



UNIVERSITÉ DE STRASBOURG



École Doctorale
des Sciences de la Vie
et de la Santé
STRASBOURG

École Doctorale des Sciences de la Vie et de la Santé



Centre de Biologie Intégrative
Institut de biologie moléculaire et cellulaire
CNRS UMR 7104 – INSERM U 1258



Étude des mécanismes moléculaires régulant l'activité des complexes formés par le récepteur nucléaire $ERR\alpha$

Une Thèse présentée par : **Pierre VILELA**

Pour obtenir le grade de **Docteur de l'université de Strasbourg**

Discipline : Sciences de la Vie

Spécialité : Biophysique et Biologie Structurale

Sous la codirection de :

Dr. Isabelle Billas, Chargé de recherche, Université de Strasbourg

Dr. Bruno Klaholz, Directeur de recherche, Université de Strasbourg

Rapporteurs:

Albane Le Maire

Catherine Venien-Bryan

Membres du Jury:

Albane Le Maire, Chargé de recherche Université de Montpellier

Catherine Venien-Bryan, Professeur, Sorbonne Université

Dominique Housset, Ingénieur de recherche, Université de Grenoble

Daniel Metzger, Directeur de recherche, Université de Strasbourg

Remerciements

Je tiens à remercier les personnes sans lesquelles ce travail n'aurait pu être mené à bien. Tout d'abord ma femme, mes proches et ma famille pour leur soutien indéfectible, tant matériel que moral. Une thèse est éprouvante à bien des niveaux, aussi leur compréhension et leur aide précieuse fut plus d'une fois salvatrice. Je remercie la ligue contre le cancer et l'association de recherche contre le cancer d'avoir financés mes travaux, ainsi que l'institut de génétique et de biologie moléculaire et cellulaire de m'avoir accueilli, et plus particulièrement l'équipe dirigée par Bruno Klaholz, ainsi que ma supérieure Isabelle Billas.

J'aimerai remercier les membres de l'équipe qui m'ont apporté connaissance et support. Merci à Isabelle Hazemann, à Vanessa Da Cleu et à Kareem Mohideen-Abdul pour les conseils et les discussions que nous avons pu avoir sur l'expression, la purification et l'étude structurale des protéines. Merci à mes collègues Charles Barchet, Samuel Holvec, Brice Beinsteiner, Nicolas Ballet et Léo Fréchin pour leurs nombreux coups de mains dans le domaine informatique, ainsi que dans le traitement d'images. Merci également à Otilie Von Loeffelholz pour son aide dans l'analyse des données et la compréhension des fichiers particules.

Merci au personnel des différentes plate-formes de l'institut. Je pense ici plus particulièrement à Pierre Poussin et Alastair Mc Ewen pour leur aide concernant l'utilisation des outils de cristallogénèse et des lignes synchrotron. Merci à Catherine Birck et Florence pour leurs conseils et leur temps concernant les analyses biophysiques, notamment sur l'utilisation de l'ultracentrifugation analytique, le FIDAbio et plus encore... Merci évidemment à Nils Maréchal et Alexandre Durand qui m'ont aidé dans la prise en main des microscopes et dans le traitement d'images, en répondant à mes très nombreuses questions et en m'accompagnant dans ma formation. De façon plus générale, je remercie les plateformes du centre de biologie intégrative et l'infrastructure FRISBI qui les supporte.

Je remercie également nos collaborateurs de l'Institut de Biologie Moléculaire et Cellulaire Karl Brillet et Eric Ennifar. Ils m'ont aidé dans la réalisation et l'analyse des résultats obtenus par l'usage de la titration calorimétrique isotherme. Je remercie également nos collaborateurs de Cronembourg, dont plus particulièrement Marie Ley, pour leur aide dans la réalisation et l'interprétation des expériences menées par spectrométrie de masses et plus particulièrement pour les expériences de HDX et de XL-MS.

Table des Matières

Liste des abréviations.....	1
Résumé.....	5
Introduction.....	9
1. Les ERR : nœuds de transcription métabolique.....	11
1.1. La régulation du métabolisme énergétique.....	11
1.2. Implication dans la production de ROS et dans la détoxification.....	13
1.3. La genèse et le maintien des mitochondries.....	13
1.4. Implication dans la réponse immunitaire.....	14
1.5. Capteurs du cycle circadien.....	14
2. Le potentiel de ERRa et ERRg dans la lutte contre le cancer.....	15
2.1. Le rôle des ERR en tant qu'outil de diagnostic.....	15
2.2. Le rôle des ERR en tant que cible thérapeutique.....	16
3. Aspects structuraux et fonctionnels des ERR.....	16
3.1. Organisation générale des récepteurs nucléaires.....	16
3.1.1. Le domaine N-terminal.....	17
3.1.2. Le domaine de liaison à l'ADN (DBD).....	18
3.1.3. La région coudée.....	24
3.1.4. Le domaine de liaison des ligands.....	24
3.1.5. Le domaine F.....	27
3.2. Positionnement des ERR parmi les récepteurs nucléaires.....	27
3.2.1. <i>La régulation du domaine N-terminal des ERR</i>	28
3.2.2. <i>Les particularités du domaine de liaison des ligands de ERR</i>	28
3.2.3. <i>Le domaine de liaison à l'ADN</i>	31
4. Les complexes impliquant les récepteurs nucléaires.....	35
4.1. Les complexes de récepteurs nucléaires entiers impliquant l'ADN.....	35
4.2. <i>Les complexes impliquant des répresseurs</i>	36
4.3. Les complexes impliquant des coactivateurs.....	37
1. <i>L'étude par diffraction aux rayons-X</i>	40
2. L'étude d'échantillons biologiques par cryo-microscopie électronique à transmission.....	43
2.1. La préparation de l'échantillon.....	43
2.2. La prise d'images.....	45
2.3. La reconstruction 3D de l'objet imagé.....	54
3. La titration calorimétrique isotherme.....	58
4. La méthode d'analyse Proteoplex.....	59
5. FIDAbio : Analyse de dispersion par flux induit.....	60
6. L'échange hydrogène-deutérium.....	62
7. La réticulation couplée à la spectrométrie de masse (XL-MS).....	63
Partie I – L'ADN dicte la liaison asymétrique de ERRa.....	67
Résumé de l'article.....	67
Conclusions.....	67

Article – Résultat (version actuelle du Manuscrit).....	69
Partie II – PGC-1a régule l’assise de ERRa sur l’ADN.....	102
Introduction.....	102
Matériels & Méthodes.....	104
1/ Choix des constructions.....	104
2/ Hybridation des ADN.....	105
3/ Expression des protéines.....	106
3.1/ Infection et expression des protéines produites en baculovirus.....	106
3.2/ Transformation et expression des protéines produites en bactéries.....	107
4/ Purification des protéines.....	107
4.1/ Purification des protéines produites en cellules d’insectes.....	107
4.2/ Purification des protéines produites en bactéries.....	108
4.3/ Purification de p300.....	108
5/ Formation des complexes.....	109
5.1/ Formation des complexes ERR-PGC-ADN.....	109
5.2/ Formation des complexes ERR-PGC-p300.....	109
5.3/ Formation des complexes ERR-PGC-PCAF et ERR-PCAF-ADN.....	109
6/ Cristallogénèse et diffraction aux rayons X.....	110
7/ Analyses biophysiques du complexe ERRa-PGC-1a.....	110
7.1/ Analyses par la méthode FIDABio.....	110
7.2/ Préparation des échantillons pour les analyses par spectrométrie de masse.....	111
7.3/ Préparation des échantillons pour analyse par la méthode Protéoplex.....	111
8/ Vitrification du complexe ERR-PGC.....	111
8.1/ Vitrification du complexe mouse_ERRa_1-422-PGC1-472.....	111
8.2/ Vitrification du complexe human_ERRa_1-423-PGC-1a_1-285.....	112
9/ Acquisitions des jeux de données et pré-traitement.....	112
Résultats.....	116
1/ PGC-1a permet l’assise de ERRa sur ses ERRE.....	116
2/ PGC-1a interagit avec ERRa au niveau d’une région très spécifique.....	120
3/ PGC-1a induit le changement conformationnel de ERRa.....	123
4/ P300 et PCAF forment un complexe avec ERRa.....	125
5/ Effet de la nature de l’échantillon et du support sur le complexe.....	127
6/ Traitement des données du jeu mouse_ERRa_1-422-PGC-1a_1-472-BE42embedded.....	130
7/ Traitement des données du jeu human_ERRa_1-423-PGC-1a_1-285-BE42embedded (Chameleon).....	134
8/ Identification des domaines individuels.....	139
9/ Amélioration des cartes du complexe ERRa-PGC-1a.....	143
10/ Description de la topologie du complexe.....	150
Discussion et Perspectives.....	154
Bibliographie.....	161
ANNEXES.....	178
Résumé.....	199

Liste des abréviations

ADN : Acide désoxyribonucléique
AF : fonction d'activation (souvent suivie du numéro 1 ou 2)
AR : Récepteur aux androgènes
ARN : acide ribonucléique
ARNm : Acide Ribonucléique messenger
ARN Pol : ARN polymérase
ATP : Adénine tri-phosphate
BET : Bromure d'éthidium
BPA : Bisphénol A
CARM : Arginine Methyl-transferase associée au coactivateur 1
CHAPS : 3-[(3-cholamidopropyl)diméthylammonio]-1-propanesulfonate
Cryo-EM : Cryo-microscopie électronique
Cs : Coefficient d'aberrations sphériques
CTE : extension Carboxy-terminale
CTF : fonction de transfert du contraste
DBD : Domaine de liaison à l'ADN
DDM : n-dodecyl- β -D-maltoside
DES : diethylstilbestrol
DLS : Diffusion dynamique de la lumière
DQE : Efficacité quantique de détection
DR : répétition directe
E : fonction enveloppe
ER : récepteur aux œstrogènes
ERE : élément de réponse aux récepteurs aux œstrogènes
ERR : Récepteur apparenté aux récepteurs aux œstrogènes
ERRE : élément de réponse aux ERR
Fig : Figure
FT : fractions non retenues sur colonne
GR : Récepteur aux glucocorticoïdes
HDX : Echanges hydrogènes-deutériums
HEPES : 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HNF-4 : facteur nucléaire hépatocytaire 4
HSP : Protéine de choc thermique
IMAC : chromatographie d'affinité aux ions métalliques
IPTG : isopropyl β -D-1-thiogalactopyranoside
IR : répétition inverse
ITC : Titration Calorimétrique isotherme
LBD : domaine de liaison des ligands
LBP : poche de liaison des ligands
MET : Microscopie à transmission électronique
MR : récepteur aux minéralocorticoïdes
MS : spectrométrie de masse

NR : Récepteur(s) nucléaire(s)
pb : paires de bases
PCAF : Coactivateur associé au facteur p300
PDB : Protein Data Bank
PGC-1 α : coactivateur-1-alpha du récepteur gamma activé par les proliférateurs du peroxyosome 1 α
PPAR : récepteur activé par les proliférateurs du peroxyosome 1 α
PMSF : fluorure de phénylméthylsulfonyle
p300 : Histone acetyltransferase p300
RAR : Récepteur à l'acide rétinoïque
RIMAC : Reverse-IMAC
RMN : Résonance magnétique nucléaire
ROR : récepteur apparenté aux récepteurs aux rétinoïdes
RXR : récepteur X des rétinoïdes
SAXS : diffusion de rayons X aux petits angles en solution
SDS : Sodium Dodecyl Sulfate
SEC : chromatographie d'exclusion stérique
SEC-MALS : SEC couplée à une diffusion de lumière multi-angles
SPA : Analyse de particules isolées
SRC : Coactivateur des récepteurs stéroïdiens
Tris : tris(hydroxyméthyl)aminométhane
Uniprot : Universal Protein Ressource
VDR : Récepteur de la vitamine D
2D : deux dimensions
3D : trois dimensions

Index des Figures

Figure 1: Les différents types de communication intercellulaire.....	9
Figure 2: Schéma récapitulatif de l'action des ERR sur le métabolisme énergétique.....	15
Figure 3: Organisation modulaire des récepteurs nucléaires.....	17
Figure 4: Organisation fonctionnelle et structurale des DBD.....	19
Figure 5: Schéma du mécanisme général des récepteurs nucléaires.....	23
Figure 6: Organisation structurale des LBD.....	25
Figure 7: Mécanisme d'action du ligand sur le LBD de RXR.....	26
Figure 8: Identité au sein de la famille des ERR.....	28
Figure 9: Mécanisme de reconnaissance de PGC-1a par ERRa comparé à celui de TIF-2 par ERa.....	29
Figure 10: Différence de sélection des agonistes inverses au sein des ERR.....	31
Figure 11: Les sites de liaison des ERR.....	32
Figure 12: Comportement de ERRa en fonction des éléments de réponse.....	33
Figure 13: Structure de ERRb obtenue par RMN sur un ERRE.....	34
Figure 14: Récepteurs nucléaires sur leurs éléments de réponses.....	35
Figure 15: Les complexes de régulation de la transcription impliquant les récepteurs nucléaires.....	37
Figure 16: Complexes d'activation de la transcription formés par ERa.....	38
Figure 17: Cartes de cryo-microscopie du récepteur aux androgènes sur son élément de réponse.....	39
Figure 18: Principe de la diffraction aux rayons-X.....	41
Figure 19: Diagramme de phases.....	42
Figure 20: Principe de la vitrification d'un échantillon pour analyse par cryo-EM.....	44
Figure 21: Système de vitrification automatisé Chaméléon.....	45
Figure 22: Anatomie d'un microscope électronique.....	46
Figure 23: Interactions entre le faisceau d'électrons et l'échantillon.....	47
Figure 24: La formation de l'image.....	48
Figure 25: Impact des caméras sur l'évolution de la qualité des images.....	49
Figure 26: Impact du defocus sur le contraste et la CTF.....	53
Figure 27: La correction de la CTF.....	54
Figure 28: Illustration du Théorème de la section centrale.....	56
Figure 29: Principe de la titration calorimétrique isotherme.....	58
Figure 30: Principe de la méthode Protéoplex.....	60
Figure 31: Schéma du protocole expérimental effectué par l'instrument FIDABio.....	61
Figure 32: Protocole suivi pour les expériences d'échanges hydrogènes-deutériums.....	62
Figure 33: Protocole suivi pour les expériences de réticulation.....	63
Figure 34: Analyse d'une première carte du complexe ERRa/PGC-1a/ADN par comparaison avec les structures existantes.....	102
Figure 35: Mise en évidence de l'effet allostérique induit par la fixation de PGC-1a.....	103
Figure 36: PGC-1a permet l'ancrage des ERR sur les ERRE.....	117
Figure 37: Effet de PGC-1a sur l'interaction entre ERRa et un élément composite.....	118
Figure 38: Représentation des réticulations intra-moléculaire de ERRa en fonction de PGC-1a et l'ADN.....	120
Figure 39: Représentation des réticulations observées entre ERRa et PGC-1a.....	121
Figure 40: Schéma des régions protégées et déprotégées sous l'influence de PGC.....	124
Figure 41: Résultats des analyses par la méthode Protéoplex.....	126
Figure 42: Migration sur gel natif du complexe human_ERRa_1-423-PCAF-wc29embedded.....	126

