signal. To compensate the loss of information in certain frequency domain regions it need to have individual images with different defocuses.

reconstruction will have exaggerated features in the size range passed by the CTF spatial frequency band. Most importantly, the definition of the whole particle against the background is affected (Figure 1.12) (Erickson and Klug, 1971; Wade, 1992; Reimer, 1997).

Procedures for CTF can either be applied to the raw data (individual images) or to the 3D volume (Orlova et al., 1997; Stark et al., 1997). Usually the CTF correction is applied to individual images at prior to start the 3D reconstruction. There are some practical observations that the numerical stability of computations on CTF corrected data is higher that on not corrected one.

For each 2D image, the amplitude spectrum is calculated and the resulting 2D spectrum is then high-pass filtered to remove the characteristic background ramps associated with spectra of EM images. This filtering is performed by 2D Fourier

transforming of the spectra and then multiplying the transforms by a rotationally symmetric Gaussian high-pass filter. At the same time, a broad Gaussian low-pass filter is applied to reduce the large amount of noise in the direct spectra. The theoretical CTF is then fitted to the treated spectra of the 2D images. After theoretical CTF determination the phase flipping is done to bring all frequency domains of the image to the same contrast.

The most effective correction is obtained by combining data sets obtained with two or more different defocus settings (Figure 1.13). Practically, CTF should be estimated on tenths or hundreds of images. Many efforts was done to make an automatic defocus and CTF determination and automatic CTF correction that can make the process of data preparation much easier for a scientist (Velázquez-Muriel et al. 2003, Sander et al. 2003; Mallick et al. 2005).

1.2.6 Alignment

Alignment is the operation on two or more images with the aim of bringing a common motif in those images into register. The images are deemed “essentially the same”. That means that all images differ in noise component and perhaps in presence or absence of some details. Therefore the alignment is directly related with the concept of visual likeness and is in its general form the arrangement identical shapes into a common position and orientation. All images are considered “aligned” when they are positioned to minimise a given functional, for example the generalized Euclidian distance or the cross correlation function. This functional could be computed in a number of different ways. One way is to compute the functional directly in between all images in the dataset, the over is to select a single image, the so-called “reference image” and compute the functional in between every single image in the dataset and the reference image (Frank, 1996).

One of the problems that are related with the practical 3D reconstruction is – all images represent different projections from the same object and therefore could be quite different from each other and in general quite non-similar to each other. That introduces the concept of homogeneous and heterogeneous alignments. The first is done on the images of similar kind i.e. on the different instances of the same projection. The latter is done on a full dataset that represent in general case the uniform representation of all projections from the object. The heterogeneous dataset alignment will have the effect that particles within all homogeneous subsets will be aligned as it was defined for homogeneous set of particles, while particles within different subsets will be brought to the orientations that are consistent with each other. For a dataset of the asymmetrical object falling in three main views, front view, side view and top view all particles with the front view will be oriented on the same way, and the same will be true for all particles belonging to side view and top view. The orientation in between any top view and any side view will be fixed. The same will be true for any pair of particles belonging to the different views (Frank, 1996). The heterogeneous alignment is also known as multiple-reference alignment (MRA) (Schats et al., 1995).

Alignment is achieved by using cross-correlation functions (CCFs) (Frank, 1980; Reimer, 1997). The basis of this approach is that the two images to be aligned are translationally shifted relative to each other by known vectors, and in each position, the product of equivalent pixels in the overlapping area are averaged over the area of overlap. This product is the CCF at the position of the shift vector (Frank, 1980.; Frank et al., 1988). The CCF shows a peak at the place where a motif present in both images overlaps. Good alignment is therefore achieved when the CCF reaches a maximum. By searching for peaks in the CCF and shifting molecules by the appropriate vector, translational alignment of the two images can be achieved

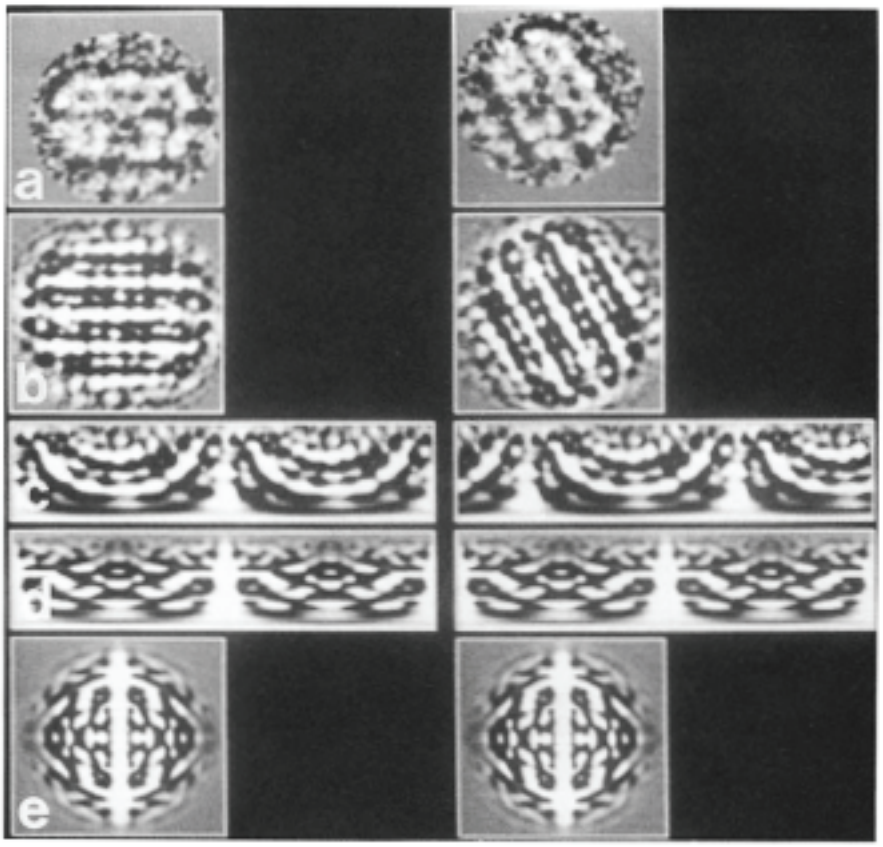


Figure 1.14 Demonstration of the invariance of the double autocorrelation function (DACF) under rotation and shift of an image. **A.** Image of worm hemoglobin in side view and the same image rotated and shifted. **B.** Autocorrelation functions of the images in A. The ACF of the image on the left is seen to be identical to that on the right, but rotated by the same angle as the image. **C.** The ACFs of B in polar coordinates (horizontal: azimuthal angle, from 0 ~ to 360~ vertical: radius). The rotation between the ACFs is reflected by a horizontal shift of the polar coordinate representation. **D.** A second, one-dimensional autocorrelation in horizontal direction produces identical patterns for both images: the DACF in a polar coordinate representation. **E.** The DACF mapped into the Cartesian coordinate system. (Frank, 1996)

(Figure 1.15). The rotational alignment is usually done by using the so-called “translational invariant“ or double autocorrelation function (Figure 1.14) (Frank, 1996).

The practical approach to the alignment in general includes two steps Reference-free alignment (Dube, 1993). Is the first alignment step in the image processing, and is used to centre the particles. The single particle data set is compared to a rotationally averaged total sum of the band-pass filtered particles, resulting in a translational alignment that centres the images. After that some

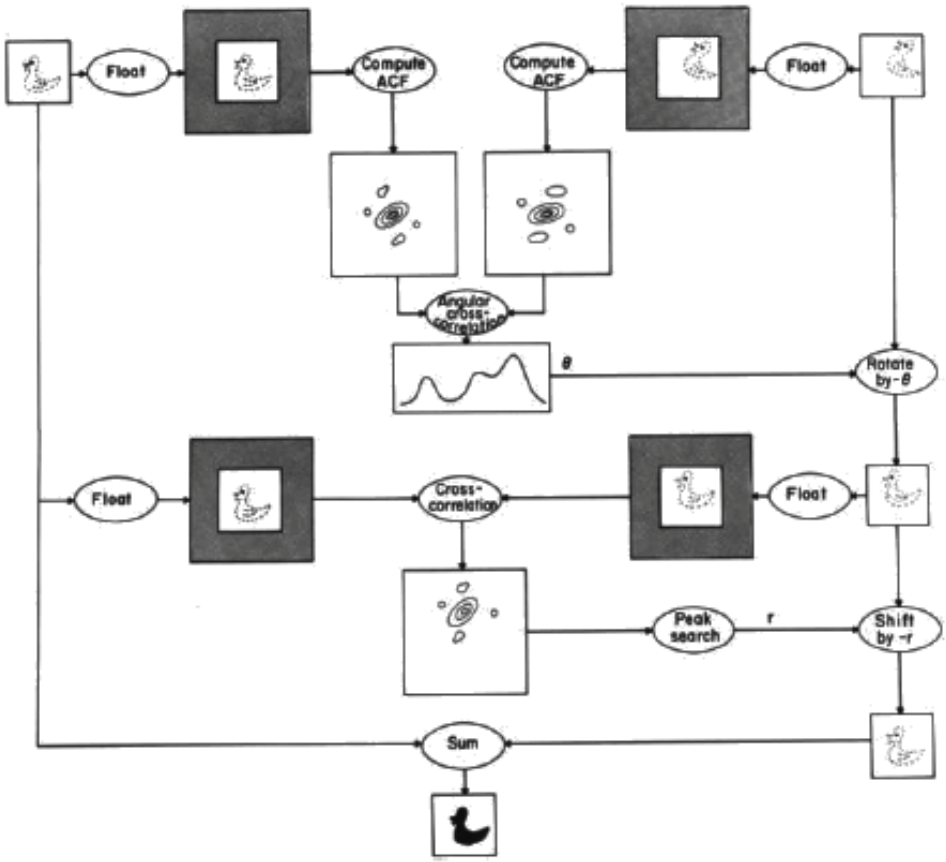


Figure 1.15 Scheme of two-step alignment utilizing the translation invariance of the autocorrelation function. Both the reference and the image to be aligned (images on the top left and top right, respectively) are "floated," or padded, into a larger field to avoid wrap-around artifacts. The ACFs are then computed and rotationally cross-correlated. The location of the peak in the rotational CCF establishes the angle between the ACFs. From this it is inferred that the image on the top right has to be rotated to bring it in the same orientation as the reference. (Note, however, that this conclusion is only correct if the motif in the images to be aligned is centrosymmetric. Otherwise, the angle between the ACFs is compatible with both the angles 0 and 180 - both positions have to be tried in the following cross-correlation. Next, the correctly rotated, padded image is translationally cross-correlated with the reference, yielding the CCF. The position of the peak in the CCF is found by a two-dimensional peak search. Finally, the rotated version of the original image is shifted to achieve alignment. (Frank, 1996)

references could be selected and used to improve the alignment of the entire dataset (Schats et al., 1995). It can be a single reference for homogeneous alignment or several references to do directly multiple-references or heterogeneous alignment.

1.2.7 Angular orientation of individual particles

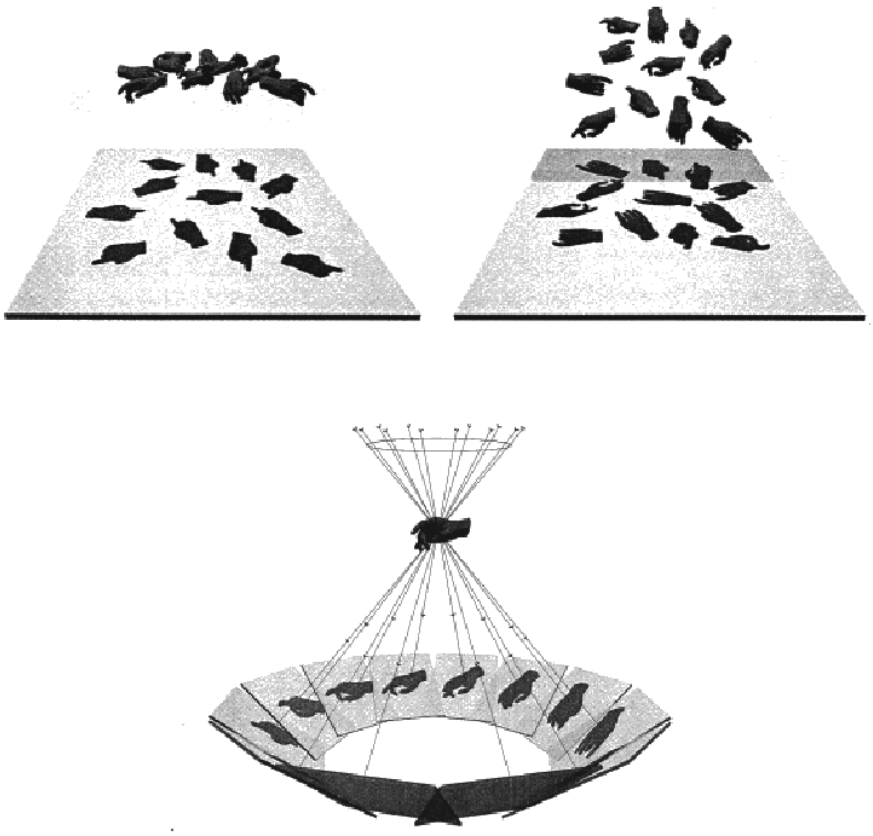


Figure 1.16 Principles of random conical data collection. First row – untilted and tilted images with molecule in preferable orientation. Bottom – equivalent projection geometry. (Frank, 1996)

As long as all individual images are aligned one needs to determine the relative orientations of projections or projections directions. That means that it needs to know the angles in between all possible pairs of projections.

The way to collect data and simultaneously additional information about projections orientations developed in 1987 by Radermacher and co-workers and is based on the random conical tilting data collection geometry (Frank, 1996) (Figure 1.16). The only constraint imposed by this technique is that the macromolecules must tend to lay onto the grid showing one or a few preferent views. Although this requirement may seem restrictive, it is fairly usual for biological molecules to interact with the film support producing a collection of preferential views. Any subset of molecules showing identical views in an untilted specimen form a rotation series with

random azimuth. When the specimen grid is tilted by a fixed angle, the above subset will appear in the micrograph as a conical projection series with random azimuths and cone angle. In the actual experiment, the specimen field is recorded twice: once tilted and once untilted (in this specific order). The first micrograph is used to extract projections for the reconstruction. The purpose of the second micrograph is twofold: to separate the particles according to their views and within each subset, determine the relative azimuths of all particles.

In the 3D projection matching scheme (Penczek et al., 1994,;Frank, 2006), reference projections are computed from low-resolution template structure such that they cover the entire angular space evenly (Figure 1.17). As template structure, an

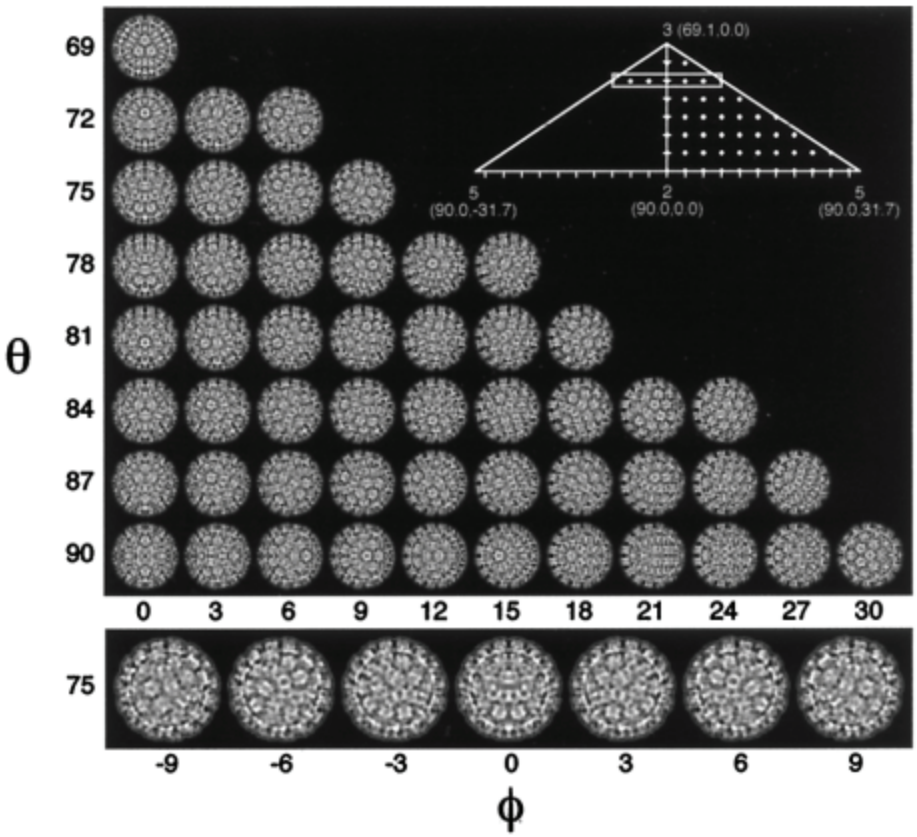


Figure 1.17 Typical set of projections that can be used for determination of Euler angles (the case of icosahedral symmetry). The correspondence of individual particle and a certain reference projection can be established by cross correlation analyses (in case of the aligned particles) or by direct multiple reference alignment. The precision of the angle that will be determined is strongly related with the number of reference projections. (Baker et al. 1999)

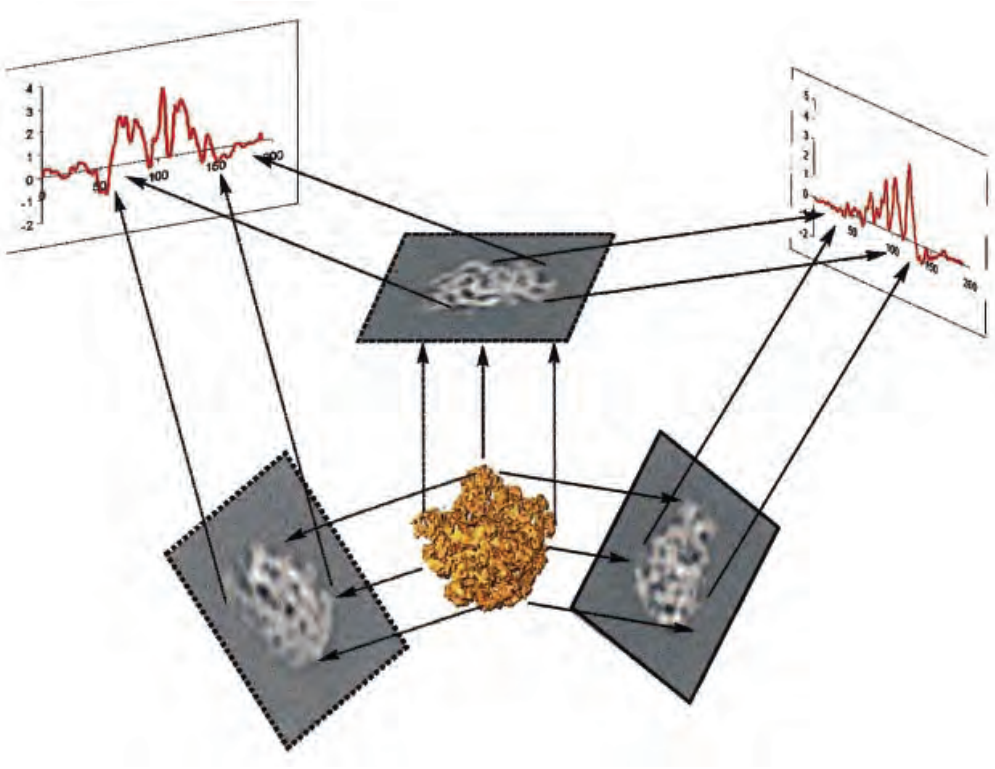


Figure 1.18 The angular reconstitution technique is based on the common line projection theorem stating that two different two-dimensional (2D) projections of the same 3D object always have a onedimensional (1D) line projection in common. From the angles between such common line projections, the relative Euler-angle orientations of set projections can be determined *a posteriori* (van Heel, 1987). For an entirely asymmetric particle like this 50S ribosomal subunit, at least three different projections are required to solve the orientation problem. (van Heel, 2000)

existing 3D volume is used, which might have been obtained by merging several random-conical projection sets. A given experimental projection is cross-correlated with all reference projections. The angle corresponding to the projection giving the largest CCF peak is the desired projection angle and will be assigned to the experimental projection. The cross-correlations function between reference and current projection gives, at the same time, the shift and azimuthal orientation of the particle under consideration. Using the new parameters for each experimental projection, a new reconstruction is computed, which normally has improved resolution. This refined reconstruction can now be used as new template, and so on (Frank, 2006).

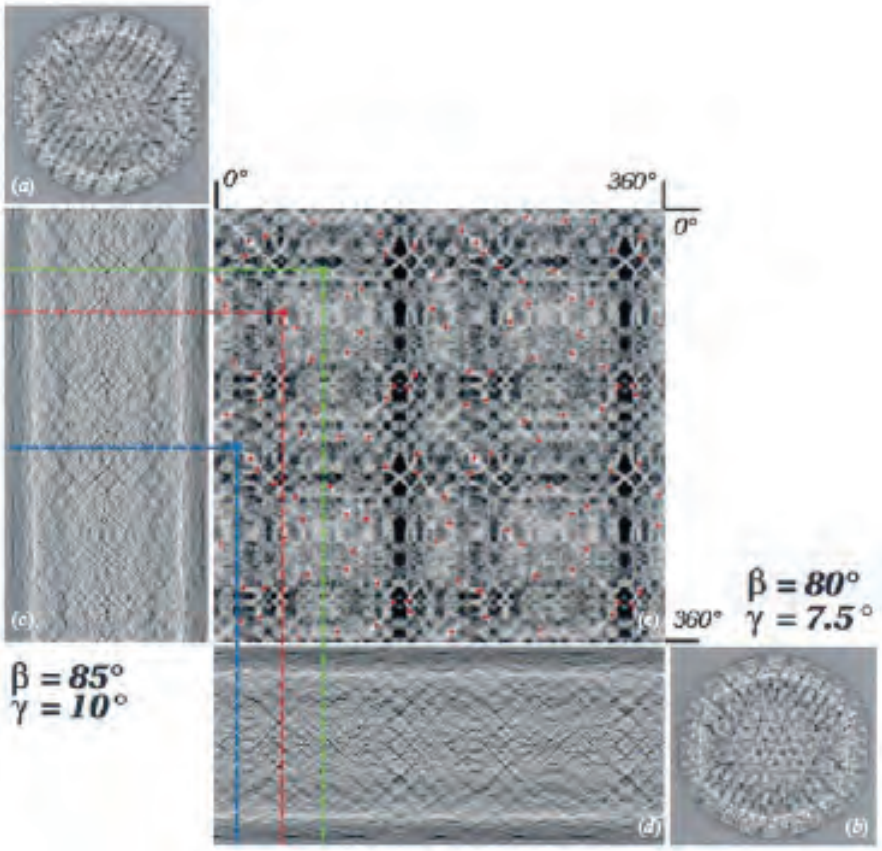


Figure 1.19 Sinograms and sinogram correlation functions. This illustration provides a graphical overview of the relations between a 2D class average (noise-reduced projection images), their `sinograms', and the sinogram correlation function between two sinograms. The images shown here (*a*, *b*) are class averages deduced a large data set of Herpes Simplex Virus Type 1 (HSV1) cryo-EM images. Each line of the sinogram images (*c*, *d*) is generated from the 2D projection image by summing all 1D lines of the 2D images, from top to bottom, after rotation of the image over angles ranging from 0° to 360°. Equivalently, the lines of the sinograms are 1D projections of the 2D images in all possible directions ranging from 0° to 360°. Each point of the sinogram correlation function contains the correlation coefficient of two lines of the two sinograms one is comparing (*e*). (van Heel, 2000)

The interested variant of the projection matching approach is polar Fourier transform (PFT), widely used for 3D reconstruction of objects with icosahedral symmetry (Baker & Cheng, 1996). It suggests the comparison of individual images and projections from model in Fourier space. All model projections and raw images are reinterpolated onto polar coordinates. Then PFT of raw images are correlated with PFT of the model.