Figure 43: Portion de micrographes des différents complexes sur différents types de grilles.....	128
Figure 44: Classification 2D des complexes selon les jeux de données sur différents supports.....	129
Figure 45: Figure du traitement de données dans RELION.....	131
Figure 46: Figure Traitement des données dans le logiciel cryoSPARC.....	133
Figure 47: Principales étapes de la reconstruction de l'objet à partir de la première acquisition de la grille Chameleon.....	135
Figure 48: Principales étapes de la reconstruction de l'objet à partir de la seconde acquisition.....	136
Figure 49: Principales étapes de la reconstruction de l'objet à partir des particules des deux acquisitions extraites en boîtes de 200 pixels.....	137
Figure 50: Principales étapes de la reconstruction de l'objet à partir des deux acquisitions.....	138
Figure 51: Mise en perspective des domaines de ERRa sur les cartes obtenues lors de la première acquisition.....	140
Figure 52: Mise en perspective des domaines de ERRa sur les cartes obtenues lors de la seconde acquisition.....	142
Figure 53: Variabilité et mouvements au sein du complexe.....	144
Figure 54: Classification 3D du complexe ERR-PGC.....	146
Figure 55: Traitement de la première classe obtenue par classification 3D.....	147
Figure 56: Traitement de la seconde classe obtenue par classification 3D.....	149
Figure 57: Affinement local de la région LBDs du complexe ERR-PGC-ADN.....	151
Figure 58: Analyse de la carte chirale.....	153
Figure 59: Comparaison de la topologie du complexe ERRa-PGC-1a avec les récepteurs USP/ECR et HNF-4.....	158

Index des tableaux

Table 1: Récapitulatif du type d'élément de réponse par récepteur et par classe.....	21
Tableau 2: Constructions destinées à la production en <i>E. coli</i>	105
Tableau 3: Constructions destinées à la production en cellules d'insecte.....	105
Tableaux 4: ADN utilisés.....	105
Tableau 5 : Conditions d'acquisition du complexe human_ERRa_1-423-PGC-1a 1-285-BE42embbeded sur grille d'or.....	113
Tableau 6: Conditions de la seconde acquisition du complexe human_ERRa_1-423-PGC-1a 1-285-BE42embbeded vitrifié au Chameleon.....	114
Tableau 7: Conditions de la seconde acquisition du complexe human_ERRa_1-422-PGC-1a 1-472-BE42embbeded vitrifié au <i>Vitrobot</i>	115

Résumé

Les récepteurs nucléaires sont une famille de facteurs de transcription impliqués dans la régulation et l'adaptation de l'expression génétique, ainsi que dans la différenciation cellulaire. Ils sont également impliqués dans le développement de maladies métaboliques telles que le diabète ou le cancer. Leur mode d'action consiste en la reconnaissance de séquences cibles sur l'ADN que l'on qualifie d'éléments de réponse. Ils s'y fixent principalement sous forme d'homodimères ou d'hétérodimères suivant le récepteur considéré. Une fois fixés, ils vont être régulés par un ensemble de coactivateurs et/ou de corépresseurs, ainsi que par des modifications post-traductionnelles selon le contexte cellulaire. Ils forment alors des complexes d'activation de la transcription. Les récepteurs nucléaires s'organisent selon une structure modulaire de quatre à cinq domaines : le domaine N-terminal, le domaine de liaison à l'ADN (DBD), la région coudée, le domaine de liaison des ligands (LBD) et le domaine F. Le domaine N-terminal est un domaine peu structuré en solution et le siège de modifications post-traductionnelles régulant l'activité du récepteur. Le domaine DBD est le domaine de reconnaissance des éléments de réponse sur l'ADN. Il est le plus conservé des domaines. La région coudée est une région flexible entre les LBDs et DBDs qui interagit, dans certains cas, avec des coactivateurs ou corépresseurs. Le LBD est un domaine qui peut reconnaître et fixer un ligand au sein d'une poche nommée la poche de liaison des ligands (LBP). La liaison du ligand occasionne le déplacement de l'hélice H12 du LBD. Le récepteur passe alors en conformation active et devient capable de lier des coactivateurs. Sur le plan thérapeutique, des molécules antagonistes sont développées pour les maintenir en conformation inactive.

Le récepteur apparenté au récepteur aux oestrogènes alpha ($ERR\alpha$) est un récepteur nucléaire orphelin, c'est-à-dire qu'il ne possède pas de ligand naturel connu. Il est constitutivement actif : la composition en acides aminés de sa LBP stabilise le récepteur en conformation active. Physiologiquement, il est impliqué dans la régulation du métabolisme énergétique des cellules. Il régule notamment l'expression de l'ensemble des enzymes impliquées dans le cycle de Krebs, dans la glycolyse, dans la B-oxydation des acides gras et la catabolise des corps cétoniques, mais se retrouve aussi impliqué dans l'intégrité mitochondriale, la détoxification des déchets de la chaîne respiratoire et l'immunité. Sur le plan de la santé, il a été montré que la surexpression de $ERR\alpha$ dans plusieurs lignées cellulaires tumorales promeut leur croissance, leur agressivité et leur survie dans les cancers du sein, du colon et de l'endomètre. $ERR\alpha$ reconnaît sur l'ADN un demi-site étendu de trois paires de bases en 5' : « TNAAGGTCA ». Ce type de séquence est qualifié d'élément de réponse aux ERR (ERRE). Cependant, $ERR\alpha$ est actif sous forme de dimère *in vivo*, ce qui ne va pas dans le sens d'une fixation à l'ADN par l'intermédiaire d'un seul demi-site. Il a été montré au sein de notre équipe en 2017 (Mohideen-Abdul *et al.*, 2017) que les DBDs du récepteur $ERR\alpha$ pouvaient naturellement se lier sous forme de dimère sur un élément de réponse constitué de deux sites inversement répétés et séparés de trois paires de bases que le récepteur aux oestrogènes peut aussi lier : « TNAAGGTCANNNTGACCT ». Cette séquence est nommée un élément de réponse composite. Notre équipe a également pu observer que l'action d'un coactivateur, le coactivateur-1-alpha du récepteur gamma activé par les proliférateurs du peroxysoxe ($PGC-1\alpha$) renforçait la liaison du deuxième DBD à l'ADN sur le second demi-site de fixation.

Ma thèse s'est focalisée sur l'étude structurale du domaine de liaison à l'ADN pour élucider les mécanismes moléculaires expliquant sa capacité à dimeriser sur l'élément de réponse composite, puis dans l'étude biophysique et structurale du complexe $ERR\alpha$ -PGC-1 α -ADN en caractérisant l'impact du coactivateur sur l'activité du récepteur nucléaire.

Dans le premier chapitre, nous avons caractérisé l'interaction du domaine de liaison à l'ADN de ce récepteur nucléaire avec l'élément de réponse composite et expliqué les mécanismes mis en jeu dans la dimérisation sous contrôle de l'ADN. L'étude par diffraction aux rayons-X révèle une carte de densité électronique à une résolution de 2.48 Å du complexe. L'interface de dimérisation y apparaît asymétrique, ce qui est une différence notable avec le récepteur aux œstrogènes. Cette interface fait notamment intervenir une nouvelle boîte fonctionnelle nommée la boîte « KRRRK », laquelle est unique et conservée dans la famille des ERR. Les deux DBDs apparaissent comme étant non équivalents. En effet, l'interface de dimérisation fait apparaître un tour d'hélice 3_{10} au niveau de l'interface de dimérisation sur le monomère liant le premier demi-site étendu. Ce même DBD est également le seul pour lequel une interaction avec le petit sillon de l'ADN est observable. L'analyse évolutive des séquences et des structures des DBDs des autres récepteurs nucléaires révèle que ce mécanisme de dimérisation asymétrique est ancestral et semble avoir été perdu entre la famille des ERR et la famille des récepteurs aux œstrogènes. Les analyses par titration calorimétrique isotherme et les migrations sur gel en conditions natives couplées aux précédentes expériences réalisées au sein de l'équipe révèlent que les caractéristiques topologiques de l'ADN sont déterminantes dans la capacité des DBD à dimériser. Dans le cas des autres ERRE, la dimérisation des DBDs n'est pas possible. Il est également montré par les profils thermodynamiques des interactions avec les différents ERRE naturels que l'affinité d'interaction du dimère avec l'élément composite est comparable à celle des autres ERRE, mais aussi que la dimérisation s'effectue de façon coopérative et non séquentielle. Ce travail est l'objet d'une publication qui va être soumise sous peu.

Dans le second chapitre, nous nous sommes intéressés au récepteur nucléaire $ERR\alpha$ entier et au rôle du coactivateur PGC-1 α sur son activité transcriptionnelle. L'objectif est d'élucider la façon dont s'effectue la communication entre domaines au sein du récepteur, l'impact du coactivateur sur la liaison à l'ADN, mais aussi de confirmer la stœchiométrie observée par d'autres expériences d'un seul coactivateur par dimère actif de $ERR\alpha$ malgré la présence de deux sites de fixation potentiels. De plus, les précédents résultats de diffraction aux rayons-X à petits angles (SAXS), de chromatographie d'exclusion stérique couplée à une diffusion de lumière multi-angles (SEC-MALS) et de spectrométrie de masse native obtenus au sein de l'équipe suggèrent que PGC-1 α permet l'assise du second DBD sur les ERRE. Ils suggèrent également que PGC-1 α renforce la liaison à l'ADN quelque soit le type d'élément de réponse. Il s'agit d'une hypothèse qui est étudiée dans cette thèse.

Les expériences d'échanges hydrogènes-deutérium (HDX) réalisées en collaboration avec l'équipe de Sarah Cianférani à l'institut pluridisciplinaire Hubert Curien sur le complexe $ERR\alpha$ -PGC-1 α mettent en évidence un grand nombre de changements conformationnels au sein du récepteur nucléaire. Tout d'abord, l'interaction connue avec l'hélice H12 est confirmée au niveau des LBDs par la protection de cette région lors de la formation du complexe. Néanmoins, celle-ci s'étend également à la boucle H11-H12 et à une partie de l'hélice H11 du récepteur. Il est également observé une protection entre la boucle H1-H3 des LBDs lors de l'ajout du coactivateur, ainsi qu'une déstabilisation de l'hélice H6. Au niveau des DBDs, on observe la protection des boîtes fonctionnelles D et T et la déprotection de l'hélice H1. Les expériences de spectrométrie de masse après réticulation et digestion des complexes menées en parallèle (XL-MS) permettent d'identifier une région de 51 résidus du coactivateur susceptible d'interagir avec le récepteur entier. Il est tout d'abord observé que PGC-1 α est réticulé avec

les LBDs au niveau des hélices H11 et H12, ainsi qu'au niveau de la boucle H1-H3, dont la protection est observée en HDX. Ensuite, il est observé que les DBDs de $ERR\alpha$ présentent plus d'interactions entre eux lors de la liaison du coactivateur. Ce résultat couplé à la protection des régions impliquées dans l'interface de dimérisation en HDX implique un effet du coactivateur sur la dimérisation des DBDs. En présence d'ADN, l'interaction entre PGC-1 α et $ERR\alpha$ se renforce par l'apparition de nouvelles réticulations au niveau des hélices H11 et H12 des LBDs. De plus, comparé à la même expérience réalisée avec l'ADN seul, un renforcement du nombre de réticulations entre les DBDs de $ERR\alpha$ est observé, ce qui confirme que PGC-1 α favorise la dimérisation en présence d'ADN. Les mesures réalisées grâce à l'instrument FIDABio, nouvel équipement de la plateforme de biologie structurale intégrative de l'Institut de Génétique et de Biologie Moléculaire et Cellulaire qui héberge l'infrastructure de recherche FRISBI, ont permis de suivre l'évolution du rayon hydrodynamique du récepteur seul et du complexe avec des concentrations croissantes d'ADN. Les profils obtenus avec plusieurs ERRE dont l'élément composite confirment le gain conséquent d'affinité occasionné par la liaison de PGC-1 α . L'affinité augmente peu sur un élément composite, mais de cinq à dix fois sur un ERRE non composite. Cela va dans le sens de l'hypothèse selon laquelle la séquence facilite la dimérisation des DBDs et les ERRE non composite repose sur la nécessaire association du coactivateur PGC-1 α pour permettre la fixation du second DBD.

L'étude structurale par cryo-microscopie électronique (cryo-EM) du complexe $ERR\alpha$ -PGC-1 α produit en bactérie sur l'élément de réponse composite met en lumière le positionnement de PGC-1 α sur le récepteur nucléaire, ainsi que l'orientation des différents domaines sur l'ADN. L'étude de ce type d'échantillon est particulièrement difficile de part sa taille, mais surtout sa flexibilité induite notamment par son domaine N-terminal dont il n'existe aucune donnée structurale chez les récepteurs nucléaires, ainsi que par sa région coudée. Le coactivateur PGC-1 α est également une protéine intrinsèquement désordonnée en solution dont le repliement n'est que partielle en présence de $ERR\alpha$. Enfin, le complexe tend à mal se distribuer sur les grilles servant de support et est prône à la dissociation lors de la vitrification. L'obtention de nouveaux appareils, tels que le Chaméléon pour permettre une vitrification rapide en s'affranchissant de l'étape d'absorption de l'excès d'échantillon ou l'installation d'un microscope électronique Glacios fonctionnant à une tension d'accélération de 200 kV pour permettre un meilleur contraste, par la plate-forme de biologie structurale intégrative furent cruciaux pour l'obtention de cartes du complexe à une résolution d'environ 10 à 12 Å. Plusieurs conformations ont été obtenues et témoignent de la flexibilité structurale des régions coudées. La région de la carte correspondant aux LBDs permet notamment d'y apprécier la non-équivalence des LBDs lors de la fixation de PGC-1 α . La résolution de la carte ne permet cependant pas d'apprécier les structures secondaires des différents domaines de façon non ambiguë, notamment à cause de la difficulté à aligner correctement l'objet en 3D. Il est cependant possible de positionner les DBDs sur l'ADN. Ces cartes ont pu être comparées aux précédents modèles SAXS obtenus, ainsi qu'aux conclusions de l'HDX et de la XL-MS afin de confirmer les interactions précédemment décrites et confirmer l'emplacement des domaines individuels.

Dans l'optique de stabiliser davantage le complexe et d'obtenir des informations supplémentaires quant aux mécanismes moléculaires mis en jeu dans la régulation de l'activité de $ERR\alpha$, des constructions des protéines PCAF et p300 ont été produites. PCAF est une protéine capable d'acétyler les DBDs de $ERR\alpha$ afin de diminuer son interaction avec l'ADN. P300 est une protéine acétylant les histones pour permettre l'ouverture de la chromatine. Son rôle a notamment été mis en évidence du point de vue structurale dans la formation de complexes avec le récepteur aux œstrogènes. La purification de ces deux protéines recombinantes a été optimisée au cours de la thèse et l'utilisation de la méthode Protéoplex chez nos collaborateurs de l'entreprise Novalix ont permis d'identifier des

conditions propices à la formation des complexes. Des conditions permettant la formation d'un complexe $ERR\alpha$ -PCAF-ADN et d'un complexe $ERR\alpha$ -PGC-1 α -p300 ont pu être mises au point. Il reste néanmoins à optimiser davantage les conditions de reconstitution des complexes.

Enfin, la thèse se conclue sur la description de l'architecture de $ERR\alpha$ sur l'élément composite et sa comparaison avec d'autres cartes de récepteurs connues. La discussion traite des aménagements réalisés pour permettre la réalisation du projet, ainsi que des limites intrinsèques liées à la nature du complexe et aux limites techniques qui y sont liées. Enfin, les limites de l'interprétation des cartes seront discutées.

Introduction

La complexité d'un organisme multicellulaire nécessite une communication entre les cellules afin d'assurer leur cohésion et leur bon développement. Celle-ci repose sur la captation de signaux endocrines ou exocrines par le biais de molécules chimiques, permettant à une cellule de modifier son expression génétique afin de répondre à des besoins particuliers. Les différents types de communication intercellulaire sont présentés sur la Figure 1 (**Fig. 1**).