The angular reconstitution technique allows one a posteriori find the relative orientations of the 2D projections of 3D structure (van Heel, 1987; Frank, 2006). If we know the orientational relationship between all projection images, we can reconstruct the 3D structure. The technique is based on the common line projection theorem stating that two different 2D projections of the same 3D object always have a one-dimensional line projection in common. From the angle between such common line projections, the relative Euler-angle orientations of projections can be determined (Fig. 12). For an entirely asymmetric particle, at least three different projections are required to solve the orientation problem. The common line projection theorem is the real-space equivalent the Fourier-space common lines theorem. The Fourier-space approach has been formulated into computer programs aimed at solving the structures of icosahedral particles (Crowther, 1971; Crowther et al., 1970).

In practice, the common line search is performed in real space with the help of the so-called sinogram; this is data table that contains in its rows the 1D projection of a 2D image exhaustively computed for all angles. The angle is found by comparing or correlating the sinograms of the two 2D projections (van Heel, 1987; van Heel, 2000) (Figure 13).

1.2.8 3D reconstruction

As long as the relative orientations of particle are know it is possible to compute a 3D densities map. Three main approaches have been developed in EM and medical tomography: algebraic reconstruction methods (mostly used in electron tomography) (Gordon et al., 1970, Jiang and Wang, 2003 and Sezan and Stark, 1982), algorithms based on Fourier transform (for the direct Fourier inversion see Shepp and Logan, 1974 and for the Fourier–Bessel approach see DeRosier and Klug, 1968, Crowther and Amos, 1971 and Crowther, 1971 and filtered back-

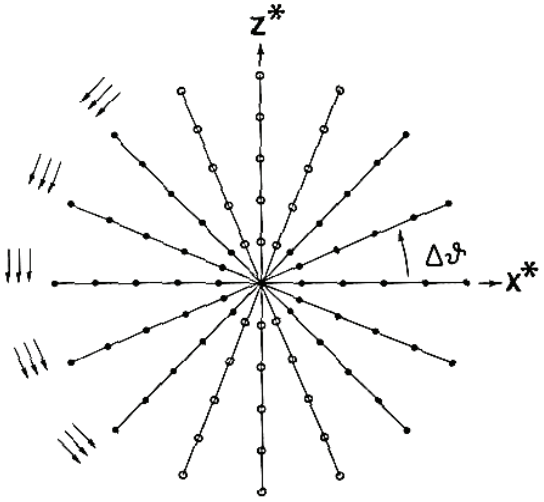


Figure 1.20 Filling of a Fourier space. Density of sampling points in Fourier space obtained by projections decreases with increasing spatial frequency (single-axis tilting case). (Frank and Radermacher, 1986)

projection (Cormack, 1973, Frank, 1996 and Harauz and van Heel, 1986, for review, see Herman, 1980 and Frank, 1996).

Fourier approaches to reconstruction utilize the projection theorem directly and regard the Fourier components of the projections as samples of the 3D transform to be determined. In most case, the positions of these samples do not coincide with the regular three-dimensional Fourier grid. This situation leads to a complicated interpolation problem, which can be stated as follows: given a number of

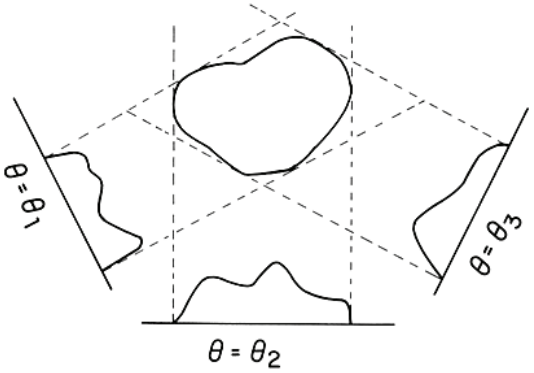


Figure 1.21 The backprojection method of a 3D reconstruction. The density distribution of projection is “smeared out” in the original direction of projection, forming a “back-projection body”. Summation of this back-projection bodies generated for all projections yields the approximation of the object. (Frank, 1995)

measurements in Fourier space at arbitrary points not lying on the sampling grid (Figure 1.20).

Numerous results in EM structural analysis have shown that methods based on the Back-Projection (BP) approach have broad usage due to their robustness, relative ease of parallelization and the possibility to process unlimited numbers of projections (Herman and Lent, 1976, Herman, 1980 and Frank, 1996). For particles with icosahedral symmetry (e.g., viruses), the Fourier–Bessel (FB) approach is used more often (Crowther, 1971 and Crowther and Amos, 1971). The algorithmic implementation of the BP method is based upon stretching a projection or an “image” along the projection direction within a three-dimensional matrix of object densities (Figure 1.21). The geometry of the stretching operation in 3D space depends on the angular divergence of the illumination source used for data acquisition and has an influence on the final 3D density precision and resolution. In the case of X-ray tomography, the geometry will be conical (Herman, 1980) or cylindrical (Egger and Morel, 1998).

1.2.9 Iterative refinement

As soon as the first reconstruction is obtained, a refinement can be performed as an iterative procedure. It has been proposed several different variant of the refinement procedure based on the common principle – the comparison of the projections of the obtained 3D density map with the existing images of the individual particles: with the increase of the resolution the similarity of projections and individual images is increased. Projections of the first reconstruction are well centered and hence can be used as new references in the procedure of raw data alignment. More accurate alignment improves classification of images. Consequently, the orientations of the `better classes' can be determined more accurately, resulting in an improved

3D structure (Serysheva et. al., 1995; Stark et al. 2000) (Figure 1.10 and Figure 1.11).

1.2.10 Evaluation of the 3D reconstruction

Single-particle reconstruction differs from crystallography in that the resolution attainable in the result must be assessed from the result itself. Examination of the power spectrum of the image reveals the resolution to which the modulation of the transfer function can be observed. This provides an upper limit to the information, which can be extracted from the image. However, there is no guarantee that the reconstruction will reach this resolution, since reconstruction is also affected by such factors as the accuracy of orientation determination and the variation between

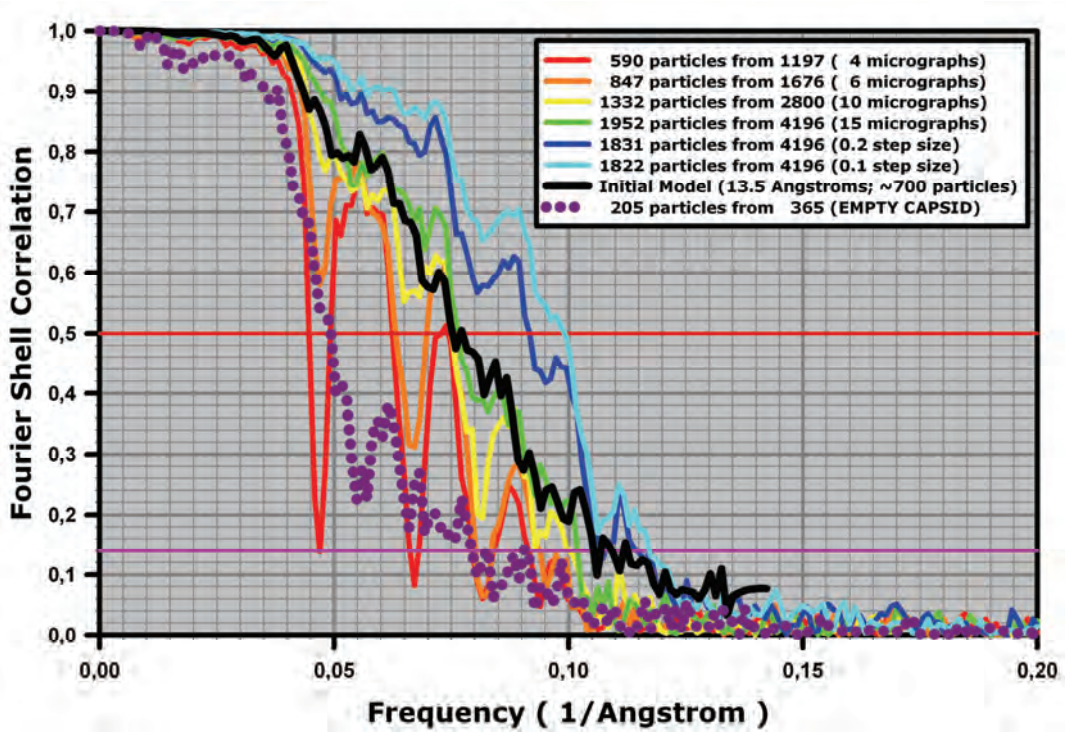


Figure 1.22 Fourier Shell Correlation – 3D reconstruction of the Rice Dwarf Virus. Red, orange, yellow and green curves – illustrate the resolution dependency from number of particles used. Distinctive gaps is a result of an incomplete CTF correction. The more particles with different defocuses is used the smaller are all gaps. Blue and cyan curves – subsequent refinements of a full dataset. The resolution of the structure is approximately 10Å by 0.5 criteria.

particles. Variation is particularly important for large complexes and may correspond to a more significant limit to the resolution achieved with some specimens than the quality of the microscope or the processing.

The most common way of resolution estimation involves dividing the data into two halves and then calculating two independent reconstructions, which are then compared by Fourier shell correlation (FSC) (van Heel and Haraus, 1986) or difference phase residual (DPR) (Baker et al., 1999). The resolution to which these measures satisfy a quality limit is taken as the resolution of the reconstruction. There is significant discussion in the literature over the choice of the appropriate quality limit (Gabashvili et al., 2000; Mueller et al.,2000). Grigorieff (2000) has pointed out that these measures are often applied in a way that gives an artificially optimistic resolution estimate. Nevertheless, this approach remains the most common in the field and represents the simplest way to compare the quality of reconstructions performed by different groups. Often, an FSC limit of 0.5 is used as a conservative measure of resolution. In well behaved reconstructions, the precise choice of cutoff value is relatively unimportant since the FSC falls sharply near the resolution limit. A second less general method is the use of internal criteria within the image processing itself. In the case of icosahedral reconstruction, for example, this could be the average phase residual for the individual images relative to the average (Stewart et al., 1991). This criterion is usually extremely conservative since the higher resolution information is relatively noisy and no increase in signal to noise similar to that obtained in crystallographic averaging occurs for the individual images.

1.2.11 Multivariate statistical analyses

The multivariate statistical analysis (MSA) technique of correspondence analysis is next used to identify the principal components of variation in the set of

images, as a preliminary to placing the images into groups of similar molecular images in similar rotational orientations (Frank and van Heel, 1982; Frank,1984 and Frank,1990; Lebart et al., 1984; van Heel and Frank, 1986).

Each image m by n pixels can be considered as a point in (m*n)-dimensional space, where each axis represents the density value at a single pixel. The entire set of images forms a cloud within this space. The distances between points in this space may be interpreted in terms of the similarity of the corresponding images. Distances between points in the (m*n)-dimensional space are measured using metrics such as the Euclidian metric (Borland and van Heel, 1990). The Euclidian distances between any two rows or columns of the input matrix are calculated, producing a symmetric matrix (van Heel and Frank, 1986).

The direct classification by pixel values is extremely computationally demanding. Instead of that straightforward approach the modern software uses the “data compression” or eigenvectors (also called eigenimages or factors) (van Heel and Frank, 1986, van Heel et al., 1996). Each image is presented as a weighted sum of vector images or eigenimeges, and the classification is accomplished according to the weigh coefficients or eigenvalues.

To determine the major components of variation, a new rotated co-ordinate system is described in which the first axis is the eigenvector representing the greatest interimage variance (associated with the largest eigenvalue), the second axis is the eigenvector representing the largest remaining interimage variance, and so on (van Heel et al., 1992). By only considering the significant eigenvectors (typically the first 5–50 (Sherman et al., 1985)), the images can be pseudo filtrated. This reduces the total amount of data for statistical analysis, and allows shifting the sensitivity of the process toward the desired details. Since the eigenvectors show the variation in the dataset, they can provide useful information not only about different

projections of the uknown structure but also about conformational flexibility or substrate binding.

2. Aims of the project

When this project was started the two domains of the nuclear receptors, ligand-binding domain and DNA-binding domain had been studied separately by X-ray crystallography in different complexes with natural and pharmacological ligands (for example Bourguet 1995; Renaud et al. 1995; Klaholz et al. 2000; Rochel et al. 2000; Egea et al. 2002; Greschik et al. 2002; Billas et al. 2003; Tocchini-Valentini et al. 2004). The difficulties of obtaining crystals of a full functional nuclear-receptor complex with DNA compelled for the usage of methods complementary to X-ray crystallography. Single-particle cryo-electron microscopy allows obtaining structural information about macromolecular complexes and hence to gain insights into the molecular mechanisms of action for hard-to-crystallize objects. The combination of cryo-EM intermediate-resolution data with high-resolution X-ray crystallography data for individual proteins was expected to reveal the architecture of a full nuclear receptor complex with DNA, ligands and possibly co-activators. The focus of this work is the molecular mechanism of interaction within the complex formed by the vitamin D receptor (VDR) and retinoid X receptor (RXR), the DNA (a response element with a direct repeat spaced by 3 nucleotides, DR3), ligands and co-activators.

3. Methods

The 3D reconstruction of the RXR/VDR/DNA complex had been done in two major steps related with the equipment used. First step had been done using a FEI Tecnai F20 200kV field emission gun cryo-electron microscope with sample preparation on home-made holey carbon grids. The plunge freezing had been done on a homemade plunger. The second step had been done on a 300kV FEI Polara cryo-electron microscope on factory-produced QUANTIFOIL holey carbon grids using automated vitrification robot VITROBOT MARK IV.

The general scheme of data collection and 3D reconstruction was one and the same for all objects and includes following steps:

1. The cryo sample preparation.
2. The analysis of the distribution of particles. Approximately 20 CCD images usually were done in different parts of the grid with different ice thickness to analyze the effective distribution and concentration of frozen particles.
3. If the concentration of particle is not acceptable – repeat steps 1 and 2 on a different sample, or with a different concentration.
4. The data collection on the CCD images or on films.
5. The visual analysis of the data collected. Only the best images, without visible drift, astigmatism or other distortions, should be used.
6. Digitization of electron microphotographs. For CCD images – conversion to a file format suitable for image processing.
7. Power spectra computation from digitized/converted images. Selection of data with best power spectra. Analyzes of power spectra usually gives higher

precision and allows determination of small drift, small astigmatism or other distortions.

8. Boxing of particles. Making separate stacks of individual particles for each scanned image or for each CCD.
9. Determination of the defocus for individual stacks and thus of the CTF. CTF-correction for individual stacks.
10. Merging or appending stacks into a single file.
11. The preliminary alignment of a full stack. Using the total sum of all particles as a reference.
12. If the quality of the alignment is not satisfactory – repeat step 11.
13. Multivariate statistical analyzis of pre-aligned particles, assigning 50-15 particles per class.
14. Select best class averages and use them as references for multiple reference alignment.
15. Repeat steps 13 and 14 until the quality of the alignment will be satisfactory. At this moment the MSA should give sharp and crisp class averages.
16. Make C1 start up procedure to determine initial Euler angles of the class averages.
17. Make the 3D reconstruction. Create the “initial model” (the 3D density map).
18. Compute reprojections (forward projections) from the model.
19. Make the MRA using those reprojections as references.
20. Make MSA and compute class averages, assigning 15 – 5 particles per class.
21. Determine Euler angles using direct projections matching or angular reconstitution procedure.
22. Compute the 3D density map.
23. Repeat steps 18 – 22 until the model will be fully refined.

3.1 Purification of the complexes

The purification of all complexes was done by the group of Natacha Rochel and Dino Moras in IGBMC. The major difference between the complexes is the DNA used for complex formation. The following paragraph describes in breath a typical purification and complex formation protocol.

The HsVDRΔ (1-427Δ166-216) was expressed as a hexahistidine fusion protein. HsRXRαΔAB (130-462) was cloned into a pACYC plasmid encoding a non-tagged protein. Ligands (9-cis retinoic acid for RXR and 1α,25-dihydroxyvitamin D3 for VDR) were added in a 2-fold excess to saturate the receptors. The annealed oligonucleotide DNA was added in a 1.2 excess to the dimer. The complex was further purified on a gel-filtration S200 column. The final buffers were Tris 20 mM pH7.5, NaCl 50 mM, KCl 50 mM, $MgCl_2$ 4mM, DTT 5mM.

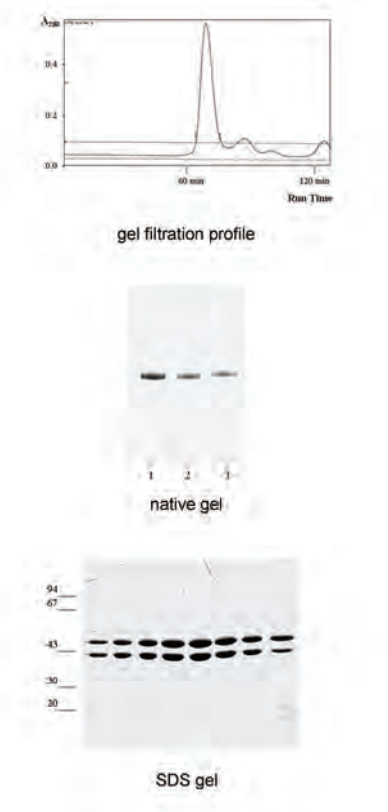


Figure 3.1 Complex purification.

3.2 Initial model

The home-made holey carbon grids were done by a standard procedure of floating the holey film on a copper support, followed by carbon evaporation and after that plastic film elimination by putting grids in a solvent. In order to obtain a thin ice to

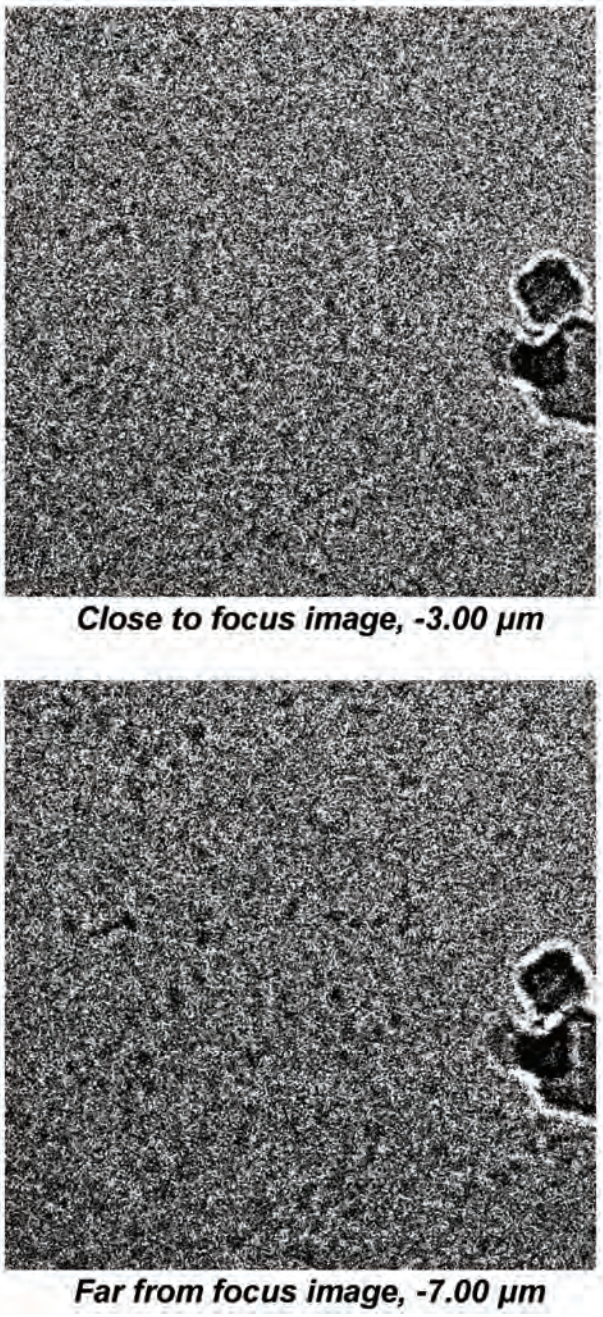


Figure 3.2 VDR/RXR/DNA cryo EM images at different defocus values(200kV acceleration voltage).

increase the transparency an unsually long hand-blotting time around 5 seconds was required.

The images have been collected at liquid-nitrogen temperature with the in-house FEI Tecnai20 field emission gun transmission cryo-electron microscope at 200 kV voltage with a dose of ~20 electrons per $Å^2$ at a magnification of 50,000 on photographic films (Kodak SO-163). In order to determine the appropriate specimen concentration and best distribution of particles in vitreous ice we screened numerous grids and grid areas best suitable for data collection. For this, CCD images were recorded at 50K magnification in different areas in order to identify areas with best visibility of particles and best transparency of ice. All data have been collected around these areas on the previously unexposed zones. For starting the first structure determination, focal pairs of microphotographs were recorded. The first close-to-focus microphotographs were taken at -2.5 - -4.5 μm underfocus. The second microphotographs were taken at -7.0 μm underfocus and were used primarily for particle identification and selection.

Micrographs were digitized with a high-precision drum scanner (Heidelberger Druckmaschinen) resulting in a final step size of 1.0 Å per pixel (scanning step size 4.66μm). The initial boxing was done on far-from-focus images using the BOXER program from the EMAN software package. In total 6,800 particles were boxed. After transfer of the box coordinates to the close-to-focus images that were done by home-maden software the boxed particles from individual microphotographs were saved. The CTF correction of boxed particles was done in CTFIT program from the EMAN software package.

The initial 3D reconstruction was done using the IMAGIC software package. The dataset was centered relative to the total sum of all images. After that the MSA analyses was done to determine the initial references for multiple reference

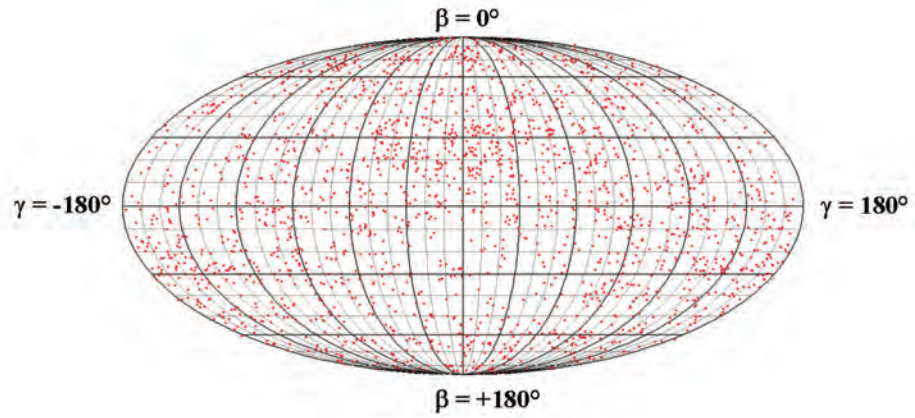


Figure 3.3 Angular distribution of particle views for 200kV reconstruction.

alignment of the dataset. Three best class averages was selected and the dataset was aligned relative to those references. After alignment, another MSA was done used to get the improved class averages. After six iterations the obtained class averages were used to make the initial model of the structure using the so-called "C1-START" procedure in IMAGIC software package. The class averages of different orientations were selected and their Euler angles were determined by common line analysis procedure implemented in IMAGIC. After that several additional class averages was added and their orientations were determined relative to the first three class averages. The 3D density map was calculated and its correctness was checked by comparison of the input class averages and the relative re-projections from the obtained 3D density map. The dataset was iteratively refined up to the resolution ~20Å.