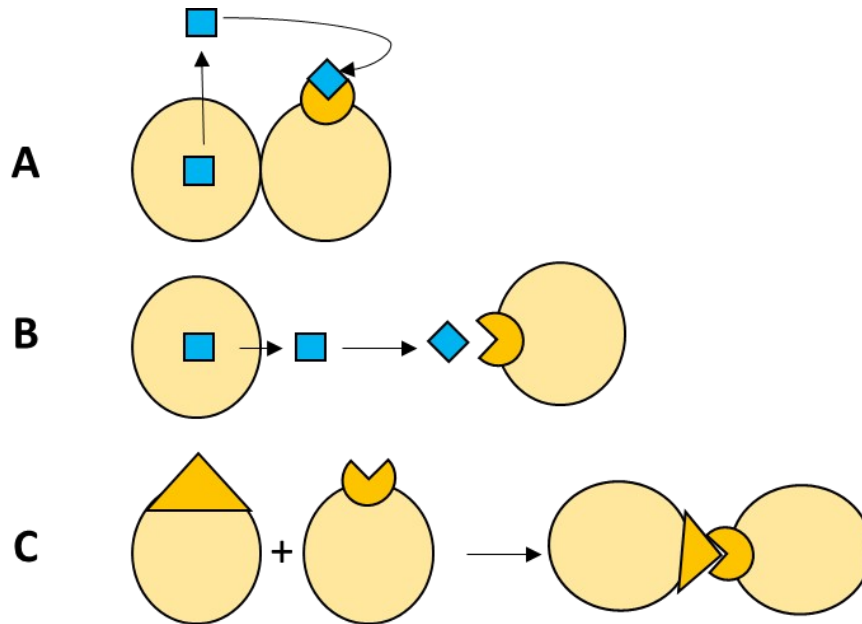


Figure 1: Les différents types de communication intercellulaire

Légende :

A/ Communication intercellulaire indirecte de proches en proches dites paracrine.

B/ Communication à distance par l'intermédiaire d'une molécule transportée dans l'organisme dite endocrine.

C/ Communication directe et physique entre les cellules dites juxtacrine.

La communication endocrine, qui est une communication distante entre les cellules, fut mise en évidence par la découverte d'une catégorie d'hormones, les œstrogènes, d'abord suspectées de réguler la synthèse d'acides ribonucléiques (ARN) (Mueller & McCorquodale, 1958), puis capables d'induire à distance la différenciation et la multiplication cellulaire, notamment par leur concentration dans les organes reproducteurs (Glascok & Hoekstra, 1959 ; Jensen & Jacobson, 1962). Il fut ensuite déterminé que des protéines liaient spécifiquement les œstrogènes et se trouvaient localisées dans le noyau cellulaire (Jensen *et al.*, 1967). Ces récepteurs sont capables de s'associer avec la chromatine et de stimuler la transcription (King *et al.*, 1966 ; Maurer & Chalkley, 1967 ; Teng & Hamilton, 1968 ; Shyamala & Gorski, 1969 ; Mainwaring, 1971 ; Mainwaring & Peterken, 1971 ; Spelsberg *et al.*, 1971 ; Steggle *et al.*, 1971). Enfin, la protéine réceptrice des œstrogènes, le récepteur aux œstrogènes (ER), fut isolée biochimiquement (Toft & Gorski, 1966) et le récepteur aux glucocorticoïdes (GR) fut le

premier a être cloné (Hollenberg *et al.*, 1985). Au cours des décennies suivantes, d'autres protéines réceptrices de molécules et impliquées dans la communication endocrines furent identifiées. On appela dès lors cette famille de protéines apparentées les « récepteurs nucléaires » (NR).

Les récepteurs nucléaires sont une famille de quarante-huit facteurs de transcription chez l'Homme (Mangelsdorf *et al.*, 1995). Leur fonction consiste à se fixer sur des séquences d'acide désoxyribonucléique (ADN) double-brins localisées au promoteur des gènes cibles ou de manière plus distante sur des séquences enhancers. Ces séquences sont qualifiées d'éléments de réponse (King & Gordon, 1972 ; Musliner & Chader, 1972 ; Toft, 1972 ; Andre & Rochefort., 1973 ; Beato *et al.*, 1973 ; Gehring & Tomkins, 1974 ; Yamamoto & Alberts, 1974 et 1975). Une fois fixés, les NR permettent le recrutement d'autres protéines et forment le complexe d'initiation de la transcription afin de stimuler l'activité de l'ARN polymérase (Nicolette, 1964 ; Kenney *et al.*, 1965 ; Means & Hamilton, 1966 ; O'Malley & al, 1968 ; Comstock & al, 1972 ; Glasser *et al.*, 1972 ; Means *et al.*, 1972 ; Rosenfeld *et al.*, 1972 ; Chan *et al.*, 1973 ; Parks *et al.*, 1974 ; Ringold *et al.*, 1975 et 1977 ; Scolnick *et al.*, 1976 ; Young *et al.*, 1977).

De plus, ces facteurs de transcription, afin de permettre l'expression de leurs gènes cibles, ont besoin de passer d'un état inactif à un état actif. Cette transition se déroule le plus souvent par l'intermédiaire d'une petite molécule hydrophobe capable de passer à travers les membranes de la cellule. Elle va interagir avec le récepteur afin de permettre son passage sous forme active. Cette dernière est qualifiée de ligand. Les ligands peuvent appartenir à quatre familles distinctes : les hormones stéroïdiennes, les hormones thyroïdiennes, les stéroïdes et les rétinoïdes. Les hormones stéroïdiennes se divisent en deux sous-catégories que sont les corticostéroïdes et les stéroïdes sexuels. Si plusieurs récepteurs ont été identifiés comme possédant un ligand, il en existe dont le ligand n'a pas encore été identifié ou qui n'ont pas besoin de ligand et sont constitutivement actifs. Ils sont qualifiés de « récepteurs orphelins » (Xu & Lambert, 2003).

Du point de vue physiologique, les récepteurs nucléaires régulent des mécanismes nombreux et variés. Nous pouvons citer en exemple la différenciation cellulaire, la multiplication cellulaire, le stockage des acides gras, la régulation du cycle circadien, l'embryogenèse, la réponse inflammatoire ou encore la détoxification cellulaire (Francis *et al.*, 2003 ; Alexander *et al.*, 2013). Ces nombreuses voies métaboliques sous leur contrôle en font des cibles thérapeutiques prometteuses. En effet, le rôle de plusieurs membres de cette famille de protéines a été démontré dans le développement de pathologies métaboliques telles que l'inflammation, le diabète, le cancer, mais aussi dans le cadre de sujets de société tels que la contraception (Tenbaum & Baniahmad, 1997 ; Moras & Gronemeyer, 1998 et 2004 ; Burris *et al.*, 2013). Des compagnies pharmaceutiques et des laboratoires de recherches ont mis au point des ligands synthétiques ciblant ces récepteurs pour réguler leur activité (Gronemeyer *et al.*, 2004 ; Xu, 2015).

I. Le récepteur orphelin apparenté au récepteur aux œstrogènes (ERR)

Si un certain nombre de récepteurs nucléaires ont pu être identifiés par l'identification de leur ligand, d'autres méthodes furent utilisées pour identifier des récepteurs orphelins. Ce fut le cas du récepteur nucléaire $ERR\alpha$ qui fut identifié pour la première fois en 1988 (Giguère *et al.*, 1988) par une stratégie d'hybridation croisée à faible stringence sur la base de la similarité entre son domaine de liaison à l'ADN (DBD) et celui du récepteur nucléaire ER, d'où il tire son nom. Puis, deux autres récepteurs nucléaires de sa famille, respectivement nommés $ERR\beta$ (Giguère *et al.*, 1988) et $ERR\gamma$, furent découverts (Heard *et al.*, 2000).

1. Les ERR : nœuds de transcription métabolique

Les ERR sont des facteurs de transcription qui se caractérisent physiologiquement par leur implication dans la régulation du métabolisme énergétique. $ERR\alpha$ et $ERR\gamma$ sont présents dans l'ensemble des types cellulaires, bien que l'expression de $ERR\alpha$ soit globalement plus importante que celle de $ERR\gamma$ (Bookout *et al.*, 2006). Cependant, ces deux protéines voient leurs niveaux d'expression fluctuer en fonction du cycle circadien. Ainsi, lors de la période diurne, l'expression de $ERR\gamma$ augmente dans le foie alors que celle de $ERR\alpha$ diminue pour augmenter durant la période nocturne, là où celle de $ERR\gamma$ diminue à son tour (Yang *et al.*, 2006). Quant à $ERR\beta$, ce dernier n'est pas aussi vastement exprimé que $ERR\alpha$ ou $ERR\gamma$. Il se retrouve essentiellement dans les cellules rénales, l'estomac ou la rétine, ainsi qu'exprimé plus fortement lors du développement embryonnaire (Bookout *et al.*, 2006).

1.1. La régulation du métabolisme énergétique

$ERR\alpha$, $ERR\beta$ et $ERR\gamma$ régulent les différentes voies métaboliques qui conduisent au cycle de Krebs et à la production d'énergie sous forme d'adénines tri-phosphates (ATP) et d'électrons à haut potentiel de transfert.

La glycolyse est l'une de ces voies, permettant le catabolisme du glucose en pyruvate au sein des cellules. $ERR\alpha$ régule un groupe de gènes liés fonctionnellement et nommé le « régulon bioénergétique de $ERR\alpha$ », lequel contient l'ensemble des enzymes impliquées dans la glycolyse (Chaveroux, C & al, 2013). $ERR\alpha$ promeut non seulement la glycolyse, mais il est également indispensable aux hépatocytes afin que ceux-ci passent d'un métabolisme oxydatif à un métabolisme glycolytique (Charest-Marcotte, A *et al.*, 2010). La glycolyse se termine généralement par la conversion des pyruvates produit à partir du glucose en acétyl-CoA. Cette réaction est sous le contrôle d'une protéine, la pyruvate déshydrogénase kinase 4 (PDK4) dont l'expression est régulée par $ERR\alpha$ (Wende *et al.*, 2005), mais aussi $ERR\gamma$ (Zhang Y, & al, 2006).

Les acides gras à longues chaînes nécessaires à l'organisme sont transportés dans la cellule par l'intermédiaire de Coenzymes A, ce qui permet notamment leur entrée au sein des mitochondries afin de produire des acétyl-CoA et d'alimenter le cycle de Krebs. Les trois ERR sont les principaux régulateurs de la transcription des gènes impliqués dans ce processus et sont même qualifiés de « maîtres exécuteurs » de cette voie métabolique (Fan W. *et al.*, 2015). Le gène *ACADM*, codant pour une enzyme impliquée dans la B-oxydation des acides gras, première étape limitante de cette voie, fut la première cible identifiée de $ERR\alpha$ (Sladek *et al.*, 1997 ; R. Vega & D. Kelly, 1997).

Les corps cétoniques sont des espèces chimiques produites transitoirement par les cellules, fournissant une source d'énergie alternative, notamment dans le cerveau, le cœur ou encore dans les muscles squelettiques durant le jeûne. Parmi les enzymes de cette voie métabolique, l'hydroxyméthylglutaryl-CoA synthase 2 et la 3-oxoacid-CoA-transférase sont régulées par ERR α (Chaveroux *et al.*, 2013 ; Svensson *et al.*, 2016).

ERR α régule également le métabolisme des acides aminés. Son inhibition est associée à une diminution de l'oxydation des leucines en conditions de fortes demandes en énergie (Sjogren *et al.*, 2021).

Les différentes voies métaboliques que nous venons d'évoquer conduisent au cycle de Krebs et à la production d'énergie sous la forme d'ATP et de huit électrons à haut potentiel de transfert, lesquels alimenteront la chaîne respiratoire. Les ERR, et plus particulièrement ERR α , ont un rôle majeur dans la régulation de ce cycle. En effet, ERR α et ERR γ ciblent des gènes impliqués dans l'intégrité mitochondriale (Eichner & Giguère, 2011) et l'inhibition de leur activité transcriptionnelle conduit à une diminution de l'intégrité des mitochondries, ainsi que de la production d'ATP (Charest-Marcotte *et al.*, 2010 ; Alaynick *et al.*, 2007 ; Dufour *et al.*, 2007 ; Sonoda *et al.*, 2007). Concernant les cellules hépatiques, ERR α y régule tous les gènes codant pour des enzymes impliquées dans le cycle de Krebs, ainsi que plus d'une centaine de gènes impliqués dans la chaîne respiratoire (Chaveroux *et al.*, 2013). De plus, les souris possédant une copie en moins du gène codant pour ERR γ dans leurs cellules musculaires sont moins performantes lors d'exercices. Cela a été associé à la diminution de leur activité mitochondriale (Rangwala *et al.*, 2010). Concernant cette fois les adipocytes, la perte des trois ERR conduit à une baisse drastique de l'expression des gènes impliqués dans les voies oxydatives et dans l'intégrité de leurs mitochondries (Gantner *et al.*, 2016) (**Fig. 2**).

Si plusieurs voies métaboliques mènent au cycle de Krebs, chacune n'est pas utilisée en même temps par la cellule, cette dernière privilégiant une source d'énergie par rapport à une autre en fonction de ses besoins. Un changement de voie connu s'opère lors d'un effort prolongé, remplaçant la glycolyse par la voie des acides gras pour alimenter le cycle de Krebs. Ce changement s'opère notamment par l'action de l'enzyme PDK4 qui est un régulateur négatif de l'oxydation du glucose. Il a été montré que ERR α augmente l'expression de PDK4 ce qui conduit à la suppression de l'oxydation du glucose dans les cellules musculaires et à l'augmentation de l'expression de gènes impliqués dans la régulation de la voie de l'oxydation des acides gras, augmentant l'utilisation des acides gras dans la production d'énergie par les muscles squelettiques. De plus, la surexpression de ERR γ conduit à la surexpression de gènes impliqués dans la chaîne respiratoire, le cycle de Krebs et le métabolisme des acides gras (Fan *et al.*, 2018). De plus, sa faculté à réguler le métabolisme oxydatif lui permet de jouer un rôle dans l'endurance des muscles squelettiques lors d'efforts (Rangwala *et al.*, 2010).

Une autre adaptation des voies métaboliques intervient lors de la transition du cœur de l'état fœtal à l'état adulte. Durant le stade fœtal, les muscles cardiaques utilisent principalement des hydrates de carbone comme source principale d'énergie, mais échangent celle-ci pour des acides gras et des lipides après la naissance. Les ERR assurent le contrôle dynamique de l'adaptation métabolique des myocardes (Vega & Kelly, 2017). ERR γ coordonne l'induction post-natal de la transition métabolique des cellules musculaires cardiaques par la régulation d'un réseau de gènes mitochondriaux (Alaynick *et al.*, 2007). Si ERR γ n'est pas nécessaire à la survie, sa délétion dans le cœur, couplée à celle de ERR α , cause une cardiomyopathie létale liée à des dysfonctionnements métaboliques de contractions et de conduction (Wang *et al.*, 2015).

En conclusion, les ERR sont des régulateurs fins et essentiels du métabolisme énergétique, régulant un ensemble vaste de gènes, mais disposant de la capacité de privilégier l'une ou l'autre des sources d'énergie en fonction des besoins cellulaires.