3.3 200kV structure

The dataset for further refinement was collected also with the in-house FEI Tecnai20 field emission gun transmission cryo-electron microscope at 200 kV voltage

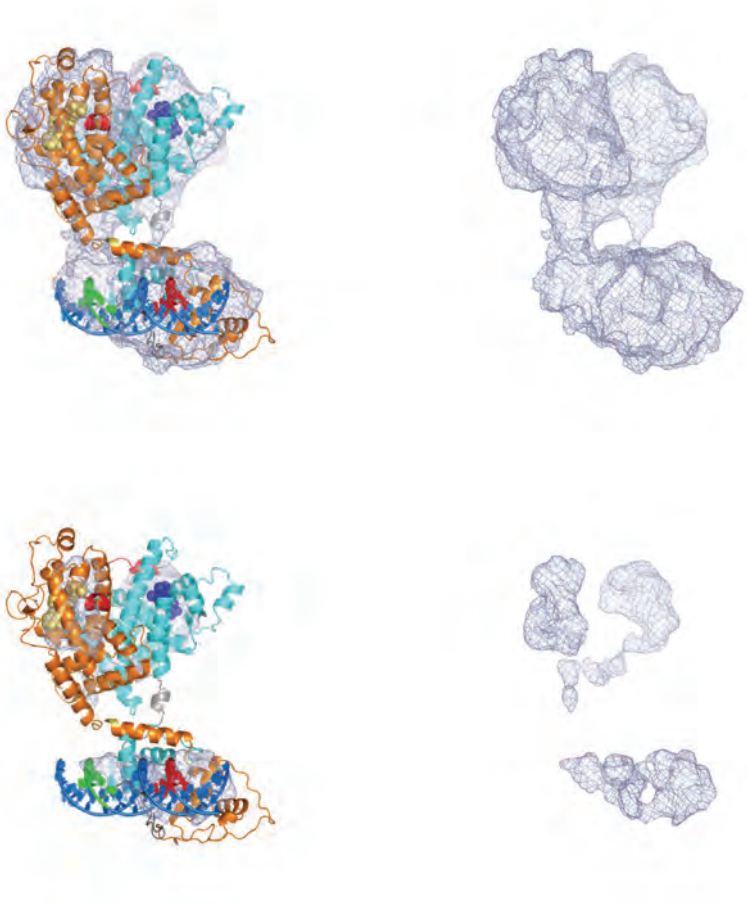


Figure 3.4 The cryo-EM map of the RXR/VDR/DR3 complex displayed at different contour levels (displayed with and without the fitted model) reveals a stronger density for the DNA (because of the phosphorus of the nucleotides) and some residual density for the core part of the LBD domains.

with a dose of ~20 electrons per $Å^2$ at a magnification of 50,000 on photographic films (Kodak SO-163). The data collection was done not using focal pairs at a defocus of -2.0 up to -4.0μm. From approximately 150 microphotographs 40 were selected and digitized as described earlier and 19 micrographs with best power-spectra selected for boxing. To facilitate particle visibility and make the auto-boxing procedure more numerically stable a smooth Gaussian real space filtering has been done, corresponding to a low-pass filter cut-off of 10 Å. The subsequent dataset has been boxed semi-automatically in BOXER with validation of all boxed particles by visual inspection. CTF correction was done by phase flipping using large particle

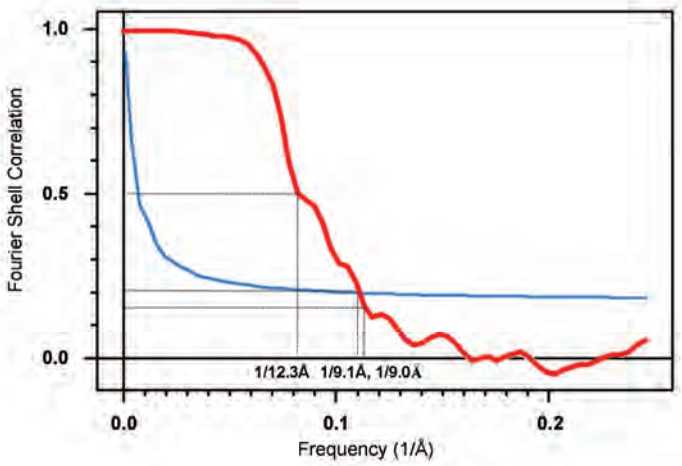


Figure 3.5 200kV structure resolution assessment according to the Fourier Shell correlation.

boxes (512 pixels) using the CTFIT program of the EMAN suite. After particle selection and CTF-correction all boxed particles were coarsened twice resulting in spatial a resolution of 2.0 Å per pixel for further image processing. The ready to process dataset contained 19,938 particles.

The structure was refined using the 3D density map from the previous dataset and the usual procedure alignment-angular reconstitution-3D reconstruction.

3.4 100kV test structure

The 100kV data at have been collected on the in-house FEI Polara which includes a field emission gun tunable between 80 and 300kV. CCD images were recorded at an underfocus of 2.0-3.0 µm and a magnification of 59k on an 4kx4k Eagle CCD camera after careful dark current and gain reference calibration (step size of 1.96 Å per pixel). About 2500 single particle images were selected from 21 CCD images, CTF-corrected (by examining the non-coherently averaged power-spectra calculated from the selected particles) and processed as described above. The structure determination was done by angular reconstitution, as an ab initio

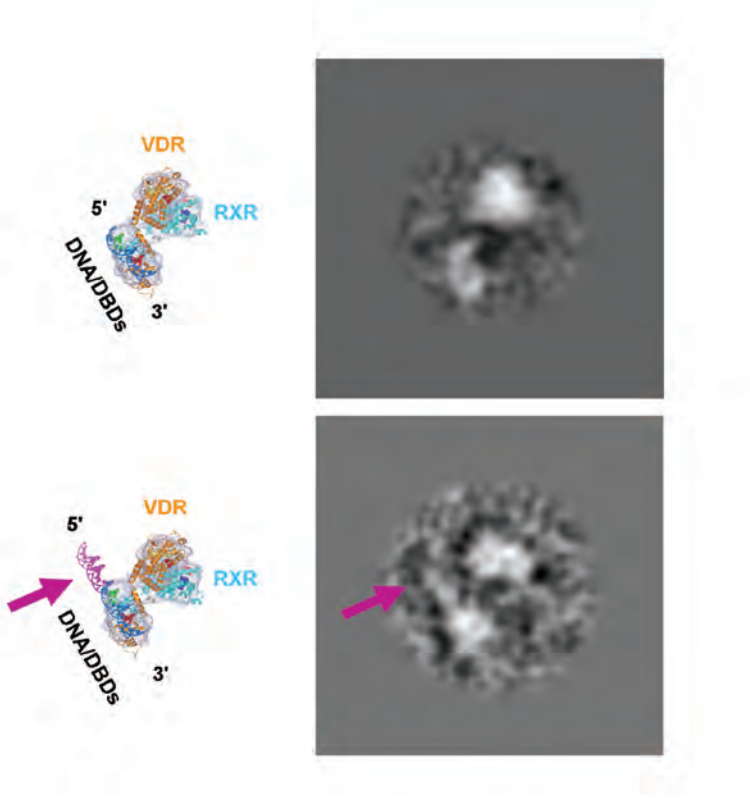


Figure 3.6 Comparison of the class averages of the RXR/VDR complex with the 20bp DNA (top lane) and with the longer DNA extended by 15 nucleotides on the 5' end (bottom lane); the corresponding views with the modeled DNA are displayed on the left.

structure independent of the structure obtained from the 200kV data set. This validates the structure, obtained at 200kV. The cross correlation coefficient of the two structures is 0.87. DNA sequences used are acAGGTCAcagAGGTCActc for the 200kV data, and cagAGGTCAtgaAGGACAtca and agaggatggagtcag-cgAGGTCAcgaAGGTCAc for the 100kV data.

3.5 100kV test structure with a 35mer DNA.

The data collection was done on the similar way, as the 100kV test dataset described above. However the concentration of the complexes that we observed was

rather low. The efficiency and the outcome of the data collection on CCDs was therefore insufficient to assemble of the reliable 3D reconstruction.

3.6 3D densities map interpretation

The interpretation of the 3D density map has been done by fitting existing crystal structures taking into account the connectivity between the two major parts of the complex. The models derived from the crystal structures of the dimer of LBD's and the DNA with DBD's have been fitted into the existing density map using the PyMOL software (http://pymol.sourceforge.net) allowing unambiguously annotating the RXR and VDR parts and assigning the handedness of the complex. The quantitative estimation of the fitting has been done in IMAGIC by computation of the cross-correlations between the 3D density map and electron densities simulated from the fitted X-ray structures. The fine scaling/magnification calibration of the map has been done according to the fitted crystal structure.

Results

4.1 3D reconstruction software

The essence of single particle analysis is the usage of many different images of individual molecules to reconstruct a three-dimensional (3D) distribution of the object densities, assuming that each image is a projection of the same object (Klug, 1982). Although the mathematical concept of restoring the 3D density from object projections was described by Radon in 1917 (Radon, 1917) it was considered for a long time as a sophisticated abstraction. Later, this theory has found many applications, and reconstruction algorithms have been described in several reviews (see Frank, 1996; Herman, 1980; Russ, 1995). All these approaches share the key assumption that while distribution of object viewing angles is random around the Euler sphere it still can be considered as the even one.

The quality of the 3D reconstruction depends on a number of images in a data set, the accuracy of the image alignment, assignment of angles to the particle images, angular distribution of the particle views, and the symmetry of the object. Electron microscopy (EM) analysis requires numerically stable and computationally effective methods so that large data sets can produce 3D reconstructions at near atomic resolution. Sample preparation used in EM both with negative staining and plunge freezing in liquid nitrogen leads to the fixation of molecules on the carbon film or within vitreous ice (cryo-EM). Molecular positions are fixed in ice and, therefore, an image of the complex is linked to the orientation of the molecule with respect to the electron beam, and can be used during the reconstruction procedure. The cryo-EM sample preparation, in addition to preserving the molecules in a hydrated, functional

state, has an advantage to negatively stained samples providing a more even distribution of orientations. However, the cryo-preparation does not ensure an even distribution of views, because hydrophobic areas on the molecule surface may lead to preferential orientations towards the air-liquid interface within the thin liquid film that spans the holes in holey carbon. The existence of preferential orientations leads to strengthening of densities in this direction and leads to distortions of structural features. Therefore, an uneven distribution of the complex orientations should be taken into account and compensated when performing 3D reconstructions. One of the possible solutions has been suggested and implemented by exact filtered back-projection approach while another approach is based on Voronoi diagrams in Fourier space (Harauz and van Heel, 1986; Penczek, 2004). The first approach is based on the assessment of cross- sections of all projections with the given one and doing pre-filtering of the projections before reconstruction according to the estimated distribution. This approach is rather efficient and can be easily parallelised, but it does not enable to deal with heavily distorted distribution of projections, like conical tilt. The method suggested by Penczek (Penczek, 2004) is based on consideration of Voronoi diagrams, but its implementation was not sufficiently optimised and became computationally time consuming.

The possibility of reconstruction of 3D object from its projections has been proved analytically by J. Radon in 1917. In this paper has been demonstrated that “the object can be reconstructed uniquely from its line projections when all of its line projections are known”. However, the Radon theorem does not address the question how to reconstruct the object from a limited number of projections. The theorem considers projections as integrals of the multi-dimensional function on a manifold of one-dimensional directions and assumes that all projections are ideal, without noise, and with infinite resolution.

In practice, having a limited number of non-ideal projections the object cannot be reconstructed up to the data quantisation frequency (with maximum precision) or up to the Niquest frequency. Each of the known 3D reconstruction algorithms either in real or in Fourier space has impression that is related to the algorithm and its realisation that can be described as the transfer function of the reconstruction process. Additional inaccuracy is caused by noise in the input data, numerical instability of the algorithm implementations, and by the limited number of projections.

The inaccuracy caused by the algorithms themselves can be named as a transfer function of the algorithm and can be interpreted as the “point spread function” (PSF) that depicts a shape and size of each point in the 3D density map. The PSF of registering systems describes a response of a detector device during recording of an image of the point, which in ideal situation with a system without any distortions and limitations in resolution should correspond to the delta (δ) - function. The PSF of a 3D object of limited resolution can be described as a Gaussian function with spherical symmetry. In Fourier space, PSF corresponds to the modulation transfer function (MTF) where deviation of the amplitudes from the constant represents the space frequencies of PSF.

The simulated example of a 30S small ribosomal subunit is shown on Figure 4.1. Two distributions are simulated – non uniform (Figure 4.1E) and uniform (Figure 4.1F). Panels 4.1A and 4.1B show the middle cross-section of the reconstructed density maps for non uniform and uniform distributions respectively. Panels 4.1C and 4.1D are the magnified fragments of panels 4.1A and 4.1B. The difference in between two 3D reconstructions is clearly visible. The non uniform distribution of projections results in an elongation of densities and all features of a 3D reconstruction starts to lose sharpness especially in the direction that corresponds to the majority of projections. In terms of PSF that situation can be interpreted as

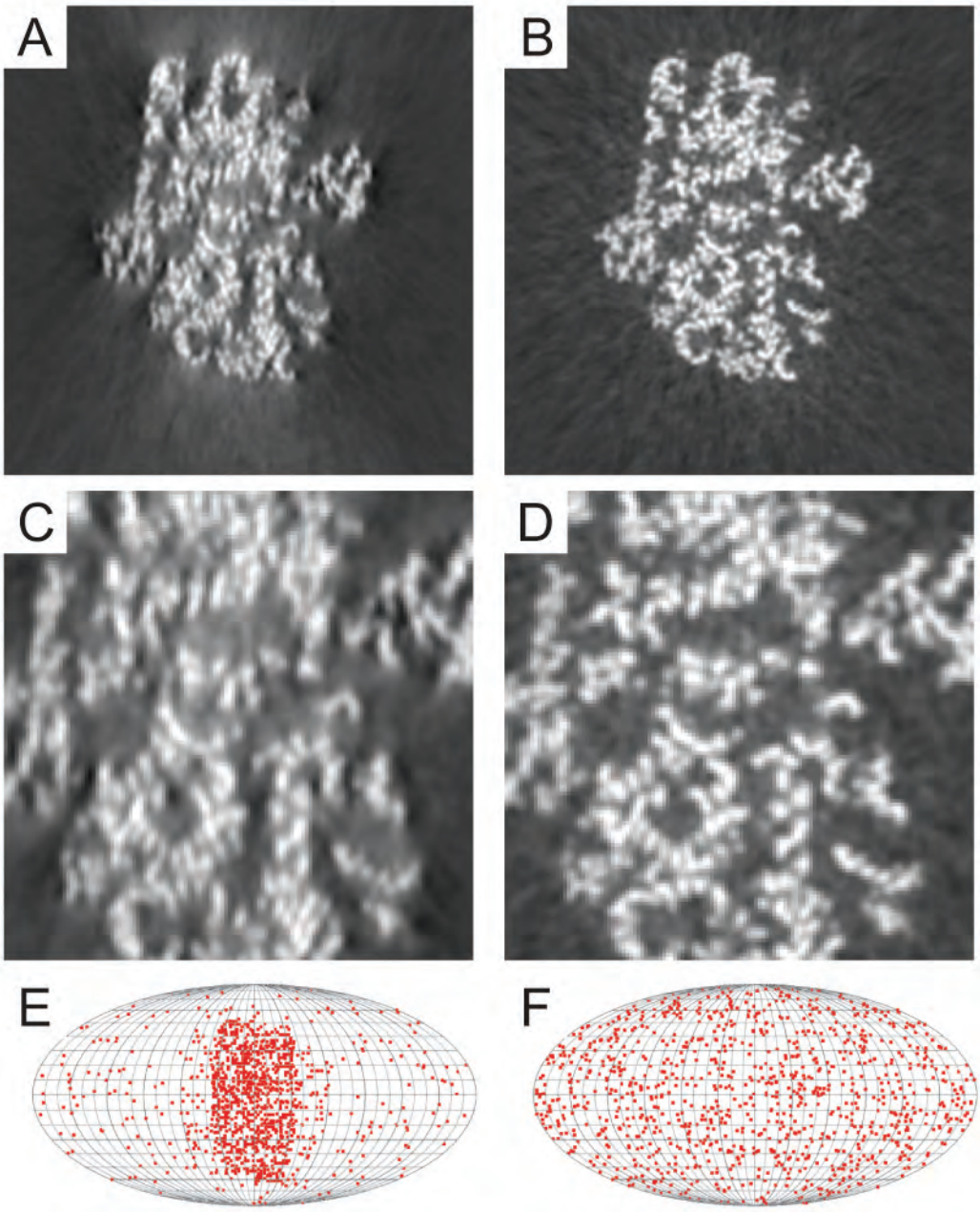


Figure 4.1 30S small ribosomal subunit test data. **A.** Central cross section of a 3D density map with preferable orientation. **B.** Central cross section – uniform distribution of projections. **C.** and **D.** Zoomed fragments of the previous panels respectively. E. and F. Simulated projections distributions – with and without peculiarities respectively.

elongation of PSF in one direction or non spherical (non uniform) point spread function.

To describe the algorithm of weighting one needs to explain the concept of Voronoi diagram on the plane (Voronoi, 1907; Aurenhammer, 1991). The Voronoi diagram on a plane is the partitioning of a plane with *n* points into convex polygons

such that each polygon contains exactly one generating point and every point in a given polygon is closer to its generating point than to any other (Figure 4.2 panels A and B). The points are called Voronoi sites. The cells are called Dirichlet regions or Voronoi cells (polygons) and edges of cells are called Voronoi edges. A Voronoi diagram is sometimes also known as a Dirichlet tessellation (Dirichlet, 1850). Each point on the boundary of the Voronoi cell, or a polygon, is equidistant from two points or Voronoi sites, and each vertex is equidistant from at least three Voronoi sites. As a result, the regions are edge to edge and vertex to vertex and therefore all regions form a polygonal partition of the plane without holes or not-partitioned spaces (Figure 4.2 panel B). The dual graph to the Voronoi diagram is the Delauney triangulation, well known in computer graphics (Delauney, 1934).

The Voronoi diagram on a sphere is equivalent and congruent to the diagram on the plane. The link is based on transformation of the plane to the sphere via, for example, the orthographical transformation. A distance between points on the sphere is defined differently from that in Euclidean space (Okabe, et al. 2000; Lee and Samet 2000). Now a set of *n* distinct points are located on the sphere with the unit radius and with the centre at the origin of the coordinate system. The shortest distance between two points *p* and *q* is defined by the length of the shortest arc of the great circle passing through *p* and q. This distance is called the “great circle distance”. The arc that defines an area of the point dominance on the sphere is a great circle that passes through the mid-point of the great circular arc linking *p* and *q* perpendicularly. This great circle is equivalent to the Voronoi edge and divides the sphere into two hemispheres. The intersection of hemispheres related with a given site gives the spherical Voronoi region or polygon associated with this site. The set of resulting spherical Voronoi polygons gives a generalised Voronoi diagram, which is

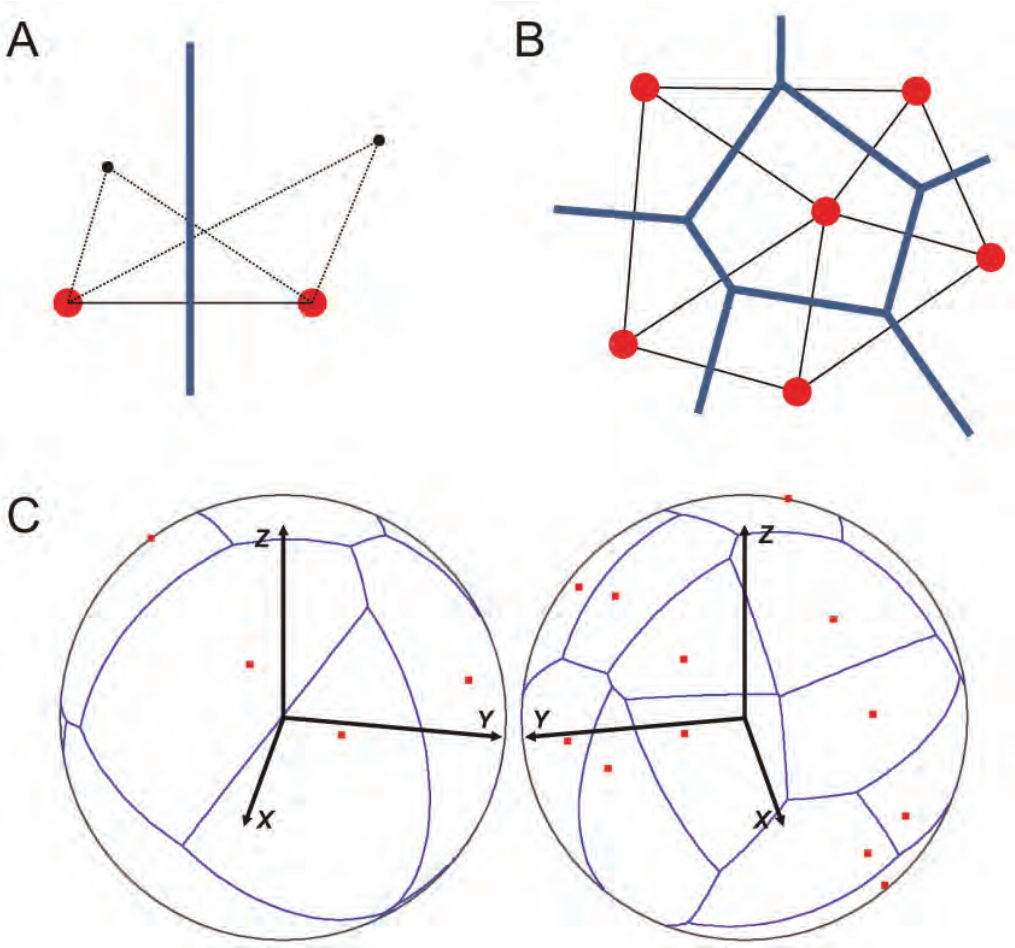


Figure 4.2 Voronoi diagrams on a plane and on a unit sphere. **A.** Basics of the Voronoi tessellation. Two Voronoi sites (red color) and one Voronoi edge (marine blue colour). **B.** Voronoi diagram on a plane. Black lines denote Delauney triangulation. **C.** Voronoi diagram on a sphere. Left – front view. Right – back view.

called the spherical Voronoi diagram which is the polygonal tessellation of a unit sphere (Figure 4.2C). The Voronoi diagram covers the sphere without gaps.

The concept of the Voronoi diagram on a sphere has been used to assign weight of each projection according to its region of dominance or Voronoi cell. Each Euler angle of a projection is represented on the unit sphere as a point and corresponds to a one facet of the Voronoi cell. The projection weight is proportional to the area of the cell. Each Voronoi cell on the sphere contains one and only one point corresponding to the single projection. If there are many projections with exactly the same angle, which is typical for projection matching, the weight is divided to the number of superimposing or overlapping projections. The “build” or the determination

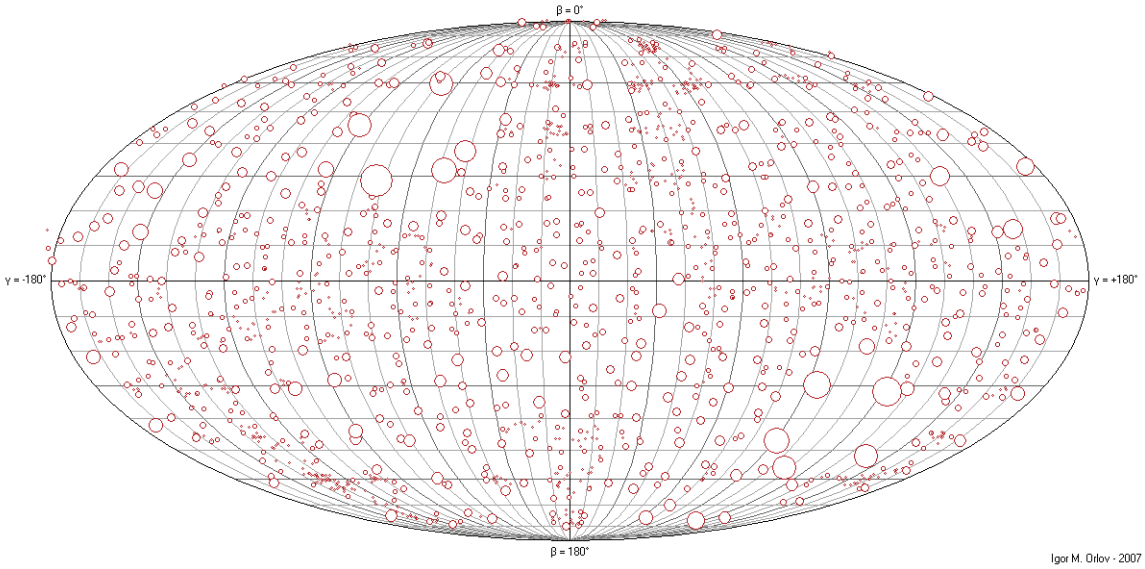


Figure 4.3 The distribution of projection on a unit sphere for the VDR-RXR-DR3 full complex structure. Circle sizes denote weights, assigned to individual projections according to the Voronoi diagram. The bigger is the circle size – the bigger is the weight assigned to a specific projection.

of the Voronoi diagram on a sphere can be done using Delauney triangulation. Computation of the Voronoi diagram is carried out as a dual transformation of the Delaunay triangulation.

This new approach has been tested on data with different point group symmetries. The software analysis the distribution of projections on the Euler sphere assigning weights for each projection according to the area of the Voronoi polygon on the Euler sphere (Figure 4.3). If there are several projections with exactly the same orientation, which is typical for projection matching approach, the software treats all those projections as one and down weights each specific projection according to the total number of projections with the same orientation.

We have modelled the effects of non-uniform 30S small ribosomal subunit from E.coli (Figures 4.4 and 4.5). The dataset has been simulated as a density map of the atomic structure (1YL4 PDB entry) in a cube 512, assuming 1Å per pixel (Figure 4.4A). The non-uniform distribution of projections had been simulated using 256 projections of the map (Figure 4.4D and 4.4E). This is a typical behaviour of

ribosomal specimen on carbon, when the preferable orientation is caused by ribosomes sticking to the carbon with one side. The distribution of projections for this 30S model dataset could be seen in 3D (Figure 4.4E) and as two-dimensional distribution on the Euler sphere (Figure 4.4D). Blue lines on Figure 4.4E show the corresponding Voronoi diagram on a sphere for this specific distribution of projections. It is clearly seen that the half sphere with high density of sampling shows smaller areas of Voronoi polygons and the opposite half sphere shows much bigger areas of Voronoi polygons.

The comparison of reconstructions with the model demonstrates clearly differences in the cross-sections of the “ideal”, non-corrected and corrected structures (Figure 4.4 A, B and C). While the main details on both structures are rather similar, the reconstructions reveal the presence of the background caused by the small amount of projections used for model calculations. An increase of the

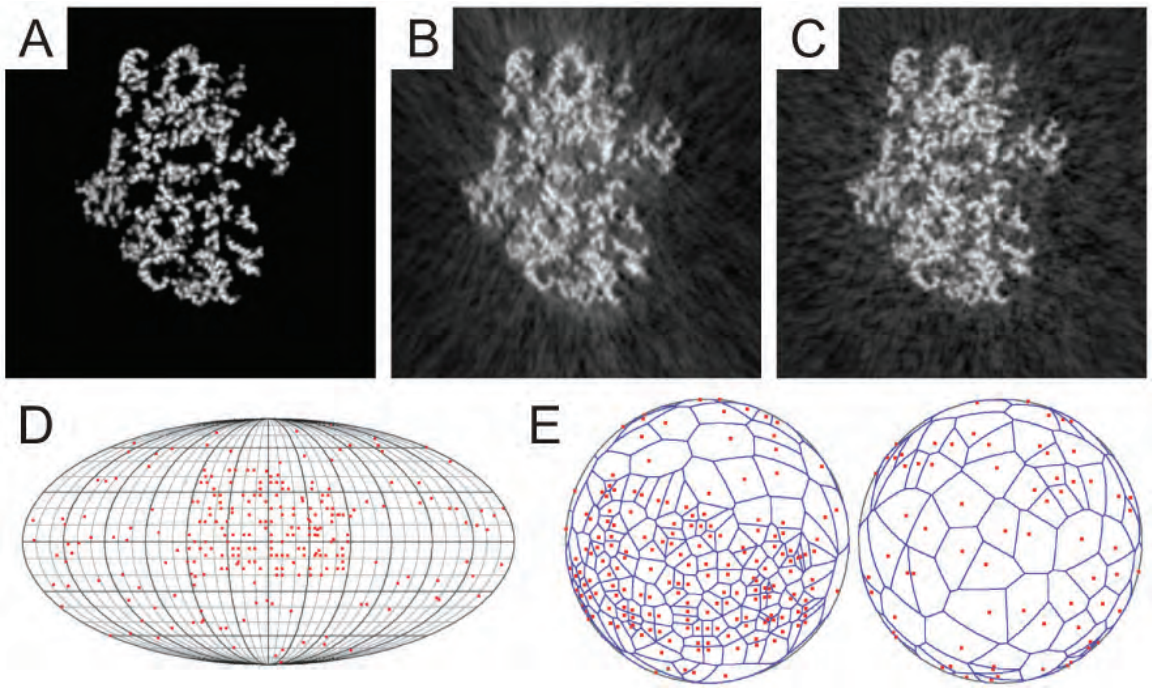


Figure 4.4 Test dataset of the 30S small ribosomal subunit with simulation of the preferable orientation. **A.** The central cross section of the ideal structure. **B.C.** Cross sections of the 3D reconstructions without and with correction of the preferable orientation respectively. **D.E.** 2D and 3D representations of the distribution of projections.

projection number improves significantly the contrast in the reconstructions, making the background less intense. The non-even distribution of projections affects the appearance of small features: in the non-corrected structure all small elements have the densities elongated in the direction where the projection directions were clustered (Figure 4.4D). In the corrected structure the corresponding areas the fine details do not reveal distortions of their shape.

We also tried to figure out the influence of the non-uniform distribution of particles to the projections of the structure. Orthogonal projections were computed from the test ribosomal reconstruction shown before and a comparison was done (Figure 4.5). The first row of Figure 4.5 presents projections from the ideal structure (Figure 4.5 panels A, B and C), the second row presents projections from the non-corrected reconstructions (Figure 4.5 panels D, E and F) and the third row presents projections from the corrected structure (Figure 4.5 panels G, H and I). The forth row presents the differences or variances between non-corrected structure projections and ideal structure (Figure 4.5 panels J, K and L). The fifth row, respectively, presents variances between corrected structure projections and ideal structure (Figure 4.5 panels M, N and O).

The non-uniform distribution of projections mostly exhibits itself in the first column of Figure 4.5 that shows the so-called “front view” of 30S ribosomal subunit. On the “front” projection from the non-corrected structure the elongated traces along the overpopulated direction are clearly visible. The corresponding variance image clearly shows that the major difference in between projections from corrected and not-corrected structures indeed is in the elongation of densities along a certain direction. Moreover, if for three-dimensional cross-sections of the density maps the difference is probably less noticeable, for projections it becomes quite strong (Figure 4.5D and corresponding difference image on Figure 4.5J). That means that in case

one uses the ordinary procedure of refinement by so-called projection matching the certain family of projections could be spoiled by those 3D reconstruction artefacts.

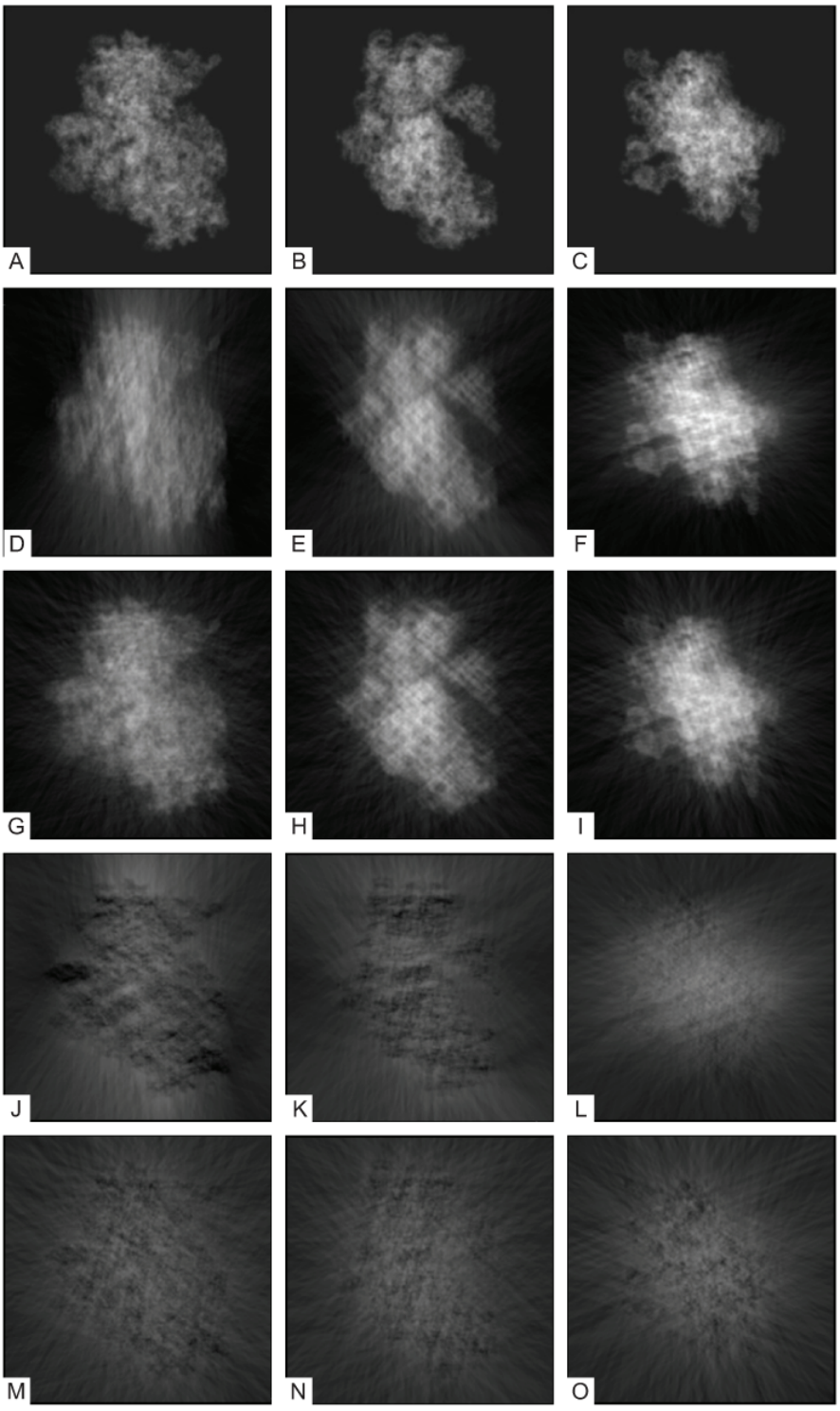


Figure 4.5 Orthogonal projections from the test dataset of the 30S ribosomal subunit. **A.B.C.** Orthogonal projections from ideal structure. **D.E.F.** Orthogonal projections from non-corrected structure. **G.H.I.** Orthogonal projections from corrected structure. **J.K.L.** Differences between non-corrected and ideal projections. **M.N.O.** Differences between corrected and ideal projections.

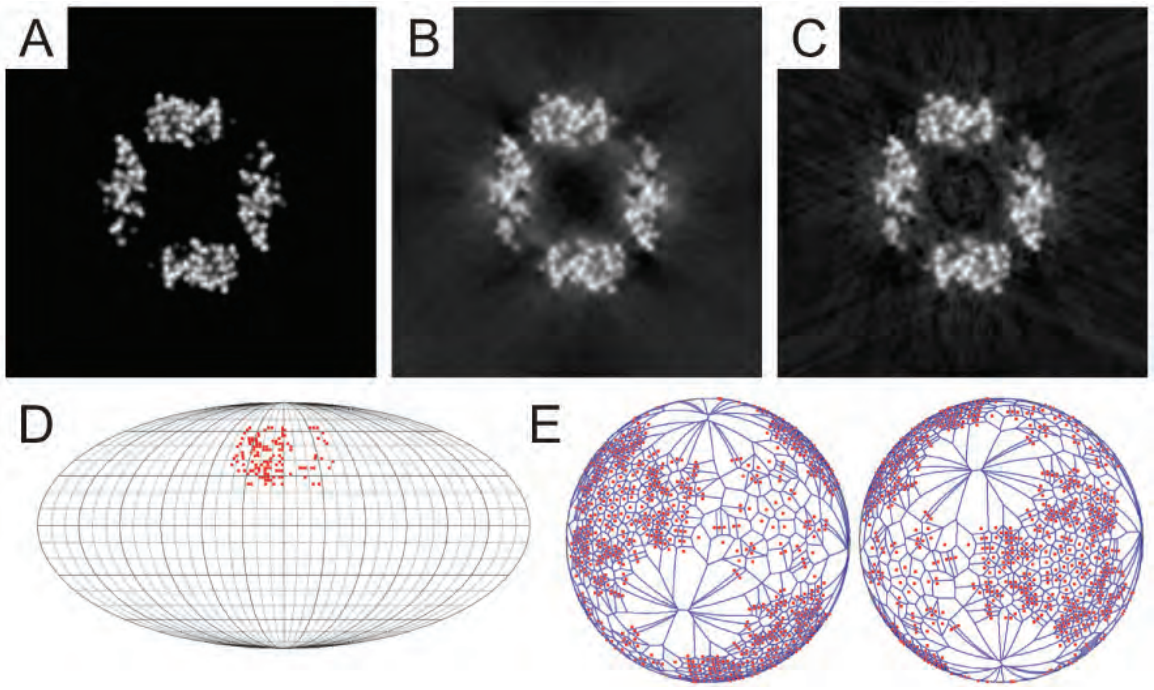


Figure 4.6 Test tetrameric reconstruction of the heat shock protein. **A.** The central cross section of the ideal structure. **B.C.** Cross sections of the 3D reconstructions without and with correction of the preferable orientation respectively. **D.E.** 2D and 3D representations of the distribution of projections.

Another example of influence of the non-uniform distribution of projections on the reconstruction with a high degree of symmetry is a heat shock protein (2BYU) which has the tetragonal symmetry (Figure 4.6). 128 projections have been computed within the tetrameric asymmetric unit (Figure 4.6AD). The corresponding Voronoi diagram for all asymmetric units is shown in Figure 4.6 panels D and E. The high degree of symmetry makes the influence of the strong preferable orientation significantly smaller. Although the strong preferable orientations in one asymmetric unit will be replicated 12 times for the tetrahedral symmetry resulting in a presence of 12 overweighed directions around the whole sphere instead of one for the non-symmetrical case.

Results clearly show the improvement of the reconstruction with PSF-compensation in comparison with the reconstructions without compensation. It is visible that the global features and the fine details resemble those of the original

reconstruction, the densities are not so elongated, and the PSF is more symmetrical. For a large datasets the real-space Voronoi approach is more computationally efficient, than the direct PSF computation in Fourier space.

4.2 VDR/RXR heterodimer complex with its target DNA

Nuclear receptors are involved in the ligand-dependent regulation of gene expression and thus represent important drug targets directly linked to severe diseases such as cancer, osteoporosis, obesity and type II diabetes (see for example Pinette et al., 2003; Friedrich et al., 2002; Fievet and Staels, 2009). The regulation of gene activity through the binding of specific ligands involves the full nuclear receptor through its DNA-binding and protein co-regulator recruitment and exchange abilities that regulate chromatin modifications and remodelling (Glass and Rosenfeld, 2000). We therefore set out to address the structural organisation of a full complex representative of NR function such as the heterodimeric RXR/VDR complex.

The RXR/VDR DNA-complex has been formed from co-purified VDR (full-length, lacking the VDR-specific insertion sequence residues 166-216) and RXR (lacking the A/B domain which is the less conserved part of NRs) and incubated with a double-stranded DNA segment of 20 nucleotides including a target response elemen. The DNA sequence comprises the consensus direct repeat acAGGTCAcagAGGTCActc with a three-nucleotide spacer, the 5’ and 3’ half-sites being recognized by RXR and VDR, respectively (Mangelsdorf and Evans, 1995). The structures of the RXR/VDR DNA-complexes (in the following, the first NR DBD named defines the 5’ position within the promoter sequence upstream of the transcription start site) have been determined by single particle cryo-EM and 3D reconstruction at a resolution of 10-12 Å (as estimated from Fourier-shell correlation and consistent with the features resolved in the map; see below and Figure 4.7). The

map reveals an L-shaped structure with one side split into a V-shape module, similar to a slingshot frame (Figure 4.7). The shape and size of the V-shaped part (~40x50x75 Å) can immediately be recognized as the LDB dimer, whereas the longer side of the L has proportions (~75x35x32 Å) characteristic for a DBD-bound DNA segment. The structure reveals that the RXR/VDR complex forms an open architecture (Figure 4.7), with the LBD hetero-dimeric unit oriented almost perpendicular to the DNA double-helix. Notably, a clear density attributable to the LBD-DBD hinge region connects the LBD part to the DNA-DBDs sub-complex.

In order to address the localization of the individual nuclear receptor domains in more detail and provide a full interpretation we fitted the available crystal structures of the LBDs and DNA-bound DBDs into the cryo-EM map. For this, we built a model of the RXR/VDR LBD based on the RXR/RAR LDB crystal structure (Bourguet et al., 2000) in which the RAR part was replaced by the VDR LBD crystal structure (Rochel et al., 2000); this takes into account the high conservation of LDB homo- and heterodimers. Moreover, since no functional DNA-bound RXR/VDR DBD structure is available for a DR3 element with RXR and VDR DBDs on the proper 5' and 3' DNA ends respectively, we built a model for the DR3 RXR/VDR DBD complex (see Methods) derived from the DR4 RXR/TR DBD (Rastinejad el al., 1995) and the VDR part of a DNA-bound 5'-VDR/RXR-3' crystal structure (Shaffer and Gewirth, 2004). Fitting the RXR/VDR LBD and DBD models allowed unambiguously annotating the RXR and VDR parts and assigning the handedness of the complex (Figure 4.7), as has been validated also by calculating the cross-correlation coefficients between the maps and the different possible fittings (Figure 4.10). Indeed, the VDR and RXR LBDs can be identified unambiguously thanks to two characteristic asymmetric features in the VDR LBD, the VDR segment comprising helix H4/loop 8/9 and RXR helix H7 and the area comprising helices H2, H3n and the ß-sheet (marked with a

star in Figure 4.7 B&C). Importantly, this features resolve the possible ambiguity residing in the pseudo-symmetry of the RXR and VDR LBD heterodimer (see view on the pseudo-two-fold symmetry axis in Figure 4.7 C). For the DNA-DBDs part, we fitted the DR3 RXR/VDR DBD model into the cryo-EM map (Figure 4.7 A,B&C). The fit determines the 5'-3' orientation of the DNA/DBD complex thanks to the precision of the DBD densities, and it is consistent with the connectivity constraint between the LBDs and DBDs of RXR and VDR, respectively. In addition, a complex with a DNA extended by 15 nucleotides on the 5' side confirms the polarity assignment of the DNA. The map clearly reveals that the DBDs lie "side-on" with respect to the DNA (Figure 4.7 B), with a part on the DNA side facing the LBDs and not opposite from them. The fitting also addresses the polarity of the complex with respect to the DNA, showing that the main anchor point to the DNA occurs on the 5' side (Figure 4.7 A) through the hinge region and the DBDs rather than for example on the centre of the response element, with the RXR DBD indeed on the 5'-end as can be anticipated for a DR3 element.

The structure of the complex reveals that the LBD heterodimer is a rather rigid and stable unit of the complex also in the presence of DBDs and DNA. This is consistent with the high sequence conservation of the LBD interface (Brelivet et al., 2004) and the stability of LBD heterodimers which has facilitated the determination of several crystal structures (RXR/RAR (Bourguet et al., 2000); RXR/PPAR (Gampe et al., 2000), RXR/LXR (Svensson et al., 2003)) but not yet for RXR/VDR. The present structure suggests that the RXR/VDR LBD adopts the same arrangement as RXR/RAR (Bourguet et al., 2000) or RXR/PPAR (Gampe et al., 2000) LBDs. Notably, the map indicates the position of the central helices of the LBD interface comprising helices H9 and H10/11 of VDR and RXR (Figure 4.7 C), and a nice mass of density is attributable to the LDB interface contact between helix H4/loop 8/9 (VDR) and helix

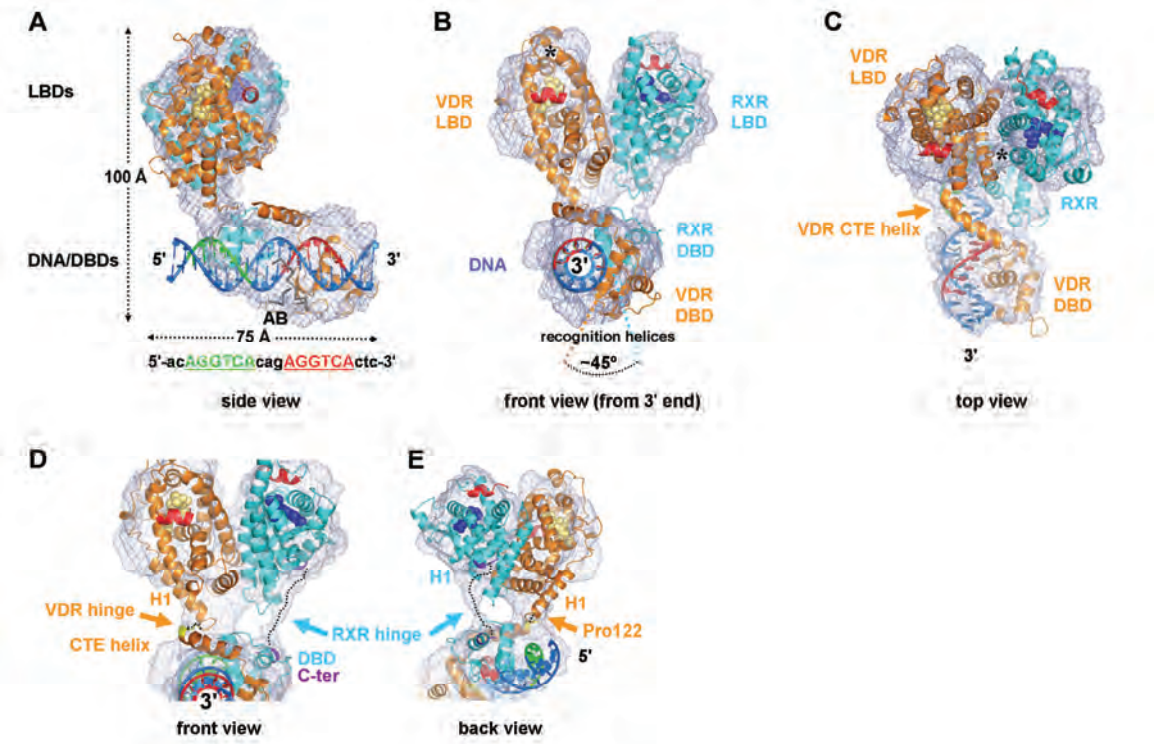


Figure 4.7 Structure of the RXR/VDR/DR3 DNA nuclear receptor complex. **A.** Side-view of the structure, with the 5' DNA end on the left. The cryo-EM map is shown in magenta, the fitted DNA/DBD and LBD heterodimer parts are shown with their backbone secondary structure (models are derived from available crystal structures, see main text). The DNA is shown in blue with the first half-site of the response element in green and the second half-site in red. The DBDs of RXR (cyan) and VDR (orange) are bound on the back, and a density close to the recognition helix of VDR is attributable to the AB domain of VDR (17 residues). The LBDs are oriented perpendicular to the DNA, anchored through the 5' side of the response element. The dimensions of the complex are indicated as well as the sequence of the response element. **B.** Front view of the complex as seen from the 3' end of the DNA (rotated 90° with respect to the view in **A.**). **C** Top view of the complex as seen along the pseudo two-fold axis through the interface of the LBDs (rotated 70° downwards with respect to the view in **B.**). **D.E.** The hinge regions between the LBDs and DBDs of RXR (cyan) and VDR (orange).