1.2. Implication dans la production de ROS et dans la détoxification

La production d'espèces porteuses d'oxygènes réactifs (ROS) est principalement associée au métabolisme énergétique et plus particulièrement au transport d'électrons au sein de la mitochondrie. La production des ROS et leur accumulation sont liées à plusieurs pathologies telles que le diabète, l'obésité ou d'autres syndromes métaboliques (Sies *et al.*, 2022). L'action d'enzymes antioxydantes prévient leur accumulation et les dommages que ces espèces réactives peuvent causer. Les ERR sont d'important régulateurs de ces enzymes (Scholtes & Giguère, 2021). $ERR\alpha$ régule notamment l'expression de gènes eux-mêmes impliqués dans la régulation de la réponse au stress oxydatif tels que les gènes Superoxyde dismutase 2 et 3, mais aussi les enzymes impliquées dans la chaîne respiratoire. La diminution de l'expression de $ERR\alpha$ est associée à une augmentation des dommages oxydatifs dans les cellules cancéreuses du sein (Deblois *et al.*, 2016). Par comparaison, la diminution de $ERR\gamma$ et/ou l'augmentation de $ERR\alpha$ conduit à une diminution des dommages causés par les ROS (Vernier *et al.*, 2020).

1.3. La genèse et le maintien des mitochondries

La biogenèse des mitochondries repose sur la transcription de gènes présent en sein du noyau et de gènes mitochondriaux. La plupart des protéines essentielles à cette étape demeurent exprimées à partir du génome nucléaire et sont importées au sein des mitochondries par l'intermédiaire d'une séquence signal et de transporteurs protéiques. $ERR\alpha$ occupe une place prépondérante dans ce phénomène puisqu'il active l'expression des facteurs de transcription mitochondriaux A, B1 et B2 nécessaires à la transcription de l'ADN mitochondrial, ainsi que les facteurs de respiration nucléaire 1 et 2, lesquels sont impliqués dans la biogenèse des mitochondries (Dufour *et al.*, 2007 ; Eichner et Giguère, 2011 ; Brown *et al.*, 2018). $ERR\alpha$ agit également de concert avec le coactivateur-1-alpha du récepteur gamma activé par les proliférateurs du peroxysoxe ($PGC-1\alpha$) afin de promouvoir la biogenèse des mitochondries (Cho *et al.*, 2013 ; Nie & Wong, 2009 ; Schreiber *et al.*, 2004). Il promeut aussi l'autophagie en cas d'infection et maintient les fonctions mitochondriales des cellules épithéliales intestinales par l'activation de l'expression génique induite par le facteur de transcription EB (Kim *et al.*, 2021). $ERR\gamma$, pour sa part, augmente la biogenèse de mitochondries en réponse à l'exercice indépendamment de $PGC-1\alpha/\beta$ (Fan *et al.*, 2018). La déficience de $ERR\gamma$ dans le cœur induit des changements dans l'expression de gènes impliqués dans la biogenèse des mitochondries, causant des défauts cardiaques au sein des embryons (Alayrick *et al.*, 2007). $ERR\beta$ est capable de lier la région codant pour le facteur de transcription mitochondrial A (Yu *et al.*, 2008) bien que son rôle exact ne soit pas encore clairement défini.

De plus, les ERR sont également impliqués dans la dynamique des mitochondries, c'est-à-dire dans les événements de fusions et de fissions des mitochondries. $ERR\alpha$ et $ERR\gamma$ régulent directement l'expression des protéines mitofusines 1 et 2 impliquées dans la fusion des membranes externes des mitochondries dans le cœur et les muscles (Cartoni *et al.*, 2005 ; Soriano *et al.*, 2006). L'expression de $ERR\alpha$ est nécessaire pour la régulation de la fission des mitochondries sous le contrôle de DRP1 (Singh *et al.*, 2018). La délétion dans le cœur de ces deux récepteurs nucléaires induit l'arrêt de

l'expression de l'ensemble des gènes liés à la fusion et à la fission des mitochondries (Wang *et al.*, 2015).

1.4. Implication dans la réponse immunitaire

ERR α est impliquée dans l'expression de gènes promouvant la défense de l'hôte durant une infection mycobactérienne par l'autophagie sélective du pathogène (Kim *et al.*, 2018).

1.5. Capteurs du cycle circadien

Le cycle circadien correspond à une « horloge biologique » d'une durée d'environ 24 heures chez l'Homme qui se caractérise par l'expression rythmique de certains gènes par des boucles de régulation interconnectées (Kim & Lazar, 2020). Sa régulation est également soumise à des capteurs du métabolisme énergétique et sa dérégulation est liée à l'apparition de pathologies métaboliques et de cancers (Panda, 2016).

ERR α est capable de lier des séquences promotrices de gènes régulant le cycle circadien (Dufour *et al.*, 2011). ERR α et ERR γ sont également exprimés de façon cyclique, ERR α étant principalement exprimé durant la période nocturne, alors que ERR γ est principalement exprimé durant la période diurne, ce qui permet l'adaptation du métabolisme énergétique par rapport à l'horloge biologique, ainsi que l'élimination efficace des ROS (Zhao *et al.*, 2014).

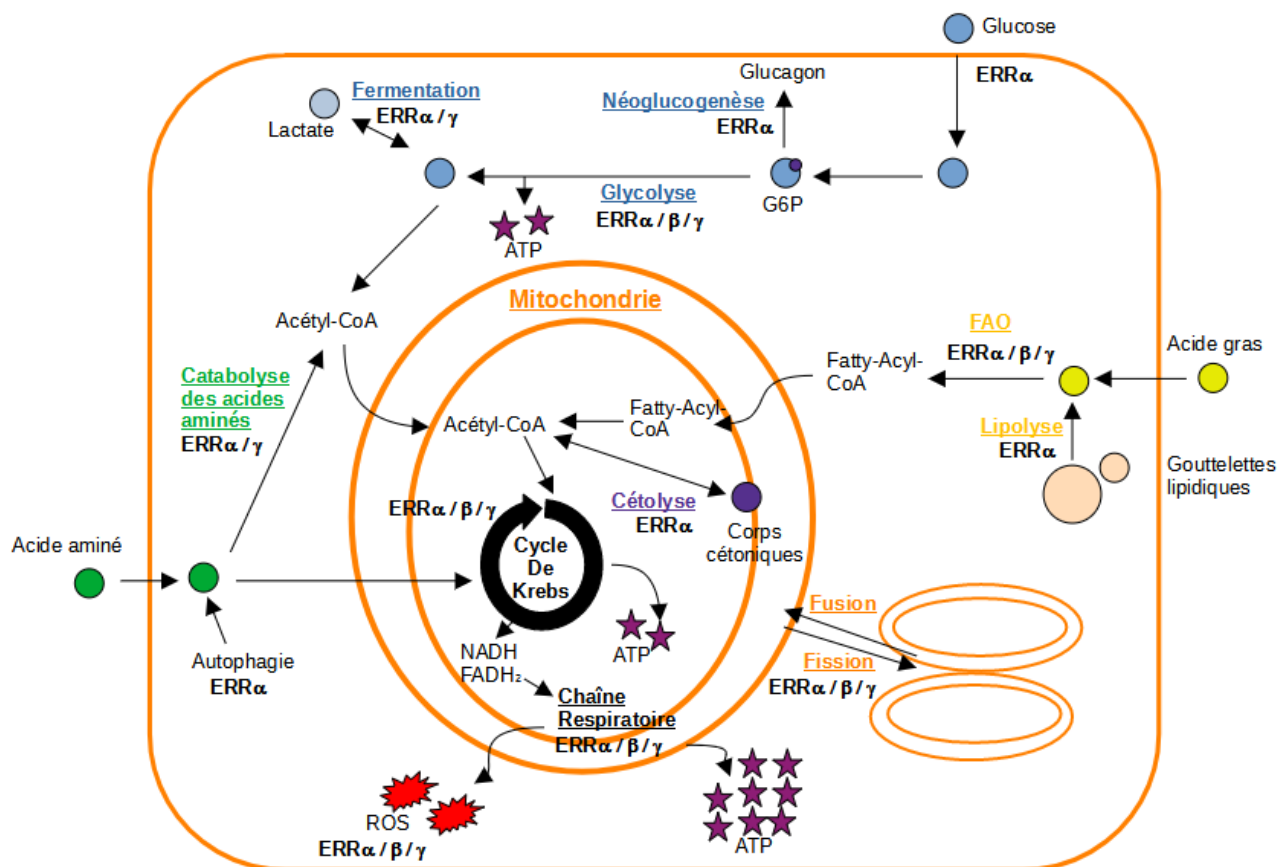


Figure 2: Schéma récapitulatif de l'action des ERR sur le métabolisme énergétique

Légende :

Différentes voies métaboliques permettent de faire fonctionner le cycle de Krebs et la chaîne respiratoire. Les ERR activent ces différentes voies selon le contexte cellulaire. Ils permettent aussi d'activer la fusion et la biogenèse des mitochondries nécessaires à la mise en place du cycle de Krebs.

2. Le potentiel de ERR α et ERR γ dans la lutte contre le cancer

Comme vu précédemment, les récepteurs nucléaires ERR α et ERR γ régulent un grand nombre d'enzymes impliquées dans les voies clés du métabolisme énergétique. De fait, leur dérégulation conduit à de nombreuses pathologies et cancers.

2.1. Le rôle des ERR en tant qu'outil de diagnostic

La surexpression de ERR α dans une large variété de lignées tumorales où ERR γ est faiblement exprimé est associée à la prolifération cellulaire et à l'émergence de capacités invasives, ce qui rend ces tumeurs particulièrement agressives (Cavallini *et al.*, 2005 ; Fujimura *et al.*, 2007; Fujimoto & Sato., 2009). Cela s'explique notamment par le rôle de ERR α dans la promotion de l'effet Warburg, effet qui permet le passage d'un métabolisme aérobie à un métabolisme reposant sur la fermentation lactique (Cai *et al.*, 2013). ERR α régule positivement un ensemble de gènes impliqués dans les fonctions oxydatives de la mitochondrie, ce qui favorise la progression du cancer du sein. Sa surexpression est donc le signe d'un pronostic défavorable pour le patient, faisant de ce dernier un marqueur

moléculaire pouvant potentiellement servir d'outil de diagnostique (Ariazi *et al.*, 2002 ; Sun *et al.*, 2006).

Dans le cadre du cancer des ovaires, la surexpression de ERR γ est associée à une diminution de la multiplication cellulaire, ce qui fait de ce récepteur un biomarqueur d'intérêt associé à un pronostique favorable pour le patient (Sun *et al.*, 2006).

Cela peut s'expliquer par la relation entre ERR α et ERR γ qui sont tous les deux des régulateurs de la balance Red/Ox au sein des cellules et témoignent de la capacité d'éliminer ou non les ROS produites durant le cycle de Krebs, élément qui s'associe à la survie cellulaire.

2.2. Le rôle des ERR en tant que cible thérapeutique

Les ERR sont des outils de diagnostique, mais aussi des cibles thérapeutiques prometteuses. En effet, dans le cadre du cancer du sein de type HER-2, lorsque apparaît une résistance aux traitements au lapatinib, un inhibiteur de la voie de signalisation HER-2, il s'avère que l'inhibition de l'activité de ERR α cause la perte de la résistance aux traitements et la diminution de la multiplication des cellules. Plus précisément, ERR α réactive la transcription par SHMT2 et promeut ainsi le développement de la résistance au lapatinib dans ce type de cancer en passant outre l'inhibition du récepteur aux œstrogènes. (Li *et al.*, 2020)

Dans le cas du cancer de la prostate, ERR α est impliqué dans l'émergence de résistance aux traitements castrateurs en permettant la biosynthèse *de novo* d'androgènes. (Xu *et al.*, 2020)

Enfin, dans le cas de la forme triple-négative du cancer du sein, la protéine STAT3 cible et active la surexpression de ERR α , promouvant la transition épithéliale-mésenchymale, la migration et l'invasion des cellules cancéreuses (Ma *et al.*, 2019). L'inhibition de ERR α par l'utilisation d'agoniste inverse dans ces cellules supprime leurs caractéristiques invasives et réduit significativement leurs croissances (Gao *et al.*, 2022).

Dans ces différents exemples, ERR α est lié à l'émergence de résistances aux cancers déjà existant. Son inhibition spécifique permet la perte des différentes résistances étudiées, si bien que cette protéine se démarque comme cible thérapeutique prometteuse.

3. Aspects structuraux et fonctionnels des ERR

3.1. Organisation générale des récepteurs nucléaires

Les récepteurs nucléaires sont organisés selon un enchaînement conservé de quatre à six domaines protéiques, adoptant une structure modulaire (Krust *et al.*, 1986) (**Fig. 3**).

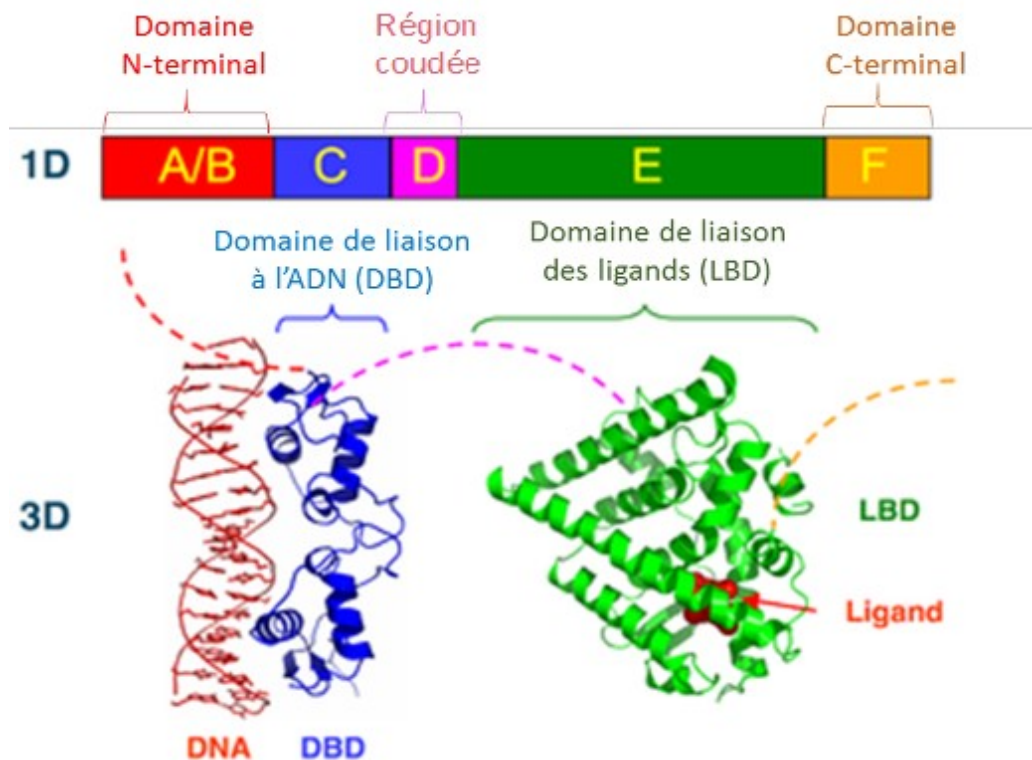


Figure 3: Organisation modulaire des récepteurs nucléaires
 (produit par Hoffmeier sur Pymol, publié sur wikipedia.org)

Légende :

Structure du facteur nucléaire hépatocytaire 4 (HNF-4) (PDB 4IQR) utilisée pour illustrer la structuration modulaire des récepteurs nucléaires qui se constitue du domaine N-terminal, du domaine de liaison à l'ADN, de la région coudée, du domaine de liaison des ligands et du domaine C-terminal (Chandra et al., 2013).

3.1.1. Le domaine N-terminal

Le domaine N-terminal, aussi appelé domaine A/B, est le moins conservé des domaines au sein des récepteurs nucléaires, présentant une taille et une composition en acides aminés très variable. On peut citer le récepteur à la vitamine D (VDR) où ce dernier ne fait que 20 résidus alors qu'il en fait plus de 600 au sein du récepteur aux minéralocorticoïdes (MR). Son rôle physiologique est lié à la régulation de l'activité des récepteurs, aussi est-il le siège de nombreuses régulations post-traductionnelles dont des phosphorylations, acétylations ou sumoylations ayant lieu au niveau de la fonction d'activation 1 (AF-1) sous l'action d'enzymes (Kumar & Brad Thompson, 2003 ; Anbalagan et al., 2012). Cette fonction d'activation est indépendante de la fixation de ligands et peut, dépendamment du NR, induire une augmentation ou une diminution de l'activité transcriptionnelle (Warnmark et al., 2003). De plus, plusieurs formes de ce domaine peuvent être obtenus en fonction de l'épissage, effet qui a notamment été observé chez les récepteurs GR et TR (Weikum et al., 2018).

Historiquement, le nom de ce domaine A/B vient de sa duplication observée au sein du récepteur nucléaire ER, d'où, pour ce dernier, l'appellation distincte de domaine A et B pour chacune des occurrences de ce domaine (Krust et al., 1986).

Structurellement, le domaine N-terminal n'est pas connu et les programmes de prédiction de structures le prédisent, pour l'ensemble des récepteurs nucléaires, comme intrinsèquement déstructuré. Les données de résonance magnétique nucléaire (RMN) indiquent également que la fonction AF-1 ne serait pas structurée en solution (Dahlman-Wright & McEwan, 1996). Néanmoins, ce dernier tend à se structurer sous la forme d'hélices alpha chez GR en présence de corégulateurs (Lavery & McEwan, 2005).

3.1.2. Le domaine de liaison à l'ADN (DBD)

Le domaine de liaison à l'ADN est le plus conservé des domaines, comprenant entre 70 et 80 résidus d'acides aminés (Mark Danielsen, 2001). Sa fonction consiste en la reconnaissance des éléments de réponses spécifiques de l'ADN propres à chaque récepteur nucléaire et en sa fixation sur ces derniers. Structurellement, les analyses par diffraction aux rayons-X et par Résonance Magnétique Nucléaire (RMN) révèlent que ce domaine contient deux motifs en doigts de zinc de type C2-C2 stabilisés par quatre cystéines (Freedman *et al.*, 1988 ; Hinrich Gronemeyer & Dino Moras, 1995). Il contient également deux hélices amphiphiles, lesquelles sont perpendiculaires l'une par rapport à l'autre (Hard *et al.*, 1990 ; Helsen *et al.*, 2012) (**Fig. 4**).

On peut subdiviser en plusieurs boîtes fonctionnelles ce domaine pour expliquer son fonctionnement (Giguère, 1999). La boîte P, dite boîte « proximale », contient le doigt de zinc qui permet l'interaction spécifique avec l'ADN au niveau du grand sillon (Schwabe *et al.*, 1993), tandis que le second doigt de zinc est impliqué dans la dimérisation des DBD au sein de la boîte D, dite boîte « distale » (Khorasanizadeh & Rastinejad, 2001 ; Luisi *et al.*, 1991). Cette interaction avec l'ADN peut également faire intervenir une extension C-terminale au domaine DBD (CTE) qui interagit avec le petit sillon de l'ADN (Solomon *et al.*, 2005 ; Weikum *et al.*, 2016). Cette extension se décompose en deux boîtes : la boîte T et la boîte A.

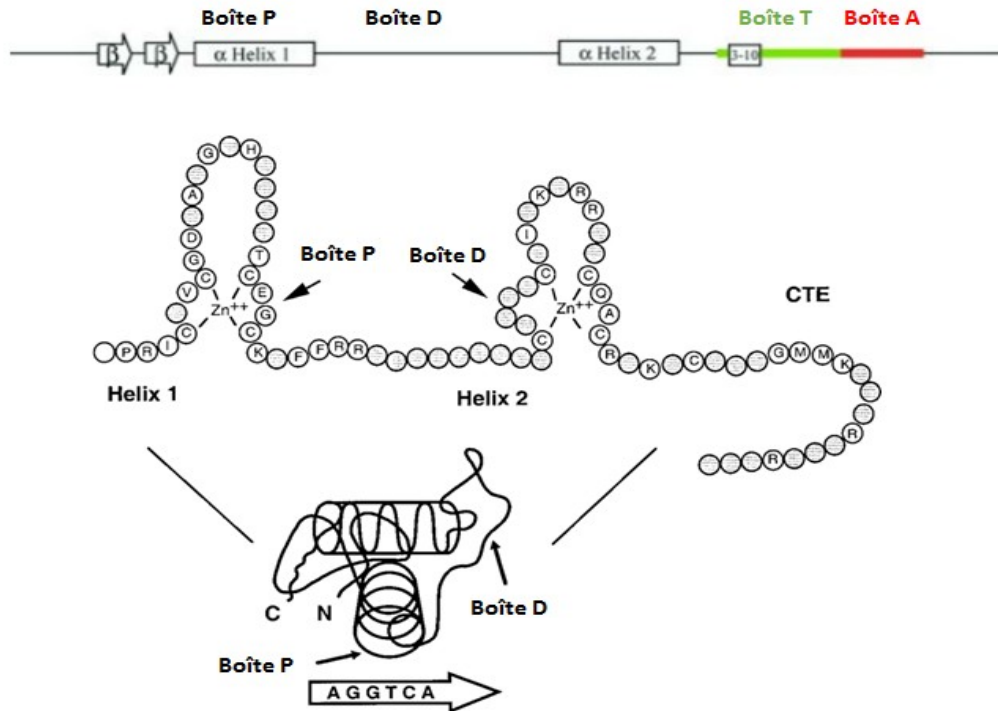


Figure 4: Organisation fonctionnelle et structurale des DBD
(selon Ana Aranda et Angel Pascual, 2001)

Légende :

Repliement caractéristique des DBD et des différentes boîtes fonctionnelles qui le constituent.

Les séquences d'ADN que reconnaissent les DBD sont appelées « éléments de réponse » et elles se composent de deux « demi-sites » dont la séquence canonique est « AGGTCA », sauf chez les oxostéroïdiens dont la séquence canonique est « AGAACA ». La discrimination entre ces deux séquences s'opère par l'intermédiaire de trois résidus d'acides aminés au sein de la boîte P correspondant au motif « Gly-Ser-Val » pour les oxo-stéroïdiens et « Glu-Gly-Ala » pour les autres récepteurs nucléaires qui reconnaissent le demi-site « AGGTCA » (Luisi *et al.*, 1991 ; Gronemeyer & Laudet, 1995). Dans la plupart des cas, chacune des sous-unités du domaine DBD qui constituent le dimère actif va se fixer sur l'un des demi-sites. Un élément de réponse donné se caractérise par la composition en bases de ses deux demi-sites, leur orientation l'un par rapport à l'autre et par l'espacement, compris le plus souvent entre zéro et six paires de bases (bp), entre ses demi-sites. Ces éléments vont définir la spécificité de reconnaissance du dimère de récepteurs nucléaires considéré pour la séquence donnée.

Les répétitions directes de deux demi-sites (DR) sont généralement des éléments de réponse reconnus par les récepteurs de la sous-famille des récepteurs à l'acide rétinoïque sous forme hétérodimérique ou, plus rarement, sous forme homodimérique dans le cas du facteur nucléaire hépatocytaire 4 (HNF-4) (Chandra *et al.*, 2013). En règle générale, dans le cas d'hétérodimères, la position des récepteurs est polarisée sur l'élément de réponse. Ainsi pour l'hétérodimère RXR-RAR, RXR occupe le demi-site en 5' sur les éléments de réponses de type DR2 et DR5, c'est à dire une répétition directe avec un espacement de deux et cinq paires de bases respectivement, alors qu'il occupe le demi-site en position 3' sur un élément de type DR1 (Rastinejad *et al.*, 2000 ; Helsen et

Claessens, 2014). Cette différence de polarité entre éléments de réponses est susceptible d'induire des spécificités de régulation au niveau transcriptionnel. Ces hétérodimères sont généralement liés sur leurs éléments de réponses dans le noyau et se voient réprimés par un ou plusieurs corepresseurs (CoR) jusqu'à ce qu'un ligand interagissent avec eux (**Table 1, Fig. 5**).

Les répétitions inverses de deux demi-sites, qui se présentent donc sous la forme de palindromes, sont des éléments de réponses reconnus par les récepteurs de la famille des stéroïdiens homodimériques. Dans le cas spécifique des oxostéroïdiens et des récepteurs aux œstrogènes, les éléments de réponses sont des répétitions inverses séparées de strictement trois paires de bases (O'Malley & Tsai, 1992 ; Schwabe *et al.*, 1993 ; Roemer *et al.*, 2006 ; Helsen *et al.*, 2012 ; Hudson *et al.*, 2014). Les récepteurs se liant sur de tels éléments de réponse se trouvent dans le cytoplasme sous forme inactive et monomérique, liés à des protéines de choc thermique (HSP). Lorsque le ligand pénètre la cellule et interagit avec le récepteur, ce dernier passe sous une conformation active, se décroche des HSP, puis dimérise avant de rentrer dans le noyau où il se fixera à son élément de réponse (**Fig. 5**). Ce type de mécanisme fut d'abord imaginé en 1971 par Bert O'Malley (O'Malley, 1971) avant d'être confirmé expérimentalement par la suite dans le cas de l'étude du récepteur aux œstrogènes (Yamamoto & Alberts, 1975 ; PrattWB. *et al.*, 2004).

Certains récepteurs nucléaires dits orphelins reconnaissent des éléments de réponses atypiques. Ces derniers se composent généralement d'un seul demi-site étendu de quelques paires de bases, généralement trois, en position 5' (Vanacker *et al.*, 1999). De part leur nature particulière, de tels sites peuvent également être reconnus par d'autres récepteurs s'ils comportent également un second demi-site en répétition directe ou inverse.

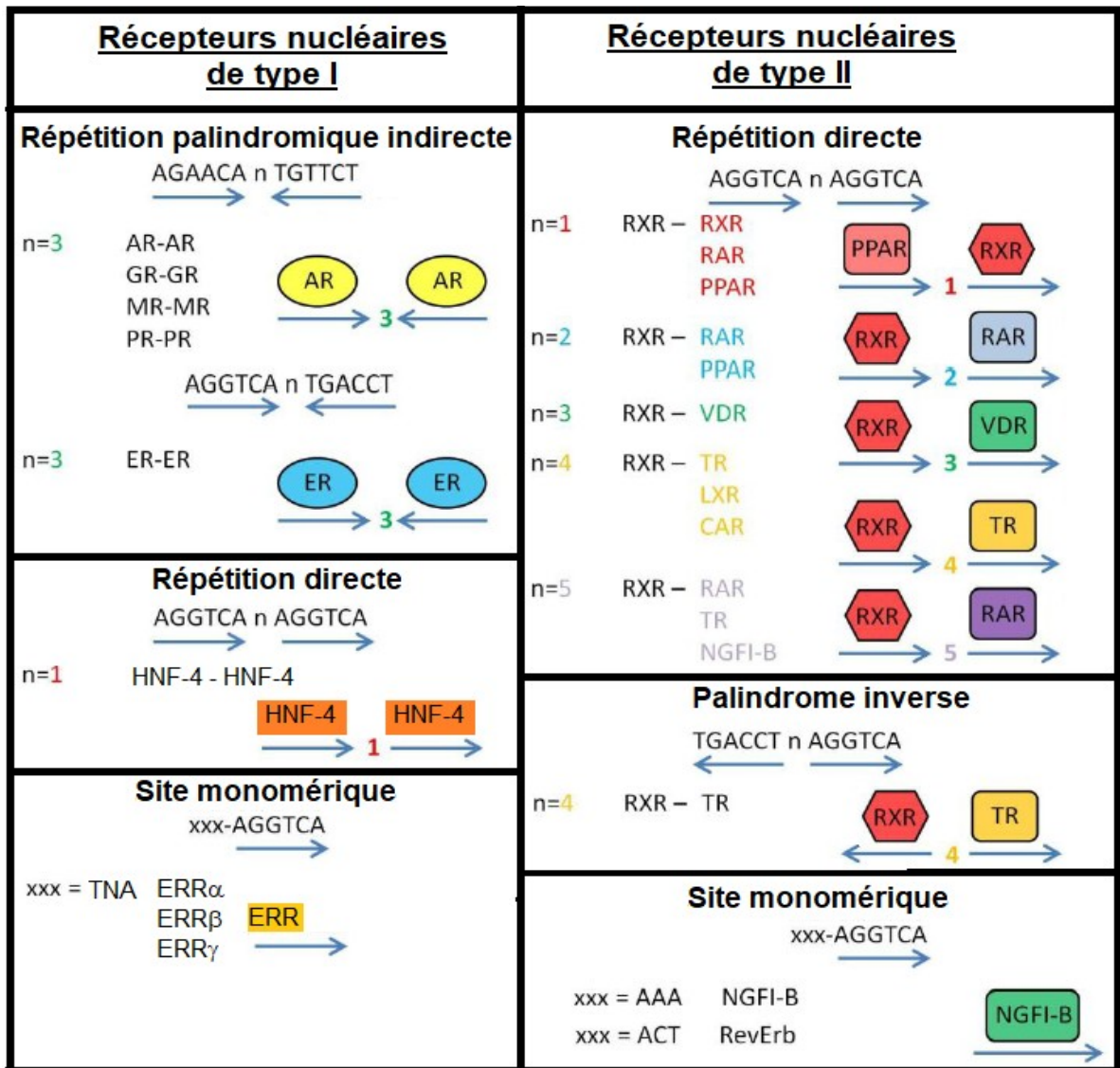


Table 1: Récapitulatif du type d'élément de réponse par récepteur et par classe
 (selon la thèse « Origine et évolution des récepteurs nucléaires et étude structurale du premier stéroïdien, ERR » de Brice Beinsteiner, (2018))

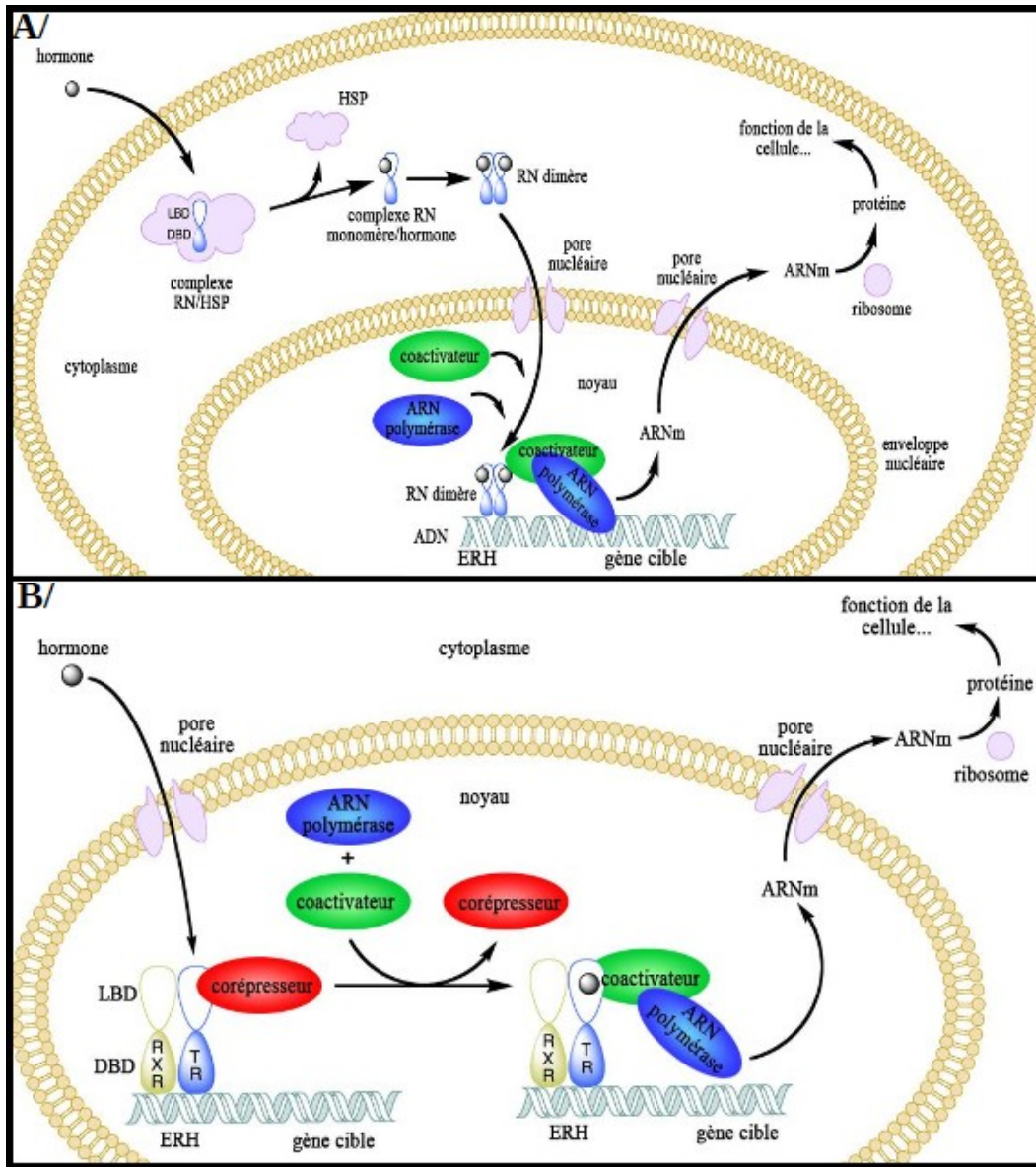


Figure 5: Schéma du mécanisme général des récepteurs nucléaires

Légende :

A/ Mécanisme d'action général d'un récepteur de classe I ligand-dépendant (homodimère). Le récepteur est présent au sein du cytoplasme lié à une HSP. La fixation d'un ligand permet sa dimérisation et son entrée dans le noyau pour se fixer à son élément de réponse.