H7 (RXR; Figure 4.7 C). The orientation of the LBD dimer is approximately perpendicular to the DNA, with only a small tilt (the pseudo two-fold symmetry axis through the LBD heterodimer is tilted ~20º towards the downstream side of the DNA, and the heterodimer is rotated by ~5º on the VDR side). In the presence of an agonist ligand as in the complex here, the trans-activation helix H12 (marked in red in Figure 4.7) is known to close the occupied ligand-binding pocket and adopt an active

conformation required for the recruitment of co-activators such as steroid receptor co-activators (SRCs (McKenna et al., 1999)). Notably, the structure of the RXR/VDR DNA complex shows that H12 and the co-activator binding site of VDR point towards the 3' direction and side-on to the DNA (opposite to the DBDs; Figure 4.7 A&B), suggesting that the co-activator (which preferably binds to the RXR-partner) acts downstream from the NR response element. This is consistent with the idea that DNA-bound nuclear receptors maintain the trans-activation helix H12 area accessible for recruiting chromatin-modifying co-regulator proteins that act on the nucleosome-bound region of the downstream DNA.

The RXR and VDR DBDs are resolved as individual entities in the map, thus revealing their precise position on the DNA. They are rotated with respect to each other by ~45º around the DNA axis which is consistent with the 3-nucleotide spacer between the 6-nucleotide direct repeats (Figure 4.7 B). The RXR DBD is positioned on the 5'-side of the DNA, while the VDR DBD is on the 3'-side, as can be recognized from the characteristic density observed for the VDR-specific C-terminal extension (CTE) helix expanding from the DBD core, while the RXR DBD has no such feature. From the proximity of RXR and VDR DBDs on the DNA is seems likely that they interact with each other. An additional density is visible in the major groove next to the VDR recognition helix which may correspond to the short A/B domain in VDR (17 residues at the N-terminus, Figure 4.7 A). The VDR CTE connects directly to the N-terminal helix H1 of the LBD (Figure 4.7 D) residue Leu 120 is the terminal residue seen in the individual LBD and DBD crystal structures). For RXR, the DBD C-terminal end extends into a discrete density connected to the RXR LBD, thus indicating the position of the hinge between the two core domains (Figure 4.7 B). The CTE of RXR has not been modelled here since its structure is unknown and it has no sequence similarity with the CTE helix region of VDR or TR. However, the features in

the map and the distance between the terminal residues of the RXR LBD and DBD (~35 Å) suggest that the RXR linker contains a rather well-defined secondary structure (possibly two helical segments with a short extended loop; see dotted line in Figure 4.7 D&E). While VDR has a straightforward connection between the DBD and LBD through the CTE helix, the RXR hinge needs to reach the LBD helix H1 at the back side of the LBD (oriented 5') and thereby crosses the tip of interface helix H9.

The first part of the two hinges protruding from helix H1 to the DBDs are resolved as two individual segments (Figure 4.7 D&E), and the hinges remain independent also when they kink into the DBD/DNA sub-complex (Figure 4.7 A). This suggests that residues of the VDR and RXR hinge region do probably not interact with each other. However, the fact that the hinges provide two independent anchor points to the DNA clearly contributes to the stabilisation of the complex in a precise conformation with respect to the DNA strand. The well-defined conformation of the complex, notably between LBDs and DBDs, is also suggested by a 3D classification test (Simonetti et al., 2008) that when applied to the data revealed no major variations in the structure. In this context, it is of particular interest that the VDR linker peptide carries a proline residue (Pro 122) right at the junction of the two helical elements (CTE and H1, sequence KDSLR**P**KLSE; Figure 4.7 E) which favours the formation of the observed kink between the DBDs and LBDs thus leading to the characteristic L-shape conformation of the complex (Figure 4.7 A). The proline residue is conserved in the VDR family and also in several other NRs (RAR, TR, PPAR, LXR, PXR, CAR) including their common heterodimeric partner RXR and therefore marks the N-terminal edge of the LBD domain. The conserved charged residues of the RPK motif (residues 121-123, tip of LBD helix H1) may contribute to an additional stabilization of the complex through interactions with the phosphate

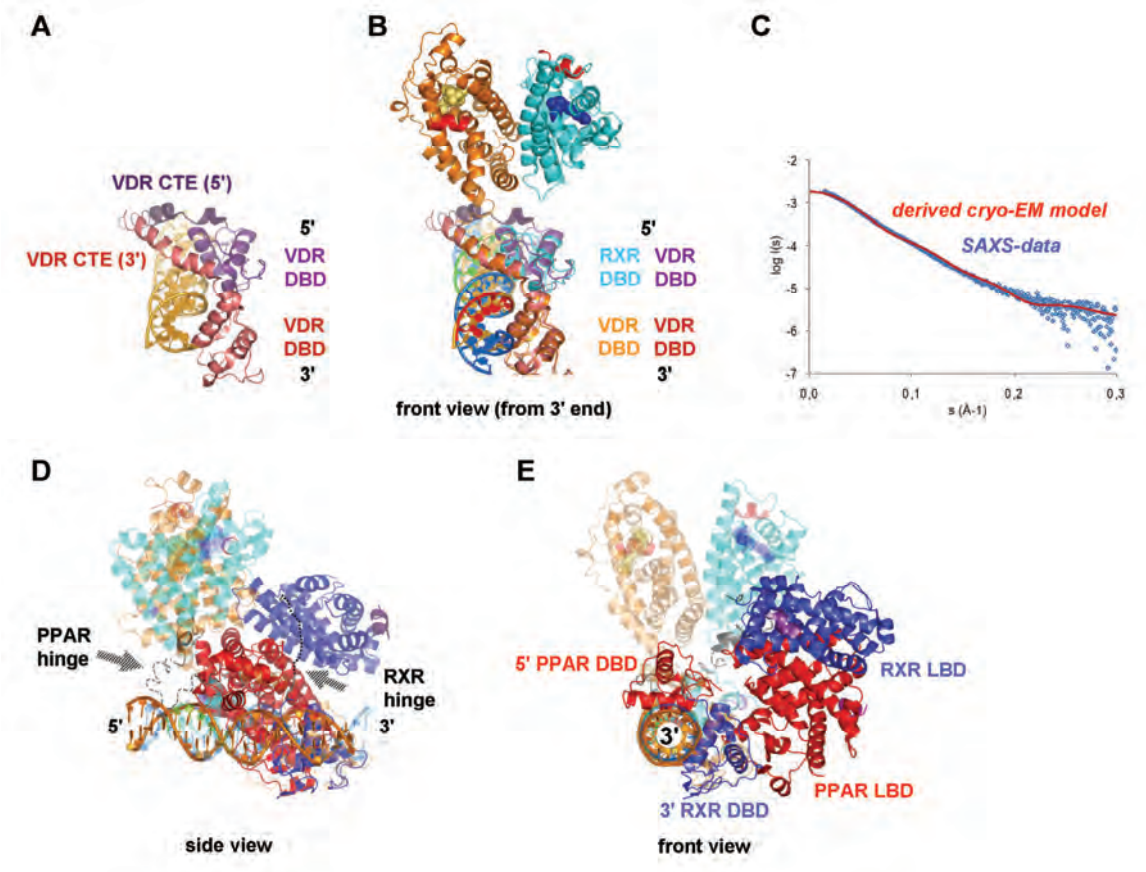


Figure 4.8 Comparison of the RXR/VDR/DR3 cryo-EM structure with SAXS data and with known crystal structures. **A.B.** Comparison of the crystal structure of the VDR homodimer DBD on a DR3 response element (PDB ID 1KB4) super-positioned to the DBD part of the cryo-EM structure of the RXR/VDR/DR3 complex. **C.** Comparison of the small angle X-ray diffraction (SAXS) curve of the RXR/VDR/DR3 complex in solution (measurements are represented as blue dots) with the theoretical curve (red) calculated from the model that was derived from the cryo-EM map. The global arrangement and conformation is rather similar, as is illustrated by the good correspondence at lower and medium s values (describing low and medium resolution components such as size and overall conformation and structure, the most prominent parameter being the radius of gyration). **D.E.** Comparison of cryo-EM structure of the RXR/VDR/DR3 complex with the crystal structure of the PPAR/RXR/DR1 complex super-positioned through the DNA and the 3' DBDs (PPAR in blue, RXR in red, co-activator peptide In magenta).

backbone of the central part of the 5' half-site, in direct vicinity to the C-terminal tip of the RXR recognition helix (Figure 4.7 E).

The organisation of the hinge region has important implications for NR function. The long, partially extended CTE region of RXR is needed for reaching helix H1 on the side opposite from the RXR DBD (i.e. further up in the 5' direction; Figure 4.7 D&E), explaining why RXR cannot have a rigid CTE helix such as VDR or TR. In

addition, the flexible linker of RXR is required for the adaptation to different DR response element types (DR2,3,4,5) with the 5'-bound hetero-dimerisation partner RXR, or DR1's with a 3'-bound RXR. In contrast, being positioned on the 3' side of the DR3 far away from the LBD (~35 Å), the VDR DBD requires a long straight segment to reach the LBD N-terminus, which is achieved through the VDR CTE helix (Figure 4.7 C). The architecture of the complex thus rationalizes the existence of the CTE helix which protrudes from the DBD core and crosses the DNA strand perpendicular to the minor groove in order to join helix H1 of the LBD (Figure 4.7 C&D). This also applies to the TR which has a CTE helix oriented in the same way over a DR4 element (Rastinejad et al., 1995). The combination of a rigid geometry of the CTE in VDR, TR and possibly other related NRs, with a more flexible RXR CTE suggests that the adaptability of NR heterodimers to different response elements resides in RXR. The structure suggests that the CTE helix stabilizes the conformation with respect to the DNA while other nuclear receptor complexes and response elements may adopt a different organization which requires adaptability of RXR (rather than intrinsic flexibility or instability of the whole complex on a given DNA element). Consistently, mutations in the hinge region of VDR and RXR show that the linker length has a critical role in VDR but not so in RXR (Shaffer et al., 2005).

The structure of the RXR/VDR DNA complex provides first insights into the architecture of a functional heterodimeric nuclear receptor complex comprising a standard direct repeat in which RXR occupies the 5' position. The complex is therefore representative of the entire class of RXR heterodimers that bind to direct repeats with 2, 3, 4 or 5 nucleotide spacers with RXR at the 5' end: RXR/RAR (DR2), RXR/VDR (DR3), RXR/TR (DR4) and RXR/RAR (DR5; reviewed in ref. 6); in contrast, DR1 elements bind RXR on the 3' side with a reversed polarity on the DNA (e.g. PPAR/RXR (Chandra et al., 2008), RAR/RXR (Rastinejad et al., 2000),

RXR/RXR (Zhao et al., 2000)). Interestingly, the arrangement seen here for an RXR/VDR DR3 complex is fully consistent with that of the VDR DBD homodimer on a DR3 element (Shaffer and Gewirth, 2002) (Figure 4.8 A&B) since in the fitted model RXR precisely takes the position of the 5' VDR segment (with the exception of the distinct CTE part), suggesting a common architecture for DR3 NR complexes independently of whether these are hetero- or homodimers. In other words, the 5' DBD in homodimers appears to play the role of the RXR DBD. In the homodimer structure the tip of the 3' VDR DBD CTE helix is bent towards the DNA axis (Figure 4.8 A&B) and thus points to the connection point of helix H1, consistent with the VDR LBD position in the present structure.

The analysis of the RXR/VDR DNA complex by small angle X-ray diffraction (SAXS) reveals an architecture and conformation rather similar to that seen by cryo-EM and the derived cryo-EM model is consistent with the SAXS curve (χ^2=2.4; Figure 4.8 C) corroborating that in solution the LBDs are approximately perpendicular to the DNA/DBDs. In contrast, the recently published crystal structure of PPAR/RXR on a DR1 response element (Figure 4.8 D&E) is rather different with respect to: (i) the architecture of the complex, (ii) the type of response element and the polarity on the DNA, (iii) the definition of the hinges, and (iv) the contacts between the individual domains. The DNA in the PPAR/RXR complex contains an idealized DR1 element, with PPAR bound to the 5' half-site, while RXR occupies the 3' position. This results in a reversed polarity when compared to the RXR/VDR complex, which is expected for a DR1 complex as opposed to DR2,3,4,5 elements. The DR1 leads to an arrangement in which the DBDs are rotated with respect to each other by ~110°, with the DBDs on either side of the DNA (Figure 4.8 E) rather than on the same side as in the RXR/VDR complex (Figure 4.7 B). Aligning the RXR/VDR and the PPAR/RXR structures with respect to the 3' DBDs reveals that the two complexes have a

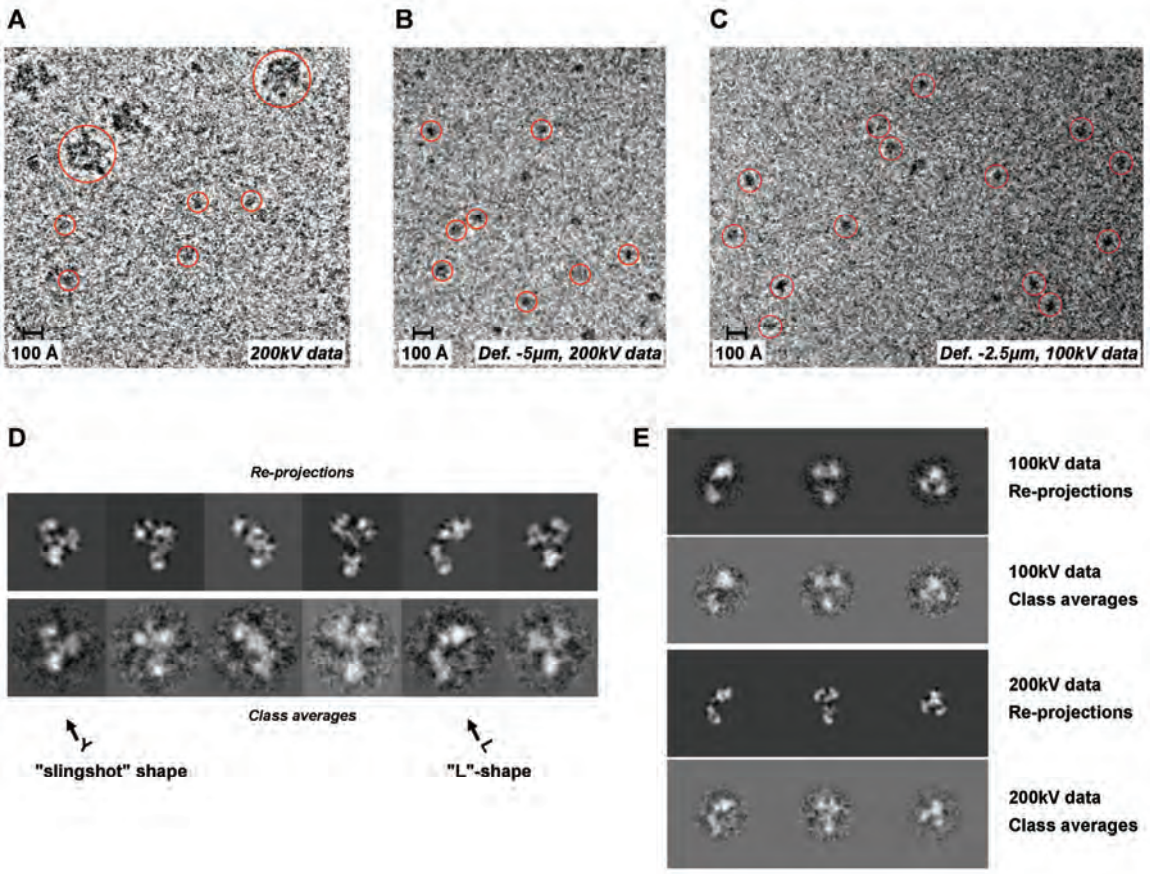


Figure 4.9 Cryo-EM single particle images of the 100kDa nuclear receptor complex. **A.** CCD image of a flash-frozen hydrated sample containing the purified nuclear receptor mixed with 50S ribosomal subunits for comparison; scale bar is 100 Å; recorded on a transmission electron microscope under cryo conditions at a defocus of ~3 µm at 200kV. Some particles are marked with red circles (the 100kDa NR particles are ~100 Å in size, while the ~1.5MDa 50S particles are up to ~220 Å in size). **B.** CCD image of a flash-frozen hydrated sample after screening for optimal ice quality and particle distribution; recorded at a defocus of ~5 µm at 200kV. **C.** CCD image of the nuclear receptor complex sample, recorded at a defocus of ~2.5 µm at 100kV. **D.** Comparison of class averages (first lane, obtained by multivariate statistical analysis and classification using the 200kV data) with the corresponding re-projections of the three-dimensional reconstruction (second lane). **E.** Class averages and corresponding orthogonal re-projections obtained for two independently determined *ab initio* structures using either 100kV or 200kV data. The two structures are very similar and fully consistent with each other (cross correlation coefficient is 0.87) and with the corresponding input class averages.

different architecture since (i) the 3' RXR DBD has no CTE helix in order to reach the LBD, and (ii) the PPAR DBD has its CTE region pointing into a completely different direction compared to the RXR CTE in the RXR/VDR complex because of the interactions with the 5' DNA extension element (Figure 4.8 E), resulting in the LBDs

facing the same side of the DNA as the DBDs (Figure 4.8 D). The two hinge regions are rather disordered in the PPAR/RXR complex (either invisible or high temperature factors characteristic of a flexible area; Figure 4.8 D), such that the stabilisation of a particular conformation cannot be explained by the hinge geometry itself. Rather, the observed conformation is probably related with an additional DBD-LBD contact between the PPAR LBD and the RXR DBD, while the RXR/VDR complex solely shows DBD-DBD and LBD-LBD contacts through the classical heterodimeric interfaces, the only DBD-LBD contacts being the well-structured hinge connectors (Figure 4.7). While SAXS and cryo-EM describe the structure of the macromolecular complex in solution, the crystal packing in the case of the PPAR/RXR complex may have led to a particular conformation of the LBDs with respect to the DBDs/DNA which is not seen in the complex studied here, nor in the PPAR/RXR complex investigated by SAXS (Rochel *et al.*, accompanying paper).

The study of a 100kDa macromolecular complex by single particle cryo-EM and 3D reconstruction to about 10-12 Å resolution has been a challenge by itself. The key parameters to overcome the difficulties related to the small molecular weight and size of the complex resulting in a low contrast in the cryo-EM images have been (i) rigorous screening of flash-frozen grids for optimal ice thickness (Figure 4.9), (ii) defocus pairs to improve the image contrast and localize the objects and get a first 3D reconstruction through angular reconstitution, (iii) validation of particle selection and alignment through multi-variate statistical analysis and classification of the images (Figure 4.9), and (iv) the usage of a lower acceleration voltage at which the image contrast is notably better (Figure 4.9). Data collected at 100kV were used to produce an independent *ab initio* reconstruction that is fully consistent with the structure refined using data collected at 200kV (inter-map correlation of 0.87; Figure 4.9 E). These data provide important evidence, through the usage of lower high

tension, for the reliability of the final structure, and thus validate the usage of cryo-EM for this relatively small macromolecular complex. Finally, the re-projections from the obtained structure are in agreement with the input class averages (Figure 4.9), and the cryo-EM map is also consistent with the size and features of the fitted LBD and DBD crystal structures (Figure 4.7). The features in the refined structure indicate the level of precision that has been reached, as illustrated by the smallest distance resolved between the LBDs (~10-12 Å), such that specific areas could be resolved (e.g. the α-helix H9/H10 bundle at the LBD interface, the hinge region of RXR) and the handedness of the complex be determined unambiguously. The study thus

	Map Orientation	***DNA Polarity***	***LBD's Position***	***Cross Correlation***
1	Direct	Correct	Correct	0,832
2	Direct	Correct	Flipped	0,823
3	Direct	Flipped	Correct	0,828
4	Direct	Flipped	Flipped	0,823
5	Mirrored	Correct	Correct	0,826
6	Mirrored	Correct	Flipped	0,812
7	Mirrored	Flipped	Correct	0,810
8	Mirrored	Flipped	Flipped	0,810

Figure 4.10 Cross correlation (CC) coefficients for different fitting possibilities of the LBDs (as a heterodimer unit) and the DNA/DBDs; flipping of the LBDs inverses the RXR and VDR positions around the pseudo two-fold axis (Figure 4.7), while flipping the DNA/DBDs inverts the polarity of the DNA. “Correct” is the orientation corresponding to the correct solution, and also represents the solution that is compatible with the connectivity constraint through the hinge region between LBDs and DBDs (specifically, solution 3 would not allow connecting the domains for distance reasons). The comparison of CC values confirms the unambiguous assignment of the components and also the handedness of the complex.

moves the generally accepted limit of 200-300kDa ahead by showing that the cryo-EM study of complexes with a molecular weight around 100kDa is possible, provided that a careful validation of the structure is done. It should also be noted that an internal symmetry would be beneficial for data processing albeit not essential as illustrated by the present work. Studies of this type may in the future be facilitated by more sensitive CCD cameras and possibly by energy filters and aberration correctors. The present work opens the way to structural studies of a whole range of relatively small macromolecular objects including those of protein / nucleic acid complexes which represent the majority of current structural biology targets but may have been difficult to crystallize or judged too small for standard cryo-EM methods. Thus, the study of NR complexes with different response element types such as inverted, everted and direct repeats can be envisaged in order to address their different, specific architectures.

5. Conclusions and perspectives

This work is dedicated to the structural study of the full-length VDR/RXR-DNA complex by single-particle cryo-EM. As a result of this work we managed to develop the methodology of an electron microscopy data collection on small proteins under cryo-conditions, i.e. hydrated, physiological conditions. We have established a reliable cryo-EM procedure for the data collection of the nuclear receptor complex. This allowed us to collect data of the VDR-RXR nuclear receptor hetero-dimer with a short DNA fragment (22 base pairs) and obtain a 3D reconstruction of the complex with 12Å resolution. This work also contributes to the methodology of a single particle 3D reconstruction we also did a software development trying to improve the reliability, speed and numerical stability of single particle 3D reconstruction procedure. We have introduced a new approach for the compensation of preferable orientations based on Voronoi diagrams that overcomes the problem of non-uniform distribution of molecular views. Weighting of projections is important in particular for the determination of asymmetrical structure (e.g. nuclear-receptors, ribosome or RNA polymerase), while it is less critical for high-symmetry cases.

The structure of the full-length VDR/RXR-DNA complex reveals that the LBD heterodimer is a rather rigid and stable unit within the complex formed with the DNA. This is consistent with the high sequence conservation of the LBD interface and the stability of LBD heterodimers which has facilitated the determination of several crystal structures (RXR/RAR; RXR/PPAR, RXR/LXR) but not yet for RXR/VDR. The present structure suggests that the RXR/VDR LBD adopts the same arrangement as RXR/RAR or RXR/PPAR LBDs. The orientation of the LBD dimer is approximately perpendicular to the DNA. Notably, the structure of the RXR/VDR DNA complex shows that transactivation helix H12 and the co-activator binding site of VDR point

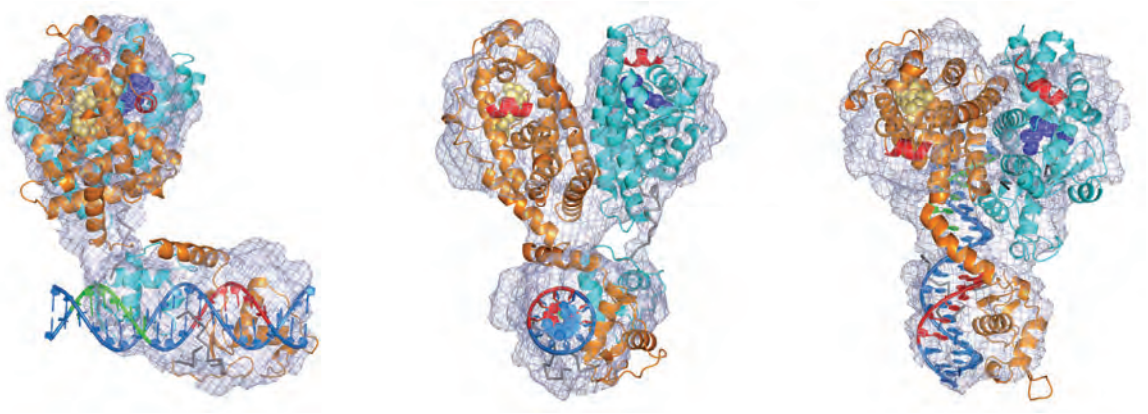


Figure 5.1 Left, front and top views of VDR-RXR-DR3 Short structure.

towards the 3' direction and side-on to the DNA (opposite to the DBDs), suggesting that the co-activator acts downstream from the NR response element. This is consistent with the idea that DNA-bound nuclear receptors maintain the trans-activation helix H12 area accessible for recruiting chromatin-modifying co-regulator proteins that act on the nucleosome-bound region of the downstream DNA (Figure 5.1).