B/ Mécanisme d'action général d'un récepteur de classe II ligand-dépendant (hétérodimère). Le récepteur est déjà fixé sur son élément de réponse mais maintenu inactif par des corépresseurs. La fixation d'un ligand permet le passage en conformation active, le détachement des corépresseurs et le recrutement de coactivateur. Ce mécanisme d'action est également similaire à celui de $ERR\alpha$, bien que celui-ci soit homodimérique.

La classification des récepteurs en classes en fonction de leur état de dimérisation a été mise évidence par Brelivet (Brelivet *et al.*, 2004). Il est en effet possible de classer les récepteurs en deux classes distinctes. La classe II regroupe les récepteurs hétérodimériques impliquant RXR comme partenaire, mais elle exclue RXR en lui-même qui appartient à la classe I. Elle se définit par la formation d'un pont salin spécifique entre la boucle H4-H5 et la boucle H8-H9 des LBDs des récepteurs nucléaires permettant le rapprochement de l'hélice H7 de RXR. La mutation de ce pont salin n'est cependant pas suffisante pour empêcher la dimérisation et il est suggéré que ce motif joue plutôt un rôle dans le recrutement spécifique de certains coactivateurs. Les récepteurs de classe I ne disposent pas de ce motif signature et incluent RXR sous forme homodimérique, $ERR\alpha$, mais aussi tous les récepteurs nucléaires présent au sein de l'organisme *C. elegans*. (**Table 1**).

3.1.3. La région coudée

La région coudée est prédite comme flexible et sa taille varie d'un récepteur à l'autre, permettant un plus ou moins grand degré de liberté du positionnement des LBDs vis-à-vis des DBDs (Pawlak *et al.*, 2012). Les quelques rares données structurales à son sujet concernent des récepteurs hétérodimériques impliquant le récepteur-X des rétinoïdes alpha ($RXR\alpha$) et $PPAR\gamma$. Cette région présente des variations structurales allant d'une boucle à des morceaux d'hélices alpha (codes PDB : 3DZU, 3DZY, 3E00) (Chandra *et al.*, 2008). Il est cependant difficile d'en tirer une conclusion générale, cette région étant peu conservée et présentant probablement peu de structuration en solution de part sa grande sensibilité aux protéases. Cette flexibilité permettrait par ailleurs à RXR de lier différents types d'éléments de réponses (Mangelsdorf *et al.*, 1995). Elle est également le siège d'un certain nombre de régulations post-traductionnelles ou de régulations induites par l'interaction avec d'autres protéines. On peut citer en exemple le récepteur AR pour lequel la région coudée peut être acétylé (Fu *et al.*, 2000) et contient un signal de localisation nucléaire, ainsi qu'une faculté de modulation de l'interaction avec l'ADN (Haelens *et al.*, 2007).

3.1.4. Le domaine de liaison des ligands

Le domaine de liaison des ligands constitue l'un des deux domaines fonctionnels, avec le DBD, des récepteurs nucléaires et le second en terme de conservation au sein de cette grande famille. Il dispose de certains éléments caractéristiques.

Tout d'abord, son repliement se trouve être un triple feuillet de onze à douze hélices alpha antiparallèles entre-elles et comportant un feuillet bêta antiparallèle entre les hélices H5 et H6. Les hélices H4, H5, H6 et H8 sont situées entre les feuillets formés d'une part par les hélices H1, H2 et H3 et d'autre part par les hélices H7, H10, H11 et H12 (Wurtz *et al.*, 1996 ; Moras & Gronemeyer, 1998). Ce type de repliement caractéristique se nomme un « sandwich d'hélices alpha antiparallèles » (**Fig. 6**). Ce domaine dispose d'un motif caractéristique qui a été identifié entre les hélices H3 et H5. Ce dernier se trouve être [(F,W)AKX₄FX₂LX₃DQX₂LL] et contient de nombreux résidus hydrophobes nécessaires au bon repliement du domaine protéique (Wurtz *et al.*, 1998).

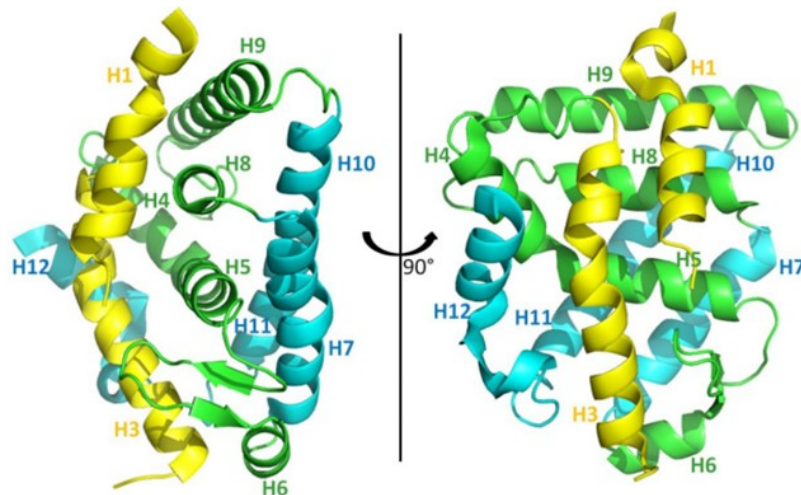


Figure 6: Organisation structurale des LBD

Légende :

Les domaines LBD des récepteurs nucléaires s'organisent en trois feuillets d'hélices alpha antiparallèles. Ici, le LBD de RXR est montré à titre d'exemple (Bourguet *et al.*, 2000). On y observe le feuillet jaune formé par les hélices H1, H2, H3, le feuillet central en vert formé par les hélices H4, H5, H6, H8, H9, et le feuillet cyan formé par les hélices H7, H10, H11 et H12.

Ce domaine est le siège principal de la dimérisation des récepteurs nucléaires, dimérisation nécessaire pour l'activation du récepteur dans la majorité des cas. L'interface de dimérisation des LBDs se trouve le long du plan formé par les hélices H10 et H11, impliquant également l'hélice H7 et la boucle entre les hélices H8 et H9 (Bourguet *et al.*, 1995, 2000 ; Gampe *et al.*, 2000 ; Egéa *et al.*, 2002 ; Svensson *et al.*, 2003 ; Chandra *et al.*, 2013). Globalement, la dimérisation peut se faire selon une interface d'interaction symétrique (récepteurs de classe I, comprenant les récepteurs qui homodimérisent) ou asymétrique (récepteurs de classe II, comprenant les récepteurs qui hétérodimérisent avec RXR). Les récepteurs de classe I disposent d'un ensemble de résidus conservés qui définissent une voie de communication entre l'hélice H12 et l'interface de dimérisation via l'hélice H1 et la boucle H9-H10 (Brelivet *et al.*, 2004, Billas *et al.*, 2013). L'homodimérisation a été identifiée comme le mécanisme ancestral de dimérisation des récepteurs nucléaires (Beinsteiner *et al.*, 2021 et 2022). Les récepteurs de classe II disposent d'une arginine conservée sur leur boucle H8-H9 capable de former un pont salin intramoléculaire avec un résidu E/D conservé présent dans la boucle H4-H5. L'interface de dimérisation est déplacée, permettant le rapprochement de l'hélice H7 de RXR de la boucle H8-H9 du partenaire et la formation d'une interface de dimérisation asymétrique. Le π -turn de l'hélice H7 de RXR a également été démontré comme nécessaire à la stabilisation de l'hétérodimère entre RXR et son partenaire. En effet, cet élément structural spécifique à RXR et HNF-4 permet la stabilisation de son hélice spécifique H10 et l'interaction avec la boucle H8-H9 du partenaire (Beinsteiner *et al.*, 2021). Cette interface est plus stable que lors de la formation d'un homodimère de RXR (Bourguet *et al.*, 2000).

Le nom de ce domaine vient historiquement de sa capacité à reconnaître et permettre la fixation de ligands. Ce ligand permet l'activation de la fonction d'activation ligand-dépendante portée par l'hélice H12 en se plaçant au niveau de la poche de liaison des ligands (LBP) (Johnson *et al.*, 2000). L'entrée du ligand au sein de ce domaine s'effectue selon deux mécanismes connus. D'une part, un mécanisme

qualifié de « piège à souris », surtout montré au sein des RXR, qui implique le repliement de l'hélice H12 lors de l'entrée du ligand au sein de la LBP (**Fig. 7**). Une fois le ligand fixé, le déplacement de l'hélice H12, qui est un effet allostérique, active la fonction d'activation ligand-dépendante de cette hélice et permet l'activation du récepteur (Renaud *et al.*, 1995 ; Bourguet *et al.*, 1995 ; Egea *et al.*, 2002). D'autre part, la stabilisation dynamique de l'hélice H12 s'effectue lors de la liaison du ligand et/ou d'un coactivateur (PDB : 4PLE) (Musille *et al.*, 2016). Le réarrangement des hélices H3, H11 et H12 offre alors une nouvelle surface d'interaction pour des coactivateurs (Renaud *et al.*, 1995 ; Moras & Gronemeyer, 1998).

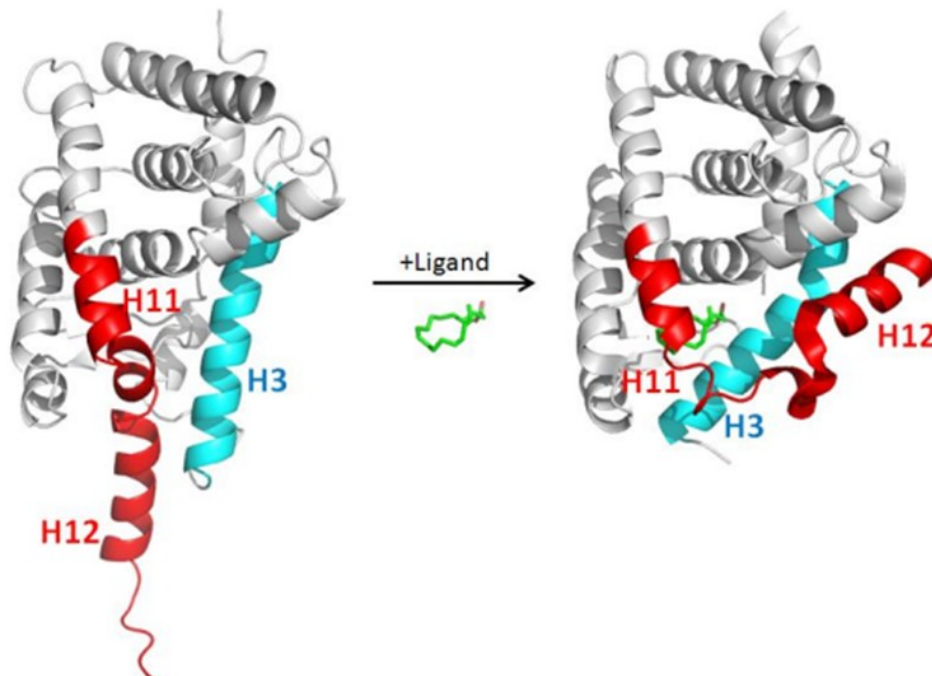


Figure 7: Mécanisme d'action du ligand sur le LBD de RXR

Légende :

Le mécanisme du piège à souris est ici montré sur RXR (PDB : 1LBD et 1DKF, figure réalisée sur Pymol). Sous l'influence du ligand, les hélices H3 et H11 se réorganisent afin de former une poche que l'hélice H12 vient refermer.

Pour la plupart des récepteurs, les chaînes latérales des résidus composant la LBP définissent la spécificité dans la reconnaissance des ligands en jouant sur différents paramètres biochimiques comme leurs contacts hydrophobes ou leur charges relatives (Wurtz *et al.*, 1996 ; Huang *et al.*, 2010). La taille de la LBP définit aussi sa capacité à reconnaître certains ligands et varie d'un récepteur à l'autre. En effet, on peut noter que la LBP dispose d'un volume de 30 Å³ pour le récepteur DHR38 alors qu'elle peut atteindre un volume de plus de 1300 Å³ chez le récepteur activé par les proliférateurs de peroxyosomes (PPAR) (Li *et al.*, 2003). La taille de la poche peut toutefois s'agrandir en présence du ligand par comparaison à la forme apo de la protéine. C'est notamment le cas du récepteur VDR (Mizwicki *et al.*, 2004).

Enfin, tous les récepteurs ne lient pas de ligand au niveau de leur LBD. Si certains récepteurs nucléaires lient une molécule ou une classe de molécule spécifique et sont effectivement qualifiés de

récepteurs « ligand-dépendant », d'autres récepteurs nucléaires sont qualifiés de récepteurs « orphelins » (Giguere *et al.*, 1988). Parmi ces derniers, des récepteurs ont été découverts sans ligand et celui-ci ne fut découvert que plus tard. On parle alors de récepteurs « adoptés ». D'autres encore sont capables de lier plusieurs molécules d'une même famille chimique de façon peu spécifique à leur LBD et sont qualifiés de récepteurs « senseurs ». Enfin, certains récepteurs n'ont pas de ligand et sont qualifiés de récepteurs « orphelins ». Parmi les 48 membres de cette famille chez l'Homme, 16 sont des récepteurs orphelins, 12 sont « d'accueil », 8 sont « adoptés » et 12 sont « ligand-dépendant » (Giguère *et al.*, 1999 ; Mullican *et al.*, 2013).

3.1.5. Le domaine F

Ce domaine, aussi appelé domaine C-terminal, n'est pas présent chez tous les récepteurs nucléaires, mais chez les oxostéroïdiens ou encore ER. Il se trouve dans la continuité du LBD et sa présence se révèle bien souvent hautement stabilisatrice, comme pour RAR ou HNF-4, ou au contraire hautement déstabilisatrice pour l'interaction entre le LBD et son ligand. Elle peut également intervenir dans l'interface de dimérisation comme observé chez ER, où il adopte une conformation d'hélice α suivie d'un feuillet β (Skafar & Zhao, 2008), ou encore dans la reconnaissance de coactivateurs et/ou corépresseurs (Chen & Evans, 1995). Dans le cas des oxostéroïdiens, ce dernier prend la forme d'un feuillet β (Williams et Sigler, 1998). Son rôle précis demeure à élucider.

3.2. Positionnement des ERR parmi les récepteurs nucléaires

D'abord considérés comme étant des isoformes d'un même récepteur, les ERR sont une famille de trois récepteurs différents codés par trois gènes différents. Cette confusion put être entretenue par leur très forte identité de séquences. En effet, l'identité entre les DBD de cette famille dépasse les 90 %, là où celle entre les LBD dépasse les 60 % (**Fig. 8**). En revanche, le domaine N-terminal présente moins de 20 % d'identité entre $ERR\alpha$ et celui des deux autres récepteurs, laissant supposer que $ERR\alpha$ est plus éloigné phylogénétiquement que $ERR\beta$ et $ERR\gamma$ (Fang Chen *et al.*, 1999 ; May F.E.B., 2014). Cette famille de récepteurs fut phylogénétiquement classée parmi les récepteurs nucléaires stéroïdiens de part sa proximité avec le récepteur aux œstrogènes (Nuclear Receptors Nomenclature Committee, 1999). La découverte de l'existence d'un récepteur aux œstrogènes ancestral en tant qu'ancêtre commun à l'ensemble des récepteurs stéroïdiens et la mise en évidence des récepteurs ERR comme récepteurs ancestraux du récepteur ER, affina cette classification, faisant des ERR actuels les descendants du plus proche ancêtre commun des récepteurs stéroïdiens (M.E. Baker., 1997 ; J. W. Thornton., 2001 ; Thornton *et al.*, 2003 ; G. V. Markov & V. Laudet., 2011 ; Baker *et al.*, 2015 ; Markov *et al.*, 2017 ; Lathe & Houston., 2018).

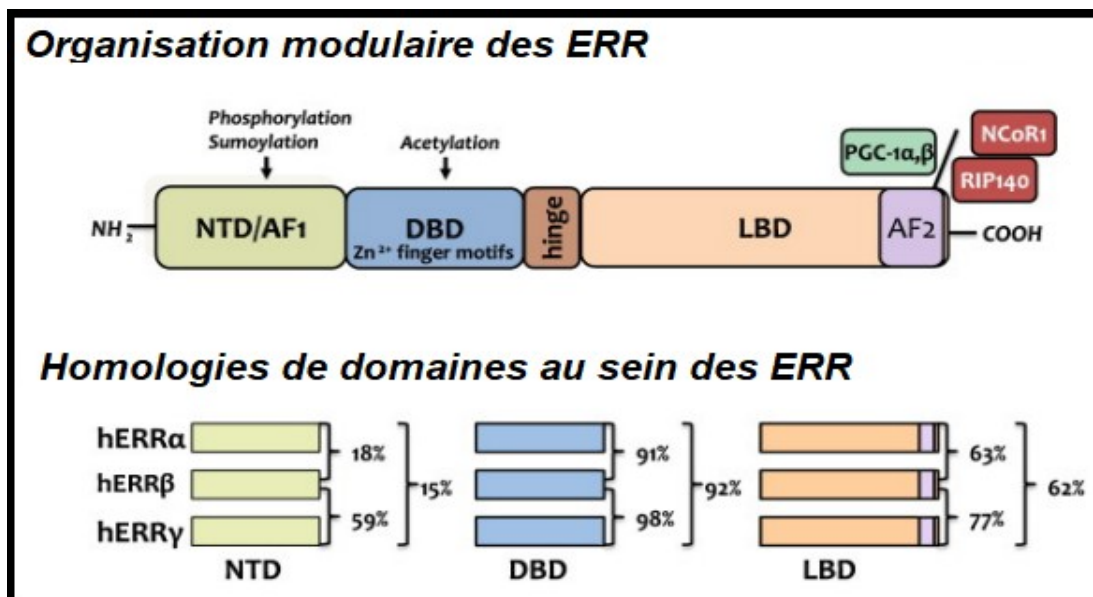


Figure 8: Identité au sein de la famille des ERR
(par Huss et al., 2015)

3.2.1. La régulation du domaine N-terminal des ERR

Le domaine N-terminal des ERR est prédit comme intrinsèquement désordonné d'un point de vue structural, bien qu'il puisse potentiellement s'ordonner au contact d'autres protéines. Il est mieux caractérisé chez ERR α où il est le siège de deux types de modifications post-traductionnelles consécutives : la phosphorylation et la sumoylation. La phosphorylation des résidus Ser19 et Ser22 est suivie par la sumoylation de la Lys14 qui diminue l'activité transcriptionnelle associée à la fixation du récepteur (Vu *et al.*, 2007 ; Tremblay *et al.*, 2008). S'il a pu être montré chez les récepteurs stéroïdiens que ce domaine pouvait réprimer leur activité en contactant le domaine de liaison du ligand ou permettre le recrutement de coactivateurs nécessaires à la formation du complexe d'activation de la transcription, un tel mécanisme n'a pas été montré pour les ERR dont le domaine est plus court que chez les récepteurs oxo-stéroïdiens (Almlof *et al.*, 1998 ; Walberg *et al.*, 2000 ; He B. *et al.*, 2003 ; Wasmuth *et al.*, 2020).

3.2.2. Les particularités du domaine de liaison des ligands de ERR

Les structures des domaines de liaison des ligands des ERR ont été obtenues par diffraction aux rayons X. Ce fut le cas en 2002 pour ERR γ (PDB : 1KV6)(Greschik *et al.*, 2002), puis en 2004 pour ERR α (PDB : 1XB7)(Kallen *et al.*, 2004) et en 2020 pour ERR β (PDB : 6LIT & 6LN4) (Yao & al., 2020). Aucun ligand naturel n'a été cristallisé au sein de la LBP de leurs LBD et ces derniers, de part le positionnement des résidus aromatiques composant leur LBP, sont maintenus dans une conformation active. Cela ne signifie cependant pas que les LBDs ne sont pas régulés.

A l'exception de ERR γ , chacun des LBDs a été cristallisé grâce à l'utilisation de peptides appartenant à un coactivateur, tel que le coactivateur des récepteurs stéroïdiens (SRC-1) ou PGC-1 α . Ce peptide correspond à un motif « LXXLL » que l'on retrouve dans les séquences des coactivateurs des récepteurs nucléaires. Structurellement, il se présente sous la forme d'une hélice alpha interagissant

principalement par l'intermédiaire d'effets hydrophobes avec les LBDs (Heery *et al.*, 1997), mais aussi par la présence d'une lysine et d'un glutamate conservés, respectivement sur les hélices H3 et H12 des récepteurs nucléaires, stabilisant la position du coactivateur par la formation de liaisons ioniques en amont et en aval du motif « LXXLL » (Darimont *et al.*, 1998). Dans le cas des ERR, il a en plus été montré que ces derniers favorisaient l'interaction avec le motif « LLXYL » en plus du motif canonique « LXXLL » présent sur le coactivateur PGC-1 α . L'interaction avec ce motif est permise par la longueur et la composition de la boucle H8-H9 des ERR. En effet, le remplacement de cette boucle, dont les résidus sont conservés au sein de la famille des ERR, par la boucle de ER α cause la perte de l'interaction entre le LBD et le motif atypique « LLXYL », sans affecter l'interaction avec le motif canonique « LXXLL ». Néanmoins, la présence seule de cette boucle n'est pas suffisante pour décrire complètement la sélectivité des coactivateurs effectuée par les ERR. En effet, le coactivateur TIF-2 présente également le motif atypique « LLXYL ». Pourtant, ce dernier n'est pas capable de se fixer sur les ERR. La composition en résidus d'acides aminés en position N-terminale du motif atypique sur le coactivateur est déterminante. En effet, les résidus précédant le motif atypique sur PGC-1 α sont Arg205, Pro206, Cys207, Ser208 et Glu209 alors qu'il s'agit des résidus Lys469, Lys470, Glu471, Asn472 et Ala473 sur TIF-2 (**Fig. 9**). Des expériences de double-hybride effectuées avec des mutants du peptide de PGC-1 α ont mis en évidence le rôle clef de la proline Pro206 et de la cystéine Cys207 dans la reconnaissance et la fixation du motif atypique de PGC-1 α par comparaison avec celui présent sur TIF-2 (Greschik *et al.*, 2008).

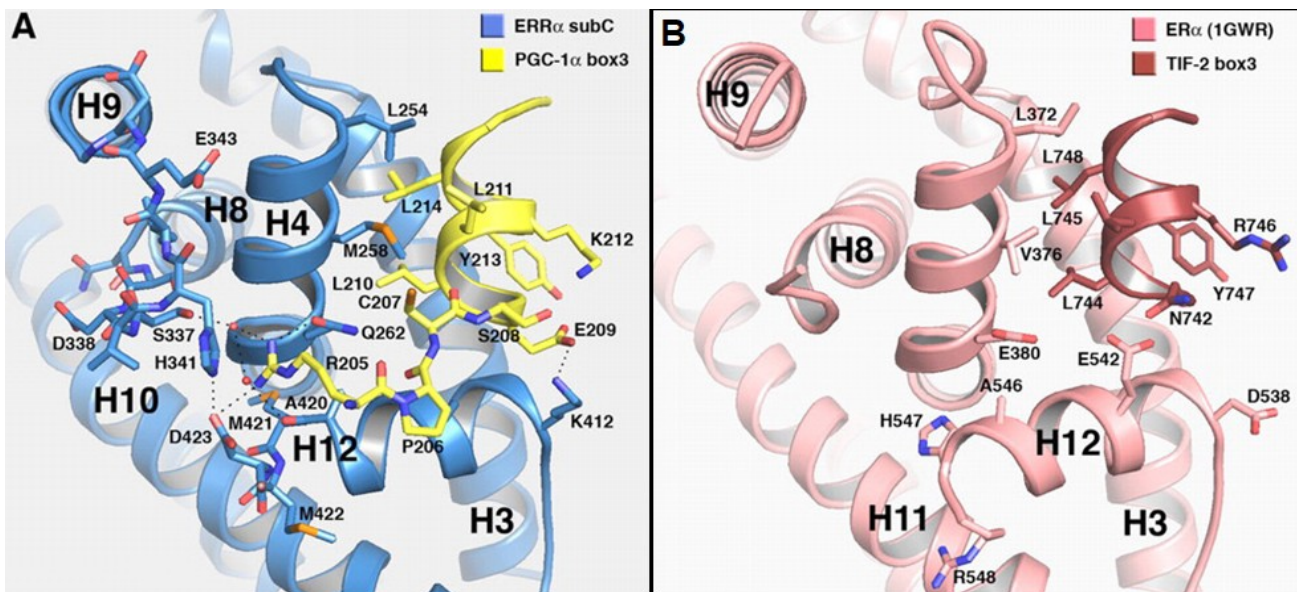


Figure 9: Mécanisme de reconnaissance de PGC-1 α par ERR α comparé à celui de TIF-2 par ER α

(Greschik *et al.*, 2008)

Légende :

A/ Interaction entre ERR α et PGC-1 α (PDB : 3D24). **B/** Interaction entre ER α et TIF-2 (PDB : 1GWR).

En plus des coactivateurs, des molécules synthétiques sont capables de se lier aux ERR et d'affecter leurs fonctions, agissant sous la forme d'agonistes inverses se fixant dans la LBP. Parmi les 22 résidus formant la LBP de ERR α , on peut relever les résidus Phe232, Glu235 et Val491,

respectivement Ala247, Glu250 et Ala406 pour ERR β et Ala287, Glu275 et Ala431 pour ERR γ . La Phe232 et la Val 491 au sein de la LBP de ERR α par comparaison avec les alanines de ces deux homologues conduit à une différence significative en terme de volume de la LBP parmi ces trois récepteurs. En effet, la poche de liaison des ligands de ERR α dispose d'un volume de 100 Å³, là où celle de ERR β fait 197 Å³ et celle de ERR γ occupe 220 Å³. A cela, nous pouvons ajouter l'orientation du résidu Glu235 de ERR α qui, s'il est le même que celui de ERR β ou ERR γ , présente néanmoins une orientation plus favorable à son interaction avec l'Arg276, stabilisant d'autant plus fortement la LBP et nuisant à l'entrée de certains agonistes tels que le BPA.

Dans le cas des LBDs de ERR α cristallisés avec le composé 1-Cyclohexyl-N-{{1-(4-Methylphenyl)-1H-Indol-3-yl}Methyl}Methanamine, le déplacement des résidus en position N-terminale de l'hélice H3 et des résidus en position C-terminale de l'hélice H11 peut être observé depuis l'intérieur de la poche vers l'extérieur de celle-ci (PDB : 2PJL). Le résidu Phe232 spécifique à ERR α se déplace alors, induisant un mouvement de la Phe510 de l'hélice H12, changeant d'orientation pour se positionner dans une position analogue à celle du motif « LXXLL » ou « LLXYL » des coactivateurs. Ce nouveau positionnement de l'hélice H12 empêche la liaison des coactivateurs sur ERR α . Chez ERR γ , le mécanisme permettant ce même changement de conformation est entièrement dépendant de la capacité du ligand à déplacer la phénylalanine de l'hélice H12, la Phe232 n'ayant pas d'équivalent (Kallen J. *et al.*, 2007). La Phe232 de ERR α intervient aussi dans la sélection d'agonistes qui empêchent la liaison du corepresseur RIP140 en modifiant la conformation de la LBP et plus particulièrement de l'hélice H3 (Peng L & al., 2011 ; Shinozuka *et al.*, 2021).

De la même manière, le Bisphénol A (BPA) illustre la différence de sélectivité entre les ERR. En effet, ce dernier se place entre les hélices H3, H7, H10 et H12 de ERR β et ERR γ et maintient les LBDs en conformation active. Cependant, ce dernier n'interagit pas avec ERR α . Cette différence peut s'expliquer structurellement par la composition en résidus d'acides aminés de la LBP de ces récepteurs. En effet, la phénylalanine en position 232 que nous évoquons plus tôt au sein de ERR α correspond respectivement aux alanines 247 et 287 chez ERR β et ERR γ . Or, le positionnement de cette phénylalanine est défavorable à l'interaction du BPA avec le récepteur. De la même manière, la Phe286 est remplacée par une tyrosine chez ERR β et ERR γ , tyrosine qui interagit par l'intermédiaire de son oxygène avec l'un des groupements phénol du BPA. Enfin, l'interaction entre le BPA et ERR β permet le repositionnement de Glu250 en face de l'Arg291, stabilisant un pont ionique au sein de ERR β alors qu'un tel pont existe déjà et est dans une position plus stable au sein de ERR α , faisant intervenir son E331 (Yao B *et al.*, 2020) (**Fig. 10**).