At the time this project was started, no crystallographic structure of a full NR complex was available. However in 2008 the structure of the PPARγ-RXRα nuclear receptor on DNA was published (Chandra et al. 2008). That is a quite interesting situation because two structural studies were done on two different NR complexes using two different methods and independently from each other. The comparison of two different approaches can potentially provide complementary information about the structure.

While the two complexes have some differences that may limit their comparison they have also some similarities. The PPAR-RXR complex is bound to the DR1 response element that has reversed polarity in comparison with DR2, 3 and 4 response elements. But, on the other hand both structures present heterodimers (PPARγ-RXRα and VDR-RXR respectively) bound to a short DNA fragments (19 and 22 mers). Both structures are expected to reveal the mechanism of interaction of NR

heterodimers with the DNA. It was expected to see in both cases the common motif of structural organization and more-or-less the same conformation as a consequence of having a common mechanism of interaction. Some minor conformational variations were expected mostly caused by a difference in response element. Therefore the difference in polarity was expected.

Surprisingly, the crystal and cryo-EM structures show quite different conformation. The heterodimers of ligand binding domains in both structures have a common position on the DNA. The DNA is positioned orthogonal to the LBDs. Ligand binding domains and DNA binding domains have a quasi-tetrameric organization. However, the crystallographic structure exhibits the unexpected contacts in between ligand binding domain of PPARγ and RXRγ DBD. The hinge regions of both PPAR and RXR are wide open and seem not to be involved in an overall complex formation. On the contrary, the EM structure while having the same pseudo-tetrameric motif shows a clear “open” conformation without apparent contacts between LBDs and DBDs. The overall complex formation seems to be based on the strong interaction of the VDR hinge region with the DNA that provides the stability to the complex without contacts between DBDs and LBDs (Figure 5.2).

As a result two hypotheses of the mechanism of such a complex formation exist. One, which is from the crystallographic structure, suggests that the NR-DNA complex formation is based on the interaction of one LBD with the DBD of RXR and that is the structural basis of the complex. The hinge regions of the NR heterodimer would play a secondary role in the complex formation. The second hypotheses is based on the cryo-EM study, assumes that the heterodimer complex formation with the DNA is mostly based on the RXR partner hinge region interaction with the DNA. The mechanism will generally be the same for all NR heterodimers where RXR is involved as a partner. The stability and the absence of major heterogeneity that was

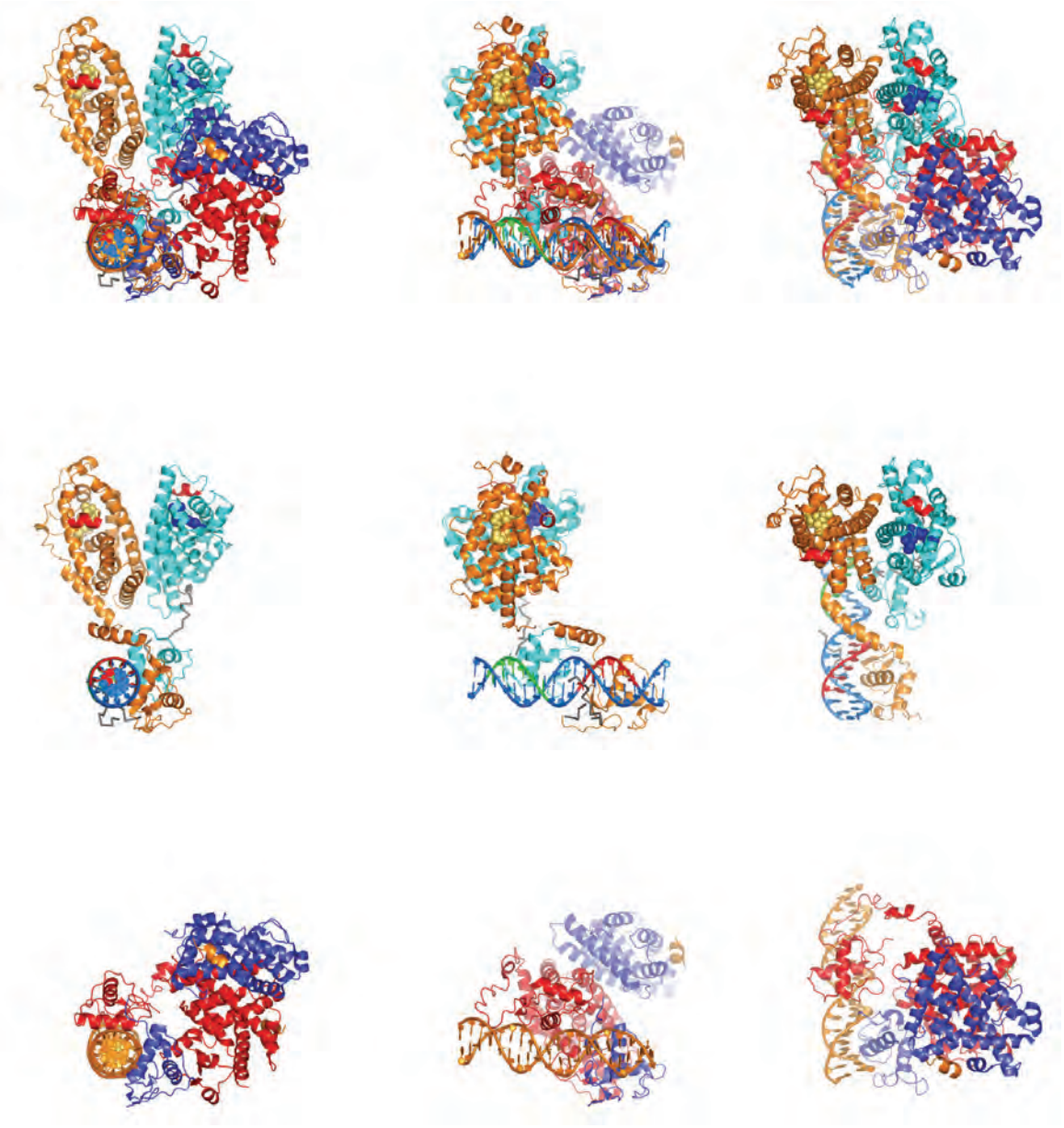


Figure 5.1 VDR-RXR-DR3 and PPARγ-RXRα-DR1 structures. First row – the superposition of two structures using DBDs and the DNA as an anchor point. Second row – three views of VDR-RXR structure. Third row – three views of PPARγ-RXRα structure.

observed in cryo-EM studies are also reinforced by the hinge region – DNA interaction. In turn, the hinge region of RXR is quite flexible and provides some flexibility to the entire construct to allow adaptivity of different RXR-related heterodimers to different hormone response elements.

Several parts in that kind of complexes however still remain unaddressed by both methods. The localization of the A/B domain, that is apparently non-structured, is quite problematical by any methods. To make structural studies on such objects

one needs to identify potential functional partners and investigate complexes with such proteins. Possibly the further functional studies will reveal the functional role of the A/B domain and it will be possible to identify functional partners needed to stabilize the A/B domain. The hinge regions of both LBDs are localized in the cryo-EM map but cannot be traced at 12Å resolution. Further developments of methodology and especially a high-resolution data collection procedure could help to improve the resolution.

Although remarkable works of Wah Chiu and Hong Zhou groups (Jiang et al. 2008, Yu et al. 2008) on icosahedral objects show the potential of cryo-EM to achieve an atomic resolutions, many things are still to be improved in the field. The exact relationships between the data collection quality and the resolution of the final structure are still to be known for small proteins, when the strong need of taking images close to focus (to obtain high-resolution information) contradicts with the visibility of particles, making it difficult not only to box the particles but also to align them. The 100kV data collection on a FEG microscope facilitates the close-to-focus data collection, but also decreases the range of defocuses to be possible to collect, because the transfer of high frequencies drops dramatically with the increase of the defocus. Moreover, in comparison with icosahedral-symmetry structures, the non-symmetrical structures are much more demanding in terms of the size and the quality of the dataset.

Nevertheless, the established procedure makes possible the investigation of similar small protein-DNA complexes by single particle cryo-electron microscopy. The study thus moves the generally accepted limit of 200-300kDa ahead by showing the possibility of studying of complexes with a molecular weight around 100kD by cryo-EM. Studies of this type may in the future also be facilitated by more sensitive CCD cameras and possibly by energy filters and aberration correctors. The present work

opens the way to structural studies of a whole range of relatively small macromolecular objects including those of protein / nucleic acid complexes which represent the majority of current structural biology targets but may have been difficult to crystallize or judged too small for standard cryo-EM methods. Thus, the study of NR complexes with different response element types such as inverted, everted and direct repeats can be envisaged in order to address their different, specific architectures.

6. References

1. Agar, A. W., R. H. Anderson and D. Chescoe. Principles and practice of electron microscope operation. Volume 2 of Practical methods in electron microscopy. North Holland, Amsterdam (1974)
2. Akiba, T., Toyoshima, C., Matsunaga, T., Kawamoto, M., Kubota, T., Fukuyama, K., Namba, K. & Matsubara, H. Three-dimensional structure of bovine cytochrome bc1 complex by electron cryomicroscopy and helical image reconstruction. Nature Struct. Biol. 3, 553-561 (1996)
3. Alland L, Muhle R, Hou H Jr, Potes J, Chin L, Schreiber-Agus N, DePinho RA. Role for N-CoR and histone deacetylase in Sin3-mediated transcriptional repression. Nature 387(6628), 49-55 (1997)
4. Aranda A, Pascual A. "Nuclear hormone receptors and gene expression". Physiol. Rev. 81 (3): 1269–304, (2001)
5. Franz Aurenhammer. Voronoi Diagrams - A Survey of a Fundamental Geometric Data Structure. ACM Computing Surveys, 23(3): 345-405, (1991)
6. Baker, T. S. & Cheng, H. A model-based approach for determining orientations of biological macromolecules imaged by cryoelectron microscopy. J. Struct. Biol. 116: 120-130, (1996)
7. Baker TS, Olson NH, Fuller SD. Adding the third dimension to virus life cycles: three-dimensional reconstruction of icosahedral viruses from cryo-electron micrographs. Microbiol Mol Biol Rev. 63(4): 862-922, (1999)
8. Bannister AJ, Kouzarides T. The CBP co-activator is a histone acetyltransferase. Nature 384(6610): 641-3, (1996)

9. Billas IM, Iwema T, Garnier JM, Mitschler A, Rochel N, Moras D. Structural adaptability in the ligand-binding pocket of the ecdysone hormone receptor. Nature 426: 91-6, (2003)

10. Benoit G, Cooney A, Giguere V, Ingraham H, Lazar M, Muscat G, Perlmann T, Renaud JP, Schwabe J, Sladek F, Tsai MJ, Laudet V. "International Union of Pharmacology. LXVI. Orphan nuclear receptors". Pharmacol. Rev. 58 (4): 798–836. doi:10.1124/pr.58.4.10, (2006)

11. Björnström L, Sjöberg M. "Estrogen receptor-dependent activation of AP-1 via non-genomic signalling". Nucl Recept 2 (1): 3. doi:10.1186/1478-1336-2-3, (2004)

12. Borland, L. and van Heel, M. Classification of image data in conjugate representation spaces. J. Opt. Soc. Am. A. 7, pp. 601–610, (1990)

13. Bourguet W, Vivat V, Wurtz JM, Chambon P, Gronemeyer H, Moras D. Crystal structure of a heterodimeric complex of RAR and RXR ligand-binding domains. Mol Cell. 5(2), 289-298, (2000)

14. Brelivet Y, Kammerer S, Rochel N, Poch O, Moras D. Signature of the oligomeric behaviour of nuclear receptors at the sequence and structural level. EMBO Rep. 5(4), 423-429, (2004)

15. Brenner, S. and Horne, R.W. A negative staining method for high resolution electron microscopy of viruses. Biochim. Biophys. Acta. 34, 103-110, (1959)

16. Brzozowski AM, Pike AC, Dauter Z, Hubbard RE, Bonn T, Engström O, Öhman L, Greene GL, Gustafsson J-Å, Carlquist M. "Molecular basis of agonism and antagonism in the oestrogen receptor". Nature 389 (6652): 753–758, (1997)

17. Busch BB, Stevens WC, Martin R, et al. "Identification of a selective inverse agonist for the orphan nuclear receptor estrogen-related receptor alpha". J. Med. Chem. 47 (23): 5593–6, (2004)

18. Chandra V, Huang P, Hamuro Y, Raghuram S, Wang Y, Burris TP, Rastinejad F. Structure of the intact PPAR-gamma-RXR- nuclear receptor complex on DNA. Nature 456, 350-356, (2008)

19. Chiu, W. What does electron cryomicroscopy provide that X-ray crystallography and NMR spectroscopy cannot? Annu. Rev. Biophys. Biomol. Struct. 22, 233-255, (1993)

20. Cormack A.M. Reconstruction of densities from their projections, with applications in radiological physics, Phys. Med. Biol. 18 (1973), pp. 195–207, (1973)

21. Cramer, H. Mathematical Methods of Statistics. Princeton University Press. (1946)

22. Crowther, R.A., Amos, L.A., Finch, J.T., DeRosier, D.J., and Klug, A. Three dimensional reconstructions of spherical viruses by Fourier synthesis from electron micrographs. Nature 226, 421-425, (1970a)

23. Crowther, R.A., DeRosier, D.J., and A.Klug The reconstruction of a three-dimensional structure from projections and its application to electron microscopy. Proc. Roy. Soc. Lond. A. 317, 319-340, (1970b)

24. Crowther, R. A. Procedures for three-dimensional reconstruction of spherical viruses by Fourier synthesis from electron micrographs. Phil. Trans. R. Soc. Lond. Ser. B 261, 221–230, (1971a)

25. Crowther, R.A. Three-dimensional reconstruction and the architecture of spherical viruses. Endeavour 30, 124–129, (1971b)

26. B. Delaunay: Sur la sphère vide, Izvestia Akademii Nauk SSSR, Otdelenie Matematicheskikh i Estestvennykh Nauk, 7:793-800, (1934)

27. DeRosier and Klug, 1968 D.J. DeRosier and A. Klug, Reconstruction of three-dimensional structures from electron micrographs, Nature 217: 130–134, (1968)

28. Gustav Lejeune Dirichlet Über die Reduktion der positiven quadratischen Formen mit drei unbestimmten ganzen Zahlen. Journal für die Reine und Angewandte Mathematik, 40:209-227, (1850)

29. Dube, P., Tavares, P., Lurz, R. & van Heel, M. The portal protein of bacteriophage SPP1: a DNA pump with 13-fold symmetry. EMBO J. 12, 1303-1309, (1993)

30. Dubochet, J., Adrian, M., Chang, J.J, Homo, J.C., Lepault, J., McDowall, A.W. and Schultz, P. Cryo-electron microscopy of vitrified specimens. Q. Rev. Biophys. 21, pp. 129–228. Full Text via CrossRef | View Record in Scopus | Cited By in Scopus (731), (1988)

31. Dubochet, J., Lepault, J., Freeman, R., Berriman, J.A. and Homo, J.C. Electron microscopy of frozen water and aqueous solutions. J. Microsc. 128, pp. 219–237, (1982)

32. Egea PF, Mitschler A, Moras D. Molecular recognition of agonist ligands by RXRs. Mol Endocrinol. 5:987-97, (2002)

33. Egea PF, Mitschler A, Rochel N, Ruff M, Chambon P, Moras D. Crystal structure of the human RXRalpha ligand-binding domain bound to its natural ligand: 9-cis retinoic acid. EMBO J. Jun 1;19(11):2592-601, (2000)

34. Egger and Morel, M.L. Egger and C. Morel, Execution times of five reconstruction algorithms in 3D positron emission tomography, Phys. Med. Biol. 43: 703–712. (1998)

35. Elad N, Clare DK, Saibil HR, Orlova EV. Detection and separation of heterogeneity in molecular complexes by statistical analysis of their two-dimensional projections. J Struct Biol. Apr;162(1): 108-20, (2008)

36. Erickson, H.P. and Klug, A. Measurement and compensation of defocusing and aberrations by Fourier processing of electron micrographs. Phil. Trans. Roy. Soc. Lond. B 261, pp. 105–118, (1971)

37. Escriva H, Langlois MC, Mendonça RL, Pierce R, Laudet V. "Evolution and diversification of the nuclear receptor superfamily". Annals of the New York Academy of Sciences 839: 143–6. doi:10.1111/j.1749-6632.1998.tb10747.x, (1998)

38. Evans RM. The steroid and thyroid hormone receptor superfamily. Science 240 (4854): 889–95. doi:10.1126/science.3283939, (1988)

39. Fiévet C, Staels B. Efficacy of peroxisome proliferator-activated receptor agonists in diabetes and coronary artery disease. Curr. Atheroscler Rep. 11(4), 281-288, (2009).

40. Frank, J. Three-Dimensional Electron Microscopy of Macro-molecular Assemblies. San Diego, London: Academic Press, (1996).

41. Frank, Three-dimensional electron microscopy of macromolecular assemblies. Oxford University Press, (2006)

42. Frank J. and van Heel, M. Correspondence analysis of aligned images of biological particles. J. Mol. Biol. 161, pp. 134–137, (1982)

43. Frank J. The role of multivariate image analysis in solving the architecture of the Limulus polyphemus hemocyanin molecule. Ultramicroscopy 13, pp. 153–164. (1984)

44. Frank J., Radermacher M., Wagenknecht T. & Verschoor A. Studying ribosome structure by electron microscopy and computer-image processing. Methods Enzymol. 164, 3-35. (1988)

45. Frank, J., Verschoor, A., Wagenknecht, T., Radermacher, M. and Carazo, J.-M., A new non-crystallographic image-processing technique reveals the architecture of ribosomes. Trends Bioch. Sci. 13, pp. 123–127. (1988)

46. Frank J. Classification of macromolecular assemblies studied as ‘single particles’. Q. Rev. Biophys. 23, pp. 281–329. (1990)

47. Frank J. The role of correlation techniques in computer image processing. In: Hawkes, P.W., Editor, , 1980. Computer Processing of Electron Microscope Images, Springer, Berlin, pp. 187–222. (1980)

48. Friedrich M, Axt-Fliedner R, Villena-Heinsen C, Tilgen W, Schmidt W, Reichrath J. Analysis of vitamin D-receptor (VDR) and retinoid X-receptor alpha in breast cancer. Histochem J. 34(1-2), 35-40 (2002)

49. Gabashvili IS, Agrawal RK, Spahn CM, Grassucci RA, Svergun DI, Frank J, Penczek P. Solution structure of the E. coli 70S ribosome at 11.5 A resolution. Cell. Mar 3;100(5):537-49. (2000)

50. Gampe RT Jr, Montana VG, Lambert MH, Miller AB, Bledsoe RK, Milburn MV, Kliewer SA, Willson TM, Xu HE. Asymmetry in the PPARgamma/RXRalpha crystal structure reveals the molecular basis of heterodimerization among nuclear receptors. Mol Cell 5(3), 545-555 (2000)

51. Giguere V. Orphan nuclear receptors: from gene to function. Endocr Rev 20: 689-725, (1999)

52. Glaeser R. M. Limitations to significant information in biological electron microscopy as a result of radiation damage. J. Ultrastruc. Res. 36, 466-482. (1971)

53. Glass CK, Ogawa S. Combinatorial roles of nuclear receptors in inflammation and immunity. Nat Rev Immunol. Jan;6(1):44-55. Review. (2006)

54. Glass CK. Differential recognition of target genes by nuclear receptor monomers, dimers, and heterodimers. Endocr Rev. Jun;15(3):391-407. (1994)

55. Glass CK, Rosenfeld MG The coregulator exchange in transcriptional functions of nuclear receptors. Genes Dev. 14(2), 121-141 (2000).

56. Gordon, R., Bender, R., Herman, G., Algebraic reconstruction techniques (ART) for three dimensional electron microscopy and X-ray photography. J. Theor. Biol. 36, 105–117. (1970)

57. Greschik H, Wurtz JM, Sanglier S, Bourguet W, van Dorsselaer A, Moras D, Renaud JP. Structural and functional evidence -for ligand-independent transcriptional activation by the estrogen-related receptor 3. Mol Cell. 2, 303-13. (2002)

58. Grigorieff N. Resolution measurement in structures derived from single particles. Acta Crystallogr D Biol Crystallogr. Oct;56(Pt 10):1270-7. (2000)

59. Gronemeyer H, Moras D. Nuclear receptors. How to finger DNA. Nature. May 18;375(6528):190-1. (1995)

60. Hinrich Gronemeyer, Jan-Åke Gustafsson & Vincent Laudet, Principles for modulation of the nuclear receptor superfamily, Nature Reviews Drug Discovery 3, 950-964 (2004)

61. HALL CE. Electron densitometry of stained virus particles. J Biophys Biochem Cytol. Jan;1(1):1-12. (1955)

62. G. Harauz and M. van Heel, Exact filters for general geometry three-dimensional reconstruction, Optik 73, pp. 146–156. (1986)

63. Harris, J. R. & Horne, R. W. Electron Microscopy in Biology, edited by J. R. Harris, pp. 203-228. Oxford: IRL Press. (1991)

64. Hawkes, P. W. & Valdrè, U. (1990). Biophysical Electron Microscopy. London: Academic Press. (1990)

65. Hawkes, P.W. Electron Optics and Electron Microscopy, Taylor and Francis: London. (1972)

66. Hawkes, P. W. In Electron Tomography: Three-Dimensional Imaging with the Transmission Electron Microscope, edited by J. Frank. New York, London: Plenum Press. (1992)

67. Hassig CA, Fleischer TC, Billin AN, Schreiber SL, Ayer DE. Histone deacetylase activity is required for full transcriptional repression by mSin3A. Cell. May 2;89(3):341-7. (1997)

68. M. A. Hayat Principles and techniques of electron microscopy: biological applications. Cambridge University Press. pp. 45–61 (2000)

69. van Heel, M. Angular reconstitution: a posteriori assignment of projection directions for 3D reconstruction. Ultramicroscopy 21, pp. 111–124. (1987)

70. van Heel M, Gowen B, Matadeen R, Orlova EV, Finn R, Pape T, Cohen D, Stark H, Schmidt R, Schatz M, Patwardhan A. Single-particle electron cryo-microscopy: towards atomic resolution. Q Rev Biophys. Nov;33(4):307-69. (2000)

71. van Heel, M. and Frank, J., Use of multivariate statistics in analysing the images of biological macromolecules. Ultramicroscopy 6, pp. 187–194. (1986)

72. van Heel, M. and Harauz, G., Resolution criteria for three dimensional reconstruction. Optik 73, pp. 119–122. (1996)

73. van Heel, M., Harauz, G., Orlova, E.V., Schmidt, R., Schatz, M. A new generation of the IMAGIC image processing system. J. Struct. Biol. 116, 17-24 (1996).

74. van Heel, M. & Schatz, M. Fourier shell correlation threshold criteria. J. Struct. Biol. 151, 250-262 (2005).

75. Heinzel T, Lavinsky RM, Mullen TM, Söderstrom M, Laherty CD, Torchia J, Yang WM, Brard G, Ngo SD, Davie JR, Seto E, Eisenman RN, Rose DW, Glass CK, Rosenfeld MG. A complex containing N-CoR, mSin3 and histone deacetylase mediates transcriptional repression. Nature. 1997 May 1;387(6628):43-8. (1997)

76. Herman, G. T. Image Reconstruction from Projections: the Fundamentals of Computerized Tomography. New York: Academic Press. (1980)

77. Herman and Lent, 1976 G. Herman and A. Lent, Iterative reconstruction algorithms, Comput. Biol. Med. 6, pp. 273–294. (1976)

78. Hoffman, D. W., Cameron, C. S., Davies, C., White, S. & Ramakrishnan, V. Ribosomal protein L9: a structure determination by the combined use of X-ray crystallography and NMR spectroscopy. J. Mol. Biol. 264, 1058-1071 (1996)

79. Hoppe, W. & Hegerl, R. Computer Proccesing of Electron Microscope Images, edited by P. W. Hawkes, pp. 127-185. Berlin, New York: Springer-Verlag. (1980)

80. HUXLEY HE. Some aspects of staining of tissue for sectioning. J R Microsc Soc. Mar-Jun;78:30-2. (1958)

81. Jiang W, Baker ML, Jakana J, Weigele PR, King J, Chiu W. Backbone structure of the infectious epsilon15 virus capsid revealed by electron cryomicroscopy. Nature. Feb 28;451(7182):1130-4. (2008)

82. Jiang and Wang, M. Jiang and G. Wang, Convergence of the simultaneous algebraic reconstruction technique (SART), IEEE Trans. Image Process. 12, pp. 957–961. (2003)

83. Kadosh D, Struhl K. Repression by Ume6 involves recruitment of a complex containing Sin3 corepressor and Rpd3 histone deacetylase to target promoters. Cell. May 2;89(3):365-71. (1997)

84. Kato S, Sasaki H, Suzawa M, Masushige S, Tora L, Chambon P, and Gronemeyer H. Widely spaced, directly repeated PuGGTCA elements act as promiscuous enhancers for different classes of nuclear receptors. Mol Cell Biol 15: 5858-5867, (1995)

85. Kato T, Goodman RP, Erben CM, Turberfield AJ, Namba K. High-resolution structural analysis of a DNA nanostructure by cryoEM. Nano Lett. Jul;9(7):2747-50. (2009)

86. Klaholz BP, Mitschler A, Belema M, Zusi C, Moras D. Enantiomer discrimination illustrated by high-resolution crystal structures of the human nuclear receptor hRARgamma. Proc Natl Acad Sci U S A. 97, 6322-7. (2000)

87. Klaholz, B. P., Myasnikov, A. G., & Van Heel, M. Visualization of release factor 3 on the ribosome during termination of protein synthesis. Nature 427, 862-865 (2004).

88. Klaholz BP, Renaud JP, Mitschler A, Zusi C, Chambon P, Gronemeyer H, Moras D. Conformational adaptation of agonists to the human nuclear receptor RAR gamma. Nat Struct Biol. Mar;5(3):199-202. (1998)

89. Klinge CM. "Estrogen receptor interaction with co-activators and co-repressors". Steroids 65 (5): 227–51. doi:10.1016/S0039-128X(99)00107-5. PMID 10751636. (2000)

90. Klinge CM, Bodenner DL, Desai D, Niles RM, Traish AM Binding of type II nuclear receptors and estrogen receptor to full and half-site estrogen response elements in vitro. Nucleic Acids Res. 25 (10): 1903–12. (1997)

91. Laherty CD, Yang WM, Sun JM, Davie JR, Seto E, Eisenman RN. Histone deacetylases associated with the mSin3 corepressor mediate mad transcriptional repression. Cell. May 2;89(3):349-56. (1997)

92. Laudet V. Evolution of the nuclear receptor superfamily: early diversification from an ancestral orphan receptor. J Mol Endocrinol 19: 207-226, (1997)

93. Lebart, L., Morineau, A., Warwick, K.M., Multivariate Descriptive Statistical Analysis. Wiley, New York, pp. 30–62. (1984)

94. Lee MS, Kliewer SA, Provencal J, Wright PE, and Evans RM. Structure of the retinoid X receptor alpha DNA binding domain: a helix required for homodimeric DNA binding. Science 260: 1117-1121, (1993)

95. Linja MJ, Porkka KP, Kang Z, et al. "Expression of androgen receptor coregulators in prostate cancer". Clin. Cancer Res. 10 (3): 1032–40. (2004)

96. Lonard DM, Lanz RB, O'Malley BW. Nuclear receptor coregulators and human disease. Endocr Rev. Aug;28(5):575-87. Epub 2007 Jul 3. (2007)

97. Ludtke, S.J., Baldwin, P.R., and Chiu, W. EMAN: semiautomated software for high-resolution single-particle reconstructions. J. Struct. Biol. 128, 82-97 (1999).

98. Luisi BF, Xu WX, Otwinowski Z, Freedman LP, Yamamoto KR, Sigler PB. Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. Nature. Aug 8;352(6335):497-505. (1991)

99. Mallick SP, Carragher B, Potter CS, Kriegman DJ. ACE: automated CTF estimation. Ultramicroscopy. Aug;104(1):8-29. Epub 2005 Mar 23. (2005)

100. Mangelsdorf David J. and Evans Ronald M., The RXR Heterodimers and Orphan Receptors. Cell 83, 841-850, (1995).

101. Mangelsdorf DJ, Thummel C, Beato M, Herrlich P, Schutz G, Umesono K, Blumberg B, Kastner P, Mark M, Chambon P, Evans RM. "The nuclear

receptor superfamily: the second decade". Cell 83 (6): 835–9. doi:10.1016/0092-8674(95)90199-X. PMID 8521507. (1995)

102. Marzi, S. et al. Structured mRNAs regulate translation initiation by binding to the platform of the ribosome. Cell 130, 1019–1031 (2007).

103. McKenna,N.J., Lanz,R.B. & O'Malley,B.W. Nuclear receptor coregulators: cellular and molecular biology. Endocr. Rev. 20, 321-344 (1999).

104. Mueller F, Sommer I, Baranov P, Matadeen R, Stoldt M, Wöhnert J, Görlach M, van Heel M, Brimacombe R. The 3D arrangement of the 23 S and 5 S rRNA in the Escherichia coli 50 S ribosomal subunit based on a cryo-electron microscopic reconstruction at 7.5 A resolution. J Mol Biol. Apr 21;298(1):35-59. (2000)

105. Moras D, Gronemeyer H. The nuclear receptor ligand-binding domain: structure and function. Curr Opin Cell Biol. Jun;10(3):384-91. (1998)

106. Naar AM, Boutin JM, Lipkin SM, Yu VC, Holloway JM, Glass CK, and Rosenfeld MG. The orientation and spacing of core DNA-binding motifs dictate selective transcriptional responses to three nuclear receptors. Cell 65: 1267-1279, (1991)

107. Nagy L, Kao HY, Chakravarti D, Lin RJ, Hassig CA, Ayer DE, Schreiber SL, Evans RM. Nuclear receptor repression mediated by a complex containing SMRT, mSin3A, and histone deacetylase. Cell. May 2;89(3):373-80. (1997)

108. Novac N, Heinzel T Nuclear receptors: overview and classification. Curr Drug Targets Inflamm Allergy 3 (4): 335–46. doi:10.2174/1568010042634541. (2004)

109. Ogryzko VV, Schiltz RL, Russanova V, Howard BH, Nakatani Y. The transcriptional coactivators p300 and CBP are histone acetyltransferases. Cell. Nov 29;87(5):953-9 (1996)

110. Olefsky JM. Nuclear receptor minireview series. J. Biol. Chem. 276 (40): 36863–4. doi:10.1074/jbc.R100047200. PMID 11459855. (2001)

111. Orlova EV, Dube P, Harris JR, Beckman E, Zemlin F, Markl J, van Heel M. Structure of keyhole limpet hemocyanin type 1 (KLH1) at 15 A resolution by electron cryomicroscopy and angular reconstitution. J Mol Biol. Aug 22;271(3):417-37 (1997)

112. Orlova, E. V. & van Heel, M. (1994). 13th International Congress of Electron Microscopy, Vol. 1, edited by B. Jouffrey & C. Collix (1994)

113. Pazin MJ, Kadonaga JT. What's up and down with histone deacetylation and transcription? Cell. May 2;89(3):325-8 (1997)

114. Penczek PA, Grassucci RA, Frank J. The ribosome at improved resolution: new techniques for merging and orientation refinement in 3D cryo-electron microscopy of biological particles. Ultramicroscopy. Mar;53(3):251-70. (1994)

115. Penczek PA, Renka R, Schomberg H. Gridding-based direct Fourier inversion of the three-dimensional ray transform. J Opt Soc Am A Opt Image Sci Vis. Apr;21(4):499-509. (2004)

116. Pinette KV, Yee YK, Amegadzie BY, Nagpal S. Vitamin D receptor as a drug discovery target. Mini Rev Med Chem. 3(3), 193-204 (2003)

117. Pike AC, Brzozowski AM, Hubbard RE, Bonn T, Thorsell AG, Engström O, Ljunggren J, Gustafsson JA, Carlquist M. Structure of the ligand-binding domain of oestrogen receptor beta in the presence of a partial agonist and a full antagonist. EMBO J. Sep 1;18(17):4608-18 (1999)

118. Radermacher, M. Three-dimensional reconstruction of single particles from random and nonrandom tilt series. J. Electr. Microsc. Tech. 9, 359-394. (1988)

119. Radermacher M, Wagenknecht T, Verschoor A, Frank J. Three-dimensional reconstruction from a single-exposure, random conical tilt series applied to the

50S ribosomal subunit of Escherichia coli. J Microsc. May;146(Pt 2):113-36 (1987)

120. Radon, J. Math. Phys. Klasse, 69, 262-277. (1917)

121. Radon, J. "Mengen konvexer Körper, die einen gemeinsamen Punkt enthalten", Mathematische Annalen 83 (1–2): 113–115 (1921)

122. Rastinejad F, Perlmann T, Evans RM, Sigler PB. Structural determinants of nuclear receptor assembly on DNA direct repeats. Nature 375, 190-191 (1995)

123. Rastinejad F, Wagner T, Zhao Q, Khorasanizadeh S. Structure of the RXR-RAR DNA-binding complex on the retinoic acid response element DR1. EMBO J. 19(5), 1045-1054 (2000). Rastinejad F, Wagner T, Zhao Q, Khorasanizadeh S. Structure of the RXR-RAR DNA-binding complex on the retinoic acid response element DR1. EMBO J. 19(5), 1045-1054 (2000)

124. Reimer, L. Transmission Electron Microscopy. Springer-Verlag, Berlin. (1989)

125. Renaud JP, Rochel N, Ruff M, Vivat V, Chambon P, Gronemeyer H, Moras D. Crystal structure of the RAR-gamma ligand-binding domain bound to all-trans retinoic acid. Nature. Dec 14;378(6558):681-9. (1995)

126. Rochel N, Wurtz JM, Mitschler A, Klaholz B, Moras D. The crystal structure of the nuclear receptor for vitamin D bound to its natural ligand. Mol. Cell 5(1), 173-179 (2000)

127. Rochel N, Wurtz JM, Mitschler A, Klaholz B, Moras D. The crystal structure of the nuclear receptor for vitamin D bound to its natural ligand. Mol Cell. Jan;5(1):173-9. (2000)

128. Rochette-Egly C, Adam S, Rossignol M, Egly JM, and Chambon P. Stimulation of RAR alpha activation function AF-1 through binding to the general transcription factor TFIIH and phosphorylation by CDK7. Cell 90: 97-107, (1997)

129. Rochette-Egly C, Gaub MP, Lutz Y, Ali S, Scheuer I, and Chambon P. Retinoic acid receptor-beta: immunodetection and phosphorylation on tyrosine residues. Mol Endocrinol 6: 2197-2209, (1992)

130. Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. J. Mol. Biol. 333, 721-745 (2003)

131. Jonathan Ruprecht and Jon Nield, Determining the structure of biological macromolecules by transmission electron microscopy, single particle analysis and 3D reconstruction, Progress in Biophysics and Molecular Biology Volume 75, Issue 3, Pages 121-164 (2001)

132. Sander B, Golas MM, Stark H. Automatic CTF correction for single particles based upon multivariate statistical analysis of individual power spectra. J Struct Biol. Jun;142(3):392-401 (2003)

133. Saxton, W. O. Computer Techniques for Image Processing in Electron Microscopy, p. 289. New York, San Francisco, London: Academic Press (1978)

134. Saxton WO, Baumeister W. The correlation averaging of a regularly arranged bacterial cell envelope protein. J. Microsc. 127(2), 127-138 (1982).

135. Saibil HR. Macromolecular structure determination by cryo-electron microscopy. Acta Crystallogr D Biol Crystallogr. Oct;56(Pt 10):1215-22. (2000)

136. Schatz M, Orlova EV, Dube P, Jäger J, van Heel M. Structure of Lumbricus terrestris hemoglobin at 30 A resolution determined using angular reconstitution. J Struct Biol. Jan-Feb;114(1):28-40. (1995)

137. Schwabe JW, Chapman L, Finch JT, Rhodes D, Neuhaus D. DNA recognition by the oestrogen receptor: from solution to the crystal. Structure. Nov 15;1(3):187-204. (1993)

138. Serysheva II, Orlova EV, Chiu W, Sherman MB, Hamilton SL, van Heel M. Electron cryomicroscopy and angular reconstitution used to visualize the skeletal muscle calcium release channel. Nat Struct Biol. Jan;2(1):18-24. (1995)

139. Sezan and Stark, 1982 M.I. Sezan and H. Stark, Image restoration by the method of convex projections. II Applications and numerical results, IEEE Trans. Med. Imaging. 1, pp. 95–101. (1982)

140. Shaffer, P. L., Donald, P., McDonnell, P., Gewirth, D. T. Characterization of Transcriptional Activation and DNA-Binding Functions in the Hinge Region of the Vitamin D Receptor. Biochemistry 44, 2678-2685 (2005).

141. Shaffer PL, Gewirth DT. Structural analysis of RXR-VDR interactions on DR3 DNA. J. Steroid Biochem. Mol. Biol. 89-90(1-5), 215-219 (2004)

142. Shaffer PL, Gewirth DT. Structural basis of VDR-DNA interactions on direct repeat response elements. EMBO J. 21(9), 2242-2252 (2002)

143. Shao D, and Lazar MA. Modulating nuclear receptor function: may the phos be with you. J Clin Invest 103: 1617-1618, (1999)

144. Shiau AK, Barstad D, Loria PM, Cheng L, Kushner PJ, Agard DA, Greene GL The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. Cell 95 (7): 927–37 (1998)

145. Sherman, M.B., Soejima, T., Chiu, W. and van Heel, M.. Multivariate analysis of single unit cells in electron crystallography. Ultramicroscopy 74, pp. 179–199 (1998)

146. L.A. Shepp and B.F. Logan, The Fourier reconstruction of a head section, IEEE Trans. Nucl. Sci. 21, pp. 21–43. (1974)

147. Simonetti A, Marzi S, Myasnikov AG, Fabbretti A, Yusupov M, Gualerzi CO, Klaholz BP. Structure of the 30S translation initiation complex. Nature 455, 416-420 (2008)

148. Slayter E., Slayter H. Light and Electron Microscopy. Cambridge University Press. (1993)

149. Sluder AE, Maina CV. "Nuclear receptors in nematodes: themes and variations". Trends in Genetics : TIG 17 (4): 206–13. (2001)

150. Stark, H., Rodnina., M. V., Wieden, H.-J., van Heel, M. & Wintermeyer, W. Large-scale movement of elongation factor G and extensive conformational change of the ribosome during translocation. Cell, 100, 301-309 (2000)

151. Stark H, Orlova EV, Rinke-Appel J, Jünke N, Mueller F, Rodnina M, Wintermeyer W, Brimacombe R, van Heel M. Arrangement of tRNAs in pre- and posttranslocational ribosomes revealed by electron cryomicroscopy. Cell. Jan 10;88(1):19-28 (1997)

152. Stewart, M. (1990), Electron microscopy of biological macromolecules. In: “Modern Microscopies” Eds: Duke P.J. and Michette A.G., Plenum, New-York.9-40. (1990)

153. Svensson S, Ostberg T, Jacobsson M, Norström C, Stefansson K, Hallén D, Johansson IC, Zachrisson K, Ogg D, Jendeberg L. Crystal structure of the heterodimeric complex of LXRalpha and RXRbeta ligand-binding domains in a fully agonistic conformation. EMBO J. 22(18), 4625-4633 (2003).

154. Taneja R, Rochette-Egly C, Plassat JL, Penna L, Gaub MP, and Chambon P. Phosphorylation of activation functions AF-1 and AF-2 of RAR alpha and RAR gamma is indispensable for differentiation of F9 cells upon retinoic acid and cAMP treatment. EMBO J 16: 6452-6465, (1997)

155. Tata JR. "One hundred years of hormones". EMBO Rep. 6 (6): 490–6. doi:10.1038/sj.embor.7400444. PMID 15940278. (2005)

156. Tocchini-Valentini G, Rochel N, Wurtz JM, Moras D. Crystal structures of the vitamin D nuclear receptor liganded with the vitamin D side chain analogues calcipotriol and seocalcitol, receptor agonists of clinical importance. Insights into a structural basis for the switching of calcipotriol to a receptor antagonist by further side chain modification. J Med Chem., 47, 1956-61. (2004)

157. Umesono K, Murakami KK, Thompson CC, and Evans RM. Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D3 receptors. Cell 65: 1255-1266, (1991)

158. Unser, M., Trus, R. L., Vrehl, M. J., Conway, J. F., Gross, M., Thevenaz, R. & Steven, A. C. Resolution Assessment of Three-Dimensional Reconstructions by Spectral Signal-to-Noise Ratio. InAbstracts of the 11th European Congress on Electron Microscopy. Dublin: U. C. D. Belfield. (1996)

159. Velázquez-Muriel JA, Sorzano CO, Fernández JJ, Carazo JM. A method for estimating the CTF in electron microscopy based on ARMA models and parameter adjustment. Ultramicroscopy. Jul;96(1):17-35. (2003)

160. Georgy Voronoi Nouvelles applications des paramètres continus à la théorie des formes quadratiques. Journal für die Reine und Angewandte Mathematik, 133:97-178, (1907)

161. Wade, R. H. The phase contrast characteristics in bright field electron microscopy. Ultramicroscopy, 46, 145-156. (1992)

162. Wang L, Mizzen C, Ying C, Candau R, Barlev N, Brownell J, Allis CD, Berger SL. Histone acetyltransferase activity is conserved between yeast and human GCN5 and is required for complementation of growth and transcriptional activation. Mol Cell Biol. Jan;17(1):519-27. (1997)

163. David F. Watson Computing the n-dimensional tessellation with application to Voronoi polytopes, The Computer Journal, Heyden & Sons Ltd., Vol 2, Num 24, pp.167-172. (1981)

164. Wagner RL, Apriletti JW, McGrath ME, West BL, Baxter JD, Fletterick RJ. A structural role for hormone in the thyroid hormone receptor. Nature Dec 14;378(6558):690-7. (1995)

165. Wärnmark A, Treuter E, Wright AP, Gustafsson J-Å Activation functions 1 and 2 of nuclear receptors: molecular strategies for transcriptional activation. Mol. Endocrinol. 17 (10): 1901–9. doi:10.1210/me.2002-0384. PMID 12893880. (2003)

166. Williams SP, Sigler PB. Atomic structure of progesterone complexed with its receptor. Nature. May 28;393(6683):392-6. (1998)

167. Wortham M, Czerwinski M, He L, Parkinson A, Wan YJ Expression of constitutive androstane receptor, hepatic nuclear factor 4 alpha, and P450 oxidoreductase genes determines interindividual variability in basal expression and activity of a broad scope of xenobiotic metabolism genes in the human liver. Drug Metab. Dispos. 35 (9): 1700–10. doi:10.1124/dmd.107.016436. PMID 17576804. (2007)

168. Wu W, Niles EG, El-Sayed N, Berriman M, LoVerde PT. Schistosoma mansoni (Platyhelminthes, Trematoda) nuclear receptors: sixteen new members and a novel subfamily. Gene 366 (2): 303–15. doi:10.1016/j.gene.2005.09.013. PMID 16406405 (2006)

169. Wu W, Niles EG, Hirai H, LoVerde PT Evolution of a novel subfamily of nuclear receptors with members that each contain two DNA binding domains. BMC Evol Biol 7 (Feb 23): 27. doi:10.1186/1471-2148-7-27. PMID 17319953. (2004)

170. Yang XJ, Ogryzko VV, Nishikawa J, Howard BH, Nakatani Y. A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. Nature. Jul 25;382(6589):319-24 (1996)

171. Yu X, Jin L, Zhou ZH. 3.88 A structure of cytoplasmic polyhedrosis virus by cryo-electron microscopy. Nature. May 15;453(7193):415-9. (2008)

172. Zhang Z, Burch PE, Cooney AJ, Lanz RB, Pereira FA, Wu J, Gibbs RA, Weinstock G, Wheeler DA "Genomic analysis of the nuclear receptor family: new insights into structure, regulation, and evolution from the rat genome". Genome Res 14 (4): 580–90. doi:10.1101/gr.2160004. (2004)

173. Zhang Y, Iratni R, Erdjument-Bromage H, Tempst P, Reinberg D. Histone deacetylases and SAP18, a novel polypeptide, are components of a human Sin3 complex. Cell. May 2;89(3):357-64. (1997)

174. Zhao Q, Chasse SA, Devarakonda S, Sierk ML, Ahvazi B, Rastinejad F. Structural basis of RXR-DNA interactions. J Mol Biol. 296(2), 509-520 (2000).

175. Zivadinovic D, Gametchu B, Watson CS "Membrane estrogen receptor-alpha levels in MCF-7 breast cancer cells predict cAMP and proliferation responses". Breast Cancer Res. 7 (1): R101–12. (2005)

Publications

1. I.M. Orlov, D.G. Morgan and and R.H.Cheng, 2005. Efficient implementation of filtered back projection algorithm on voxel-by-voxel base. Journal of Structural Biology, Volume 154, Issue 3, June 2006, Pages 287-296.
2. I.M. Orlov, N. Rochel, D. Moras, B. P. Klaholz, Cryo-electron microscopy structure of the 100kDa full nuclear receptor RXR/VDR heterodimer complex with its target DNA. *Submitted to Nature.*

Oral Presentations

1. I.M. Orlov, A.G. Myasnikov and B. P. Klaholz, Implementation of a fast weighted back-projection 3D reconstruction for a single particle cryo-EM. The workshop "New algorithms in macromolecular crystallography and electron microscopy", Leiden, May 13-17, 2008.
2. I.M. Orlov, A.G. Myasnikov and B. P. Klaholz, 3D reconstruction software with compensation of preferable orientation suitable for any point symmetry group. GRC Conference on Three Dimensional Electron Microscopy, Grenoble, Dec. 3-4, 2007.

Posters

1. I.M. Orlov, N. Rochel, D. Moras and B. P. Klaholz, Cryo-electron microscopyv structure of the 100kDa full-length nuclear receptor RXR/VDR heterodimer complex with its target DNA. GRC Conference on Three Dimensional Electron Microscopy, 2010, June 20-25, Il Ciocco Hotel and Resort, Lucca (Barga), Italy.
2. I.M. Orlov, N. Rochel, D. Moras and B. P. Klaholz, Cryo-electron microscopyv structure of the 100kDa full-length nuclear receptor RXR/VDR heterodimer complex with its target DNA. GTBIO, 2009, December 8-11, Institut Pasteur, Paris, France.
3. I.M. Orlov, A.G. Myasnikov and B. P. Klaholz, 3D reconstruction software with compensation of preferable orientation suitable for any point symmetry group. GTBIO, 2009, December 8-11, Institut Pasteur, Paris, France.
4. I.M. Orlov, A.G. Myasnikov and B. P. Klaholz, 3D reconstruction software with compensation of preferable orientation suitable for any point symmetry group. GRC Conference on Three Dimensional Electron Microscopy, 2008, June 15-20, Il Ciocco Hotel and Resort, Lucca (Barga), Italy.
5. I.M. Orlov, A.G. Myasnikov and B. P. Klaholz, Implementation of a fast and highly precise weighted back-projection 3D reconstruction. GRC Conference on Three Dimensional Electron Microscopy, 2006, June 25-30, Il Ciocco Hotel and Resort, Lucca (Barga), Italy.