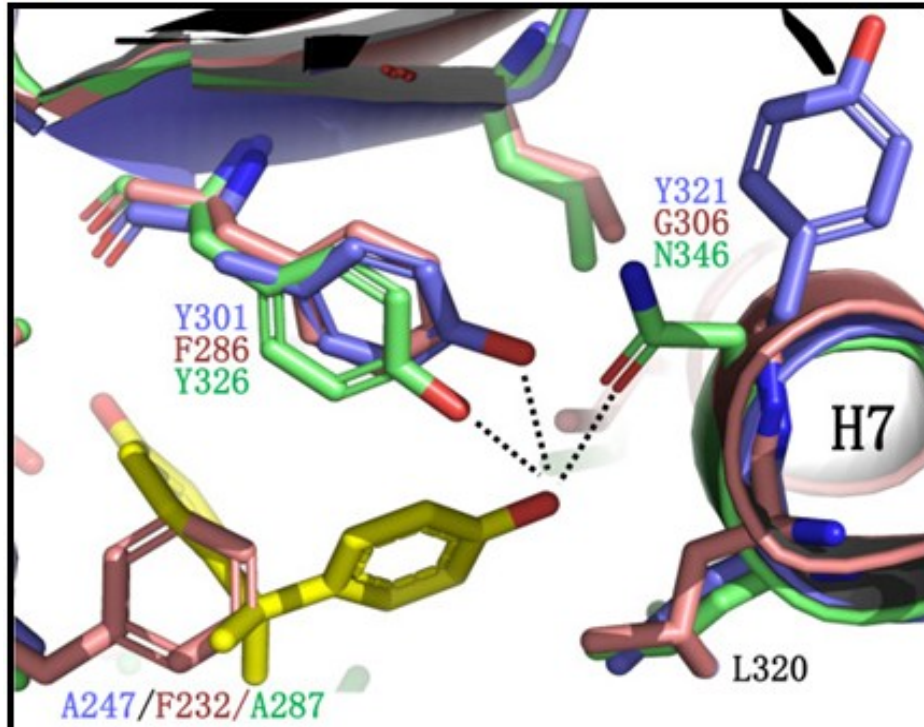


Figure 10: Différence de sélection des agonistes inverses au sein des ERR
(Yao et al., 2020)

Légende :

Le site de fixation du BPA (en jaune) au sein de ERR γ est ici montré. Les trois ERR ont été superposés avec, respectivement, ERR α (rose), ERR β (bleu) et ERR γ (vert).

Concernant le pont formé par le Glu250 et l'Arg291, il dispose d'une autre fonction caractérisée chez ERR γ . Sous l'action d'un ligand synthétique nommé le diethylstilbestrol (DES), cette interaction disparaît pour permettre un réarrangement de la LBP. Le changement conformationnel des chaînes latérales induit un élargissement de la LBP jusqu'à une taille de 330 Å³, permettant l'entrée de la molécule. Un tel mécanisme n'est pas recensé chez ERR α et pourrait s'expliquer par le fait que la « seconde poche » que possède ERR γ et qui fusionne avec la première sous l'action du DES n'existe pas chez ERR α , empêchant la déformation de la LBP et donc l'entrée du ligand (Greschik *et al.*, 2004). Il est par ailleurs notable que les agonistes et ligands peuvent s'affecter positivement ou négativement dans le cadre de l'interaction avec les LBDs de chaque monomère composant le dimère actif d'un récepteur nucléaire. Ce mécanisme a été montré, entre autre, au sein de l'hétérodimère RAR-RXR (Le Maire *et al.*, 2019). Dans le cas de ERR β et ERR γ , le BPA semble pouvoir améliorer l'entrée d'autres agonistes synthétiques dans la poche de liaison des ligands du récepteur.

3.2.3. Le domaine de liaison à l'ADN

Le domaine de liaison à l'ADN des ERR est identique à plus de 70 % avec celui de ER α et à plus de 90 % entre les ERR. Il reconnaît des éléments de réponse aux récepteurs apparentés aux récepteurs aux œstrogènes (ERRE) qui se composent d'un demi-site de fixation étendu de trois paires de bases en 5' : « TNAAGGTCA ». Ce dernier fut identifié par comparaison des sites reconnues par les ERR et par ER, notamment *in vivo* au sein des souches MCF-7. Un élément composite permettant la liaison des

ERR, mais également des ER fut également identifié par cette méthode :

« TNAAGGTCANNNTGACCT », où la seconde base est préférentiellement un C ou un G (Deblois *et al.*, 2009) (**Fig. 11**). Ce type de séquence représente 6 à 10 % des sites reconnus par ERR α .

La structure du DBD monomérique de ERR β , obtenue par RMN, nous indique que l'extension Carboxy-terminale du DBD interagit avec l'extension de trois paires de bases « TNA » de l'élément de réponse et donc avec le petit sillon de l'ADN (Gearhart *et al.*, 2003), comme cela put être observé, plus tard, pour d'autres récepteurs (Solomon *et al.*, 2005 ; Weikum *et al.*, 2016). Néanmoins et contrairement à ce qui était connu avant, il fut montré par des expériences de migration en gel natifs, de SEC-MALS, de RMN et de spectrométrie de masse que les éléments composites permettaient la fixation de dimères de DBD de ERR α (Mohideen-Abdul *et al.*, 2017) (**Fig. 12**). L'hypothèse selon laquelle les paramètres structuraux de l'ADN qui sont séquences-dépendants sont cruciaux pour la fixation du dimère de DBDs a été confirmée par la mutation des séquences d'ADN et le suivi de leurs interactions avec les DBDs.

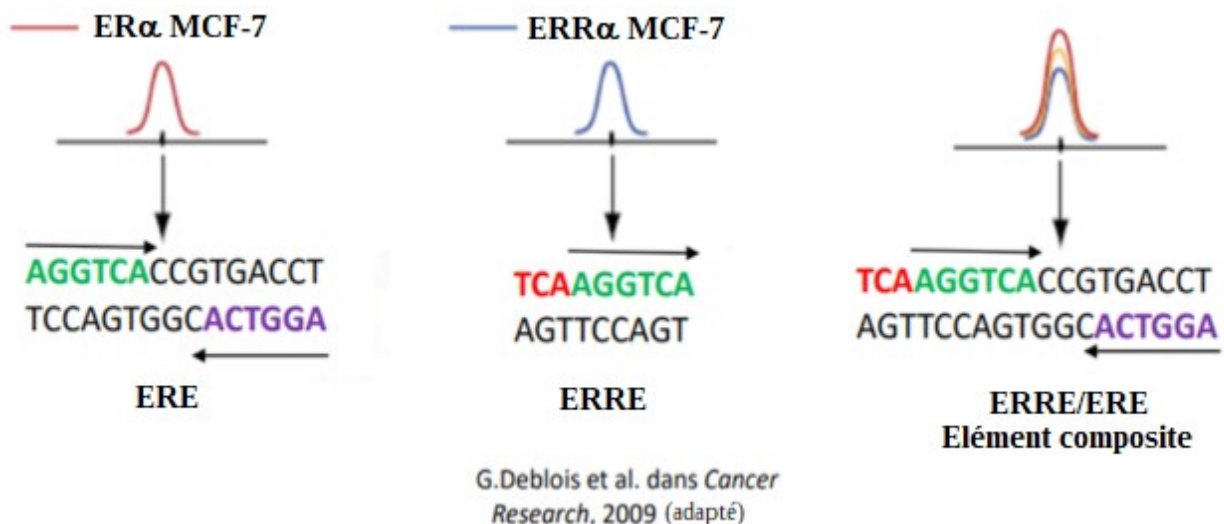
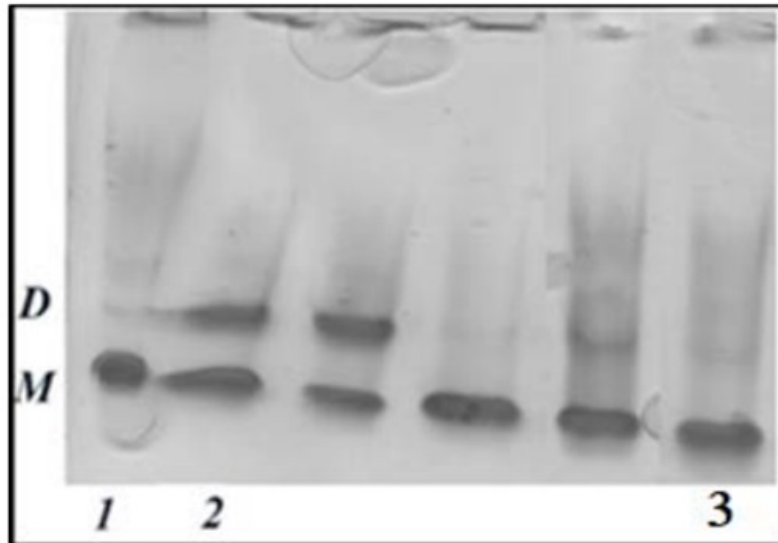


Figure 11: Les sites de liaison des ERR
(adapté de Deblois *et al.*, 2009)

Légende :

Sites de fixation de ER α et ERR α . ERE désigne l'élément de réponse aux récepteurs aux oestrogènes (ERE), ERRE est le demi-site étendu spécifique à ERR α (ERRE) et ERRE/ERE est un élément composite sur lequel ERR α et ER α sont capables de se lier.



K. Mohideen-Abdul et al. dans *Frontiers in Endocrinology*, 2017 (modifié)

Pistes	Nom	Sequences
1	tra ERRE	TTG CAAGGTCA CACAGTGAGTTACT AACAGTTCCAGTGTGTCACCTCAATGA
2	emb ERRE/IR3	ATG CAAGGTCA CCGTGACCTTTACG TACAGTTCCAGTGGC ACTGGAA AATGC
3	tff1 ERE/IR3	CCT GCAAGGTCA CGGTGGCCACCCGC GGACGTTCCAGTGCC ACCGT GGGCG

Figure 12: Comportement de $ERR\alpha$ en fonction des éléments de réponse
(adapté de Mohideen-Abdul et al., 2017)

Légende :

1/ Fixation du DBD sur un ERRE naturel sous forme monomérique (M). 2/ Fixation du DBD sur un élément composite sous forme dimérique (D) et monomérique (M). 3/ Fixation du DBD sur un élément de réponse au récepteur aux oestrogènes sous forme monomérique (M)

Sur un élément de type ERRE, l'interaction se fait au niveau des bases du demi-site « AGGTCA » situé sur le grand sillon de l'ADN (**Fig. 13-A**). Les Lys124 et 128 interagissent avec les bases G8 et T9, mais aussi le résidu Glu121 avec la base C19 du brin complémentaire et l'Arg179 avec la base T21 du brin complémentaire. Ces interactions impliquent donc la boîte P du DBD et le grand sillon de l'ADN, chose commune à d'autres NR (**Fig. 13-B**).

Concernant l'extension en 5' « TCA », l'interaction se fait par les résidus Gly180 et Arg182, lesquels interagissent respectivement avec T3 et C4 pour l'Arg182 et G23 pour la Gly180. Ces résidus font partis d'une boîte que l'on qualifie de boîte A de la CTE et dont la conformation particulière permet l'entrée des glycines du motif « RGGR » au cœur du petit sillon formé par l'extension « TGA ». D'autres résidus sont impliqués dans l'interaction avec l'ADN, mais ces derniers interagissent surtout avec les riboses et/ou le squelette phosphaté (**Fig. 13-C**).

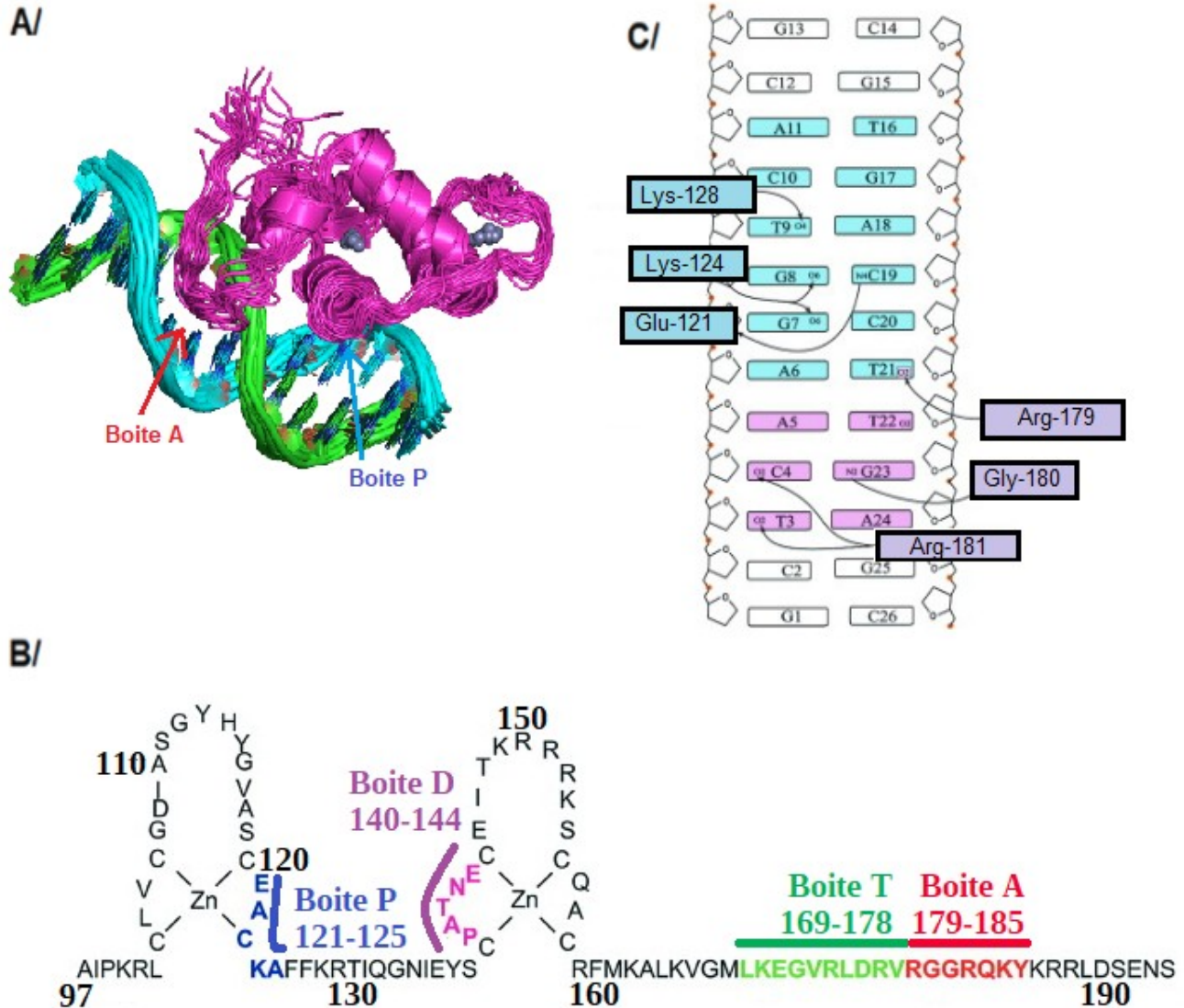


Figure 13: Structure de $ERR\beta$ obtenue par RMN sur un ERRE
(Gearhart et al., 2003)

Légende :

A/ Structure RMN de $ERR\beta$ sur un ERRE de 13 pb. Les boîtes P et A interagissant avec le ERRE y sont annotées. **B/** Séquence du DBD de $ERR\beta$ avec la mise en évidence des boîtes fonctionnelles. **C/** Sites d'interactions du DBD sur le ERRE.

L'intervention de modifications post-traductionnelles régulent l'activité des DBDs au cours du temps. Par exemple, l'acétylation des Lys129, Lys138, Lys160 et Lys162, lorsqu'elles se produisent, réduisent significativement l'activité du récepteur nucléaire. De telles modifications sont apportées, entre autre, par le coactivateur associé au facteur p300 (PCAF), identifiée pour sa capacité à acétyler et inhiber, *in vitro*, l'activité du récepteur (Wilson et al., 2010). De telles modifications n'ont pas été répertoriées pour les récepteurs $ERR\beta$ et $ERR\gamma$.

Point notable