Implementation of a fast and highly precise weighted back-projection 3D reconstruction

Igor M. Orlov, Alexander G. Myasnikov and Bruno P. Klaholz

Department of Structural Biology and Genomics, IGMBC, 1 rue Laurent Fries, 67404 Illkirch CEDEX, France

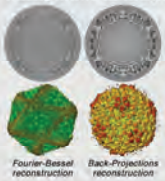
Efficient implementation of a filtered back-projection algorithm using a voxel-by-voxel approach

Igor M. Orlov[1,2], David Gene Morgan[3] and R. Holland Cheng[2,3]

[1] Department de Biologie et Genomique Structurales, IGMBC, 1 rue Laurent Fries, 67404 Illkirch CEDEX, France
[2] Department of Biosciences at NOVUM, Karolinska Institutet,14157 Huddinge, Sweden
[3] Molecular & Cellular Biology, University of California, 1 Shields Avenue, Davis, CA

Background of the project (1)

One of the interesting and unexpected problems revealed during our work on the cryo-EM 3D-reconstruction of the Rice Dwarf Virus was the presence of Moiré-like patterns in the Fourier-Bessel reconstructions which is widely used for icosahedral object reconstruction. The presence of periodical pattern on a Fourier-Bessel reconstruction and overall slowness of traditional back-projection implementations suggested us to develop a new optimised version of the filtered back-projections approach, called the BKPR program.

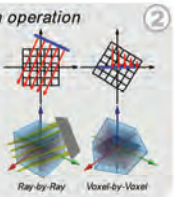


Fourier-Bessel reconstruction — Back-Projections reconstruction

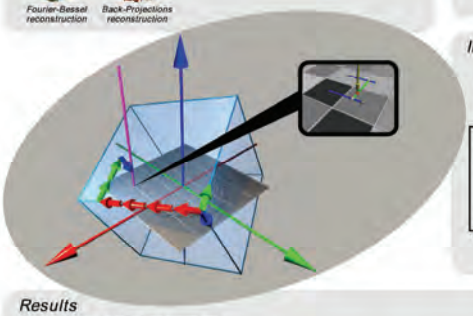
Principles of the backprojection operation (2)

In the ray-by-ray backprojection approach, densities of the image are stretched out (back-projected) through the whole volume of the object being reconstructed in the direction that corresponds to the orientation found for this image. If the direction lines not correspond exactly to one of the coordinate axes of the object matrix, the density of the nodes must be determined by 3D interpolation.

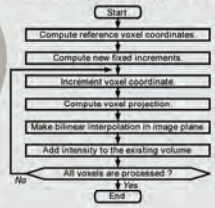
The voxel-by-voxel approach is based on the direct projection of each voxel to the image, that has to be backprojected. In this case the density is determined by bilinear interpolation in the image plane.



Ray-by-Ray — Voxel-by-Voxel



Incremental voxel-by-voxel approach (3)



The orientation angles found for the projection image are used to determine the rotation matrix and unit vectors in the system of coordinates related to the projection. Then fixed axial increments are determined, and used for computation of voxel coordinates. Processing of voxels is performed one by one to compute projections of each voxel on the image plane and the increment of the voxel density will be determined by two-dimensional interpolation of neighboring pixels in the image. The calculated density will be added to the voxel density of the 3D map assigned during the processing of the previous image. The program then proceeds to the next voxel.

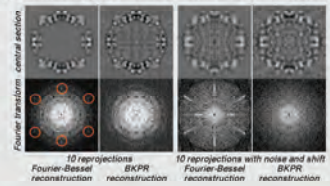
Results (4)

The precision and stability of our program was tested on synthetic and real data.

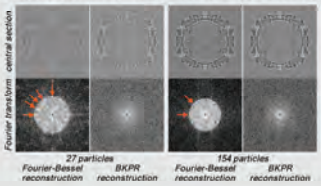
A model was created from an X-ray structure (PDB entry 1A62) by applying 532 symmetry to the atomic co-ordinates using IMAGIC. The final model imitated a virus protein envelope and was placed in the center of a cube with a size of 1293 voxels (1Å/voxel).

As a real data we used cryo-EM images of empty capsids of the Rice Dwarf Virus, recorded on a 300 kV FEG electron microscope.

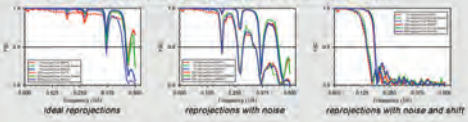
Central sections and Fourier transforms of reconstructions from the simulated 3D density map

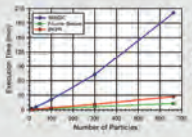


Central sections and Fourier transforms of reconstructions from empty capsids of Rice Dwarf Virus



Fourier Shell Correlations between original synthetic 3D model and different 3D reconstructions





Benchmarks for time estimation of the IMAGIC, Fourier-Bessel and BKPR reconstruction programs were computed using an Athlon MP 2200 processor (RedHat Linux 9.0). Time was estimated for the sets of 10, 30, 100, 300 and 662 reprojections from the model structure using each software package.

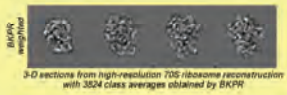
Igor M. Orlov, David Gene Morgan and R. Holland Cheng,
Efficient implementation of a filtered back-projection algorithm using a voxel-by-voxel approach.
Journal of Structural Biology, Volume 154, Issue 3, June 2006, Pages 287-298

Weighted Back-Projections (5)

One of the serious problems in single particle analysis is the non-uniform orientation distribution in ice or on carbon film. This leads to distorted densities in three-dimensional reconstructions because the point spread function is non-spherical. This difficulty can be overcome by weighting of the projections according to their Euler angle distribution (Harauz and van Heel, 1986).

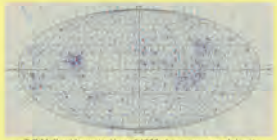
Here we suggest a new approach based on Voronoi diagrams for the determination of exact projection weights. A Voronoi polygon on the Euler sphere identifies a region whose interior comprises all points on the sphere which are closer to a given projection point than to any other, thus determining an "influence area" for each projection. This approach allows generating the quasi-equal density distribution so that the area of each Voronoi polygon on a sphere defines the weight for the corresponding projection.

(5A)

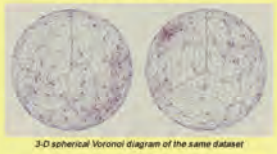


3-D sections from high-resolution 70S ribosome reconstruction with 3824 class averages obtained by BKPR

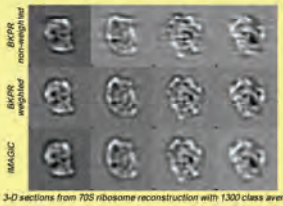
(5B)



2-D Mollweide projection of 1300 class-averages dataset and 2-D Voronoi diagram



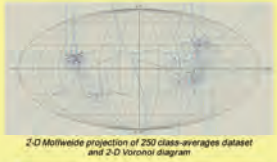
3-D spherical Voronoi diagram of the same dataset



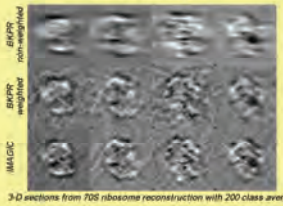
3-D sections from 70S ribosome reconstruction with 1300 class averages

In the case of strong preferential orientations of the particles, the effect of elongated densities in the reconstruction becomes so strong, that the structure gets completely distorted. As expected, the weighting of the projections significantly improves the density distribution. Comparison with IMAGIC-5 indicates that BKPR produces comparable results in the case of severe overweighting of certain directions and small number of projections, but the exact weighting results in even fewer stripes on the sections of the 3D reconstruction.

(5C)



2-D Mollweide projection of 250 class-averages dataset and 2-D Voronoi diagram



3-D sections from 70S ribosome reconstruction with 200 class averages

References (6)

Aurenhammer, F., 1991. Voronoi diagrams—a survey of a fundamental geometric data structure. ACM Computing Surveys (CSUR), 23, 345-405.
Brown, K.Q., 1979. Voronoi diagrams from convex hulls. Inform. Process. Lett., 9, 223-228.
Chen, D.M., Cho, Z.H., Wang, C.Y., 1996. A fast implementation of the incremental backprojection algorithms for parallel beam geometries. IEEE Trans. Nucl. Sci. 43, 3328-3334.
Crowther, R.A., 1971. Procedures for three-dimensional reconstruction of spherical viruses by Fourier synthesis from electron micrographs. Philos. Trans. Roy. Soc. Lond. B. 261, 221-230.
Frank, J., 1996. Three-dimensional electron microscopy of macromolecular assemblies. Academic Press, San Diego, California.
Fuller, S.D., Butcher, S.J., Cheng, R.H., Baker, T.S., 1996. Three-dimensional reconstruction of icosahedral particles-the uncommon line. J. Struct. Biol. Jan-Feb, 116(1), 48-55.
Gordon, R., Bender, R., Herman G., 1970. Algebraic reconstruction techniques (ART) for three dimensional electron microscopy and X-ray photography. Journal of Theoretical Biology, 36, 105-117.
Harauz, G., van Heel, M., 1986. Exact filters for general geometry three-dimensional reconstruction. Optik 73, 146-156.
Herman, G., Lent, A., 1976. Iterative reconstruction algorithms. Comput. Biol. Med. 6, 273-294.
Na, H., Lee, C., and Cheong, O. 2002. Voronoi diagrams on the sphere. Comput. Geom. Theory Appl. 23, 183-194.

The program source code and executable can be downloaded from the following WEB address:
http://www-igbmc.u-strasbg.fr/recherche/Dep_BGS/Eq_PBohm/Publi/Paper1.html

BKPR: a 3D reconstruction software with compensation for preferable orientations suitable for any point symmetry group.

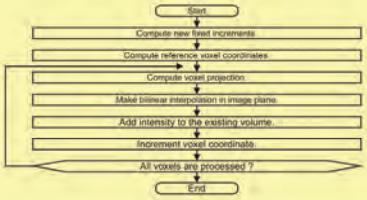
Igor M. Orlov, Alexander G. Myasnikov and Bruno P. Klaholz

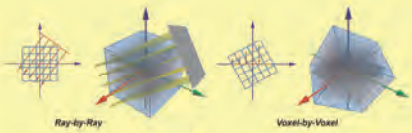
Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGMBC), 1 rue Laurent Fries, 67404 Illkirch CEDEX, France

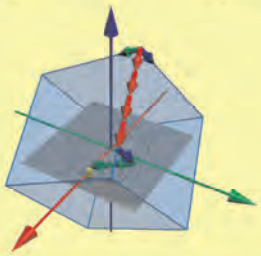
Fast Weighted Back-Projections

In the ray-by-ray backprojection approach, densities of the image are stretched out (back-projected) through the whole volume of the object being reconstructed in the direction that corresponds to the orientation found for this image. If the direction does not correspond exactly to one of the coordinate axes of the object matrix, the density of the nodes must be determined by 3D interpolation.

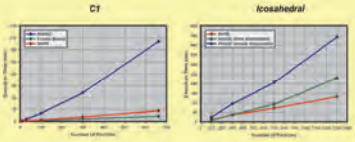
The voxel-by-voxel approach is based on the direct projection of each voxel to the image that has to be backprojected. In this case the density is determined by bilinear interpolation in the image plane.







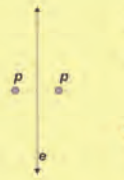
The orientation angles found for the projection image are used to determine the rotation matrix and unit vectors in the coordinate system related to the projection. Then fixed axial increments are determined, and used for computation of voxel coordinates. Processing of voxels is performed one by one to compute projections of each voxel on the image plane and the increment of the voxel density is determined by two-dimensional interpolation of neighboring pixels in the image. The calculated density is added to the voxel density of the 3D map assigned during the processing of the previous image. The program then proceeds to the next voxel.



Voronoi Diagrams on Plane and on a Sphere

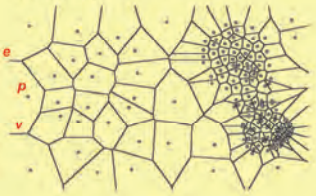
One of the serious problems in single particle analysis is the non-uniform orientation distribution in ice or on carbon film. This leads to distorted densities in three-dimensional reconstructions: the point-spread function does not have a spherical symmetry. This difficulty can be overcome by weighting of the projections according to their angular distribution on the Euler sphere.

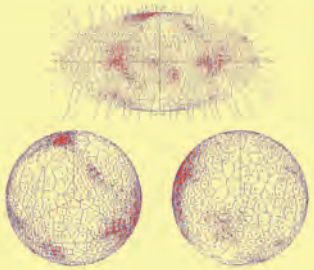
Hereby we suggest a new approach based on Voronoi diagrams for the determination of projection weights. A Voronoi polygon on the Euler sphere identifies a region whose interior comprises all points on the sphere which are closer to a given projection point than to any other, thus determining an "influence area" for each projection. This approach allows generating the quasi equal-density distribution so that the area of each Voronoi polygon on a sphere defines the weight for the corresponding projection.



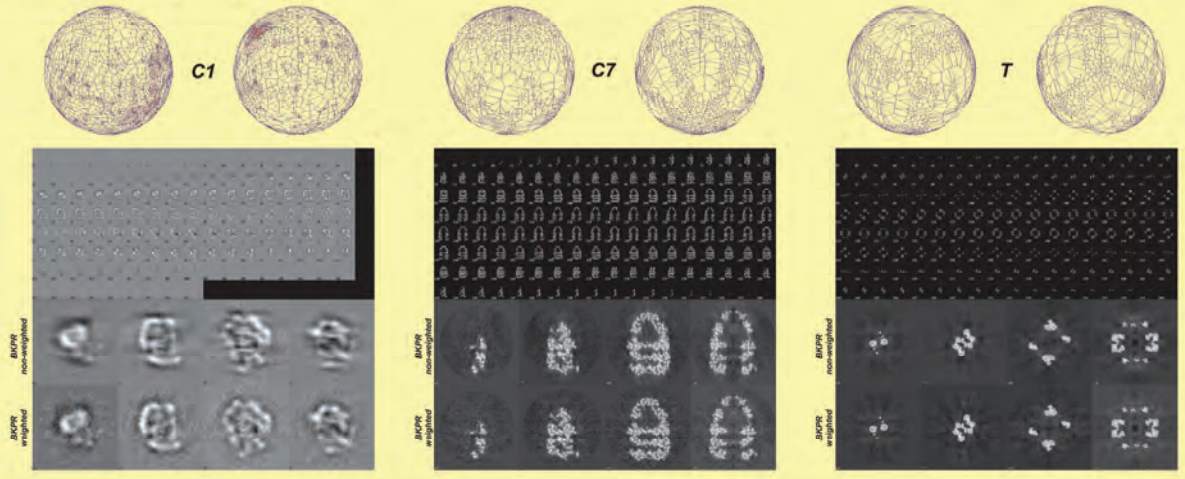
p: sitepoints
v: Voronoi vertex
e: Voronoi edge

- *Let P be a set of n distinct points (sites) in the plane.*
- *The Voronoi diagram of P is the subdivision of the plane into n cells, one for each site.*
- *A point q lies in the cell corresponding to a site pi ∈ P if Euclidean_Distance(q, pi) < Euclidean_Distance(q, pj), for each pj ∈ P, j ≠ i.*





Different Point Group Symmetrys



References

Aurenhammer, F., 1991. Voronoi diagrams—a survey of a fundamental geometric data structure. ACM Computing Surveys (CSUR), 23, 345-405.
Brown, K.Q., 1979. Voronoi diagrams from convex hulls. Inform. Process. Lett., 9, 223-228.
Chen, C.M., Cho, Z.H., Wang, C.Y., 1996. A fast implementation of the incremental backprojection algorithms for parallel beam geometries. IEEE Trans. Nucl. Sci. 43, 3328-3334.
Crowther, R.A., 1971. Procedures for three-dimensional reconstruction of spherical viruses by Fourier synthesis from electron micrographs. Philos. Trans. Roy. Soc. Lond. B. 261, 221-230.
Frank, J., 1996. Three-dimensional electron microscopy of macromolecular assemblies. Academic Press, San Diego, California.
Fuller, S.D., Butcher, S.J., Cheng, R.H., Baker, T.S., 1996. Three-dimensional reconstruction of icosahedral particles-the uncommon line. J. Struct. Biol. Jan-Feb, 116(1), 48-55.
Gordon, R., Bender, R., Herman G., 1970. Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography. Journal of Theoretical Biology, 36, 105-117.
Hatzuz, G., van Hasi, M., 1986. Exact filters for general geometry three-dimensional reconstruction. Optik 73, 146-150.
Herman, G., Lent, A., 1975. Iterative reconstruction algorithms. Comput. Biol. Med. 6, 273-294.
Na, H.-Lee, C., and Cheong, O. 2002. Voronoi diagrams on the sphere. Comput. Geom. Theory Appl. 23, 183-194.

Igor M. Orlov, David Gene Morgan and R. Holland Cheng.
Efficient implementation of a filtered back-projection algorithm using a voxel-by-voxel approach.
Journal of Structural Biology, Volume 154, Issue 3, June 2006, Pages 287-296.

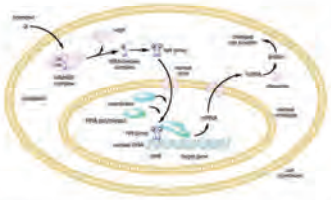
CNRS CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE
ulp UNIVERSITÉ LOUIS PASTEUR STRASBOURG
Inserm Institut national de la santé et de la recherche médicale
Région Alsace

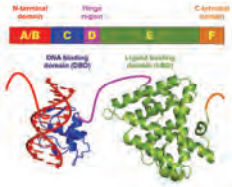
Cryo-electron microscopy structure of the 100kDa full nuclear receptor RXR/VDR heterodimer complex with it's target DNA

Igor Orlov, Natacha Rochel, Dino Moras, Bruno P. Klaholz

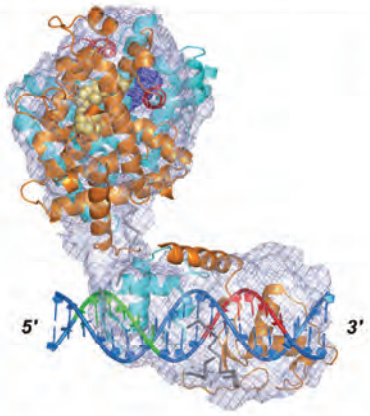
Department of Structural Biology and Genomics, IGMBC, 1 rue Laurent Fries, 67404 Illkirch CEDEX, France

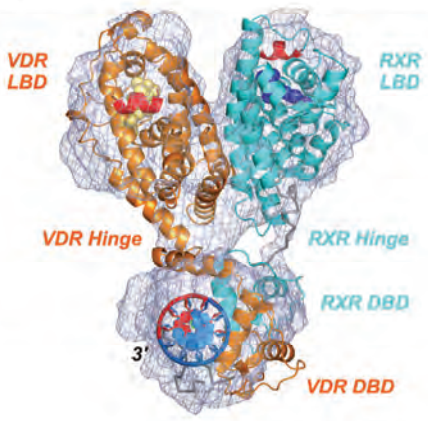
Transcription of the genetic information is a fundamental mechanism of gene expression regulation with important impact on human health. The nuclear receptors are ligand-dependant transcription regulators that bind to the promoter region of the DNA of a target gene. The activation of nuclear receptors by hormones (steroids etc.) or vitamins (vit. D, retinoic acid) or other agonists induces the binding of specific co-activators to the nuclear receptor. The co-activation proteins, like SRC (steroid co-activator receptor) have a particular role to modify the chromatin to make the target gene accessible to general transcription factors and the DNA-polymerase. Understanding the molecular mechanism of gene expression regulation by nuclear receptors, their ligands and associated co-activators is a fundamental question for future medical applications. Indeed, the nuclear receptors are directly involved in mechanisms such as cell growth, differentiation and apoptosis, thus explaining their therapeutic impact in the treatment of diseases like psoriasis, osteoporosis and different types of cancers.



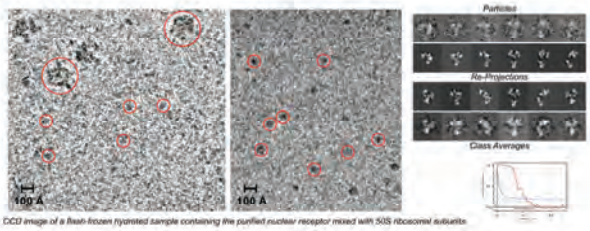


Until now, structural studies of nuclear receptors been focused mainly on their two core domains: the ligand-binding domain (LBD) and the DNA-binding domain (DBD). Here we present the architecture of the full nuclear receptor heterodimeric complex of the ligand-bound vitamin D receptor (VDR) and retinoid X receptor (RXR) bound to a consensus DNA response element forming a direct repeat (DR3). The structure of the RXR/VDR/DNA complex has been determined by single particle cryo-electron microscopy (cryo-EM) whilst overcoming technical challenges related to the 100kDa molecular weight of the object (Orlov et al., submitted). The structure reveals an open conformation with the DBD and LBD parts connected through well-resolved hinges and provides insights into the mechanism of ligand-dependant transcription regulation.



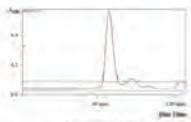


The images have been collected at liquid-nitrogen temperature with the in-house FEI Tecnai20 cryo-electron microscope at 200 KV voltage with a dose of ~20 electron per Å^2 at a magnification of 50,000 on photographic films. In order to determine the appropriate specimen concentration and best distribution of particles in vitreous ice we screened numerous grids and grid areas best suitable for data collection. For starting the first structure determination, focal pairs of microphotographs were recorded. The first close-to-focus microphotographs were taken at -2.5 - -4.5 µm underfocus. The second microphotographs were taken at -7.0 µm underfocus and were used primarily for particle identification and selection. Once the particle selection and initial structure determination had been validated with the defocus pair data set (6,800 particles) further data were collected at a defocus of -2.0 up to -4.0 µm (19,938 particles).



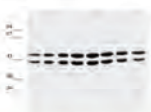
CCD image of a flash-frozen hydrated sample containing the purified nuclear receptor mixed with 50S ribosomal subunits

The HsVDRΔ (1-427Δ166-216) was expressed as a hexahistidine fusion protein. HsRXRαΔAB (130-462) was cloned into a pACYC plasmid encoding a non-tagged protein. Ligands (9-cis retinoic acid for RXR and 1α,25-dihydroxyvitamin D3 for VDR) were added in a 2-fold excess. The annealed oligonucleotide DNA was added in a 1.2 excess to the dimer. The complex was further purified on a gel-filtration S200 column. The final buffers were Tris 20 mM pH7.5, NaCl 50 mM, KCl 50 mM, $MgCl_2$ 4mM, DTT 5mM.



Gel Filtration Profile

Native Gel

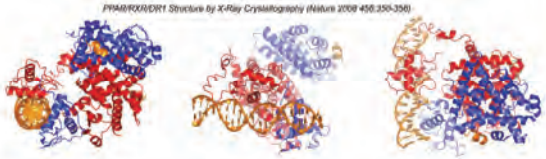


SDS Gel

The structure reveals an open conformation rather different from the crystallized PPAR/RXR size complex with RXR and VDR LBDs and DBDs adopting a parallel organization, with the LBDs perpendicular to the DNA and positioned at the DNA 5'-end of the response element. The DBD and LBD parts are connected through well-resolved hinges without additional contacts between DBDs and LBDs suggesting that the hinges provide independent links to the DBDs and the DNA and thereby stabilize the complex in a precise conformation on its target DNA. The DBDs are bound side-on to the DNA partially facing the LBDs, consistent with a general model according to which nuclear receptors could bind to their target DNA while keeping the area of transactivation helix H12 (marked in red) accessible for recruiting chromatin-modifying co-regulator proteins on the side oriented downstream of the DNA.

VDR/RXR/DR3 structure by Cryo-EM

PPAR/RXR/DR1 Structure by X-Ray Crystallography (Nature 2008 456:350-356)



This work was supported by the Agence National pour la Recherche (ANR), the Centre National pour la Recherche Scientifique (CNRS), the Institut National de la Santé et de la Recherche Médicale (INSERM), the European Molecular Biology Organization (EMBO) Young Investigator Programme (YIP), the Institut du Développement et des Ressources en Informatique Scientifique (IDRIS, France), and the European Commission as SPINE2-complexes (contract n°. LSHG-CT-2006-031220). I. O. was supported by a PhD fellowships from the Alsace Region and by the ANR. The electron microscope facility is supported by the Alsace Region, the Fondation pour la Recherche Médicale (FRM), INSERM, CNRS and the Association pour la Recherche sur le Cancer (ARC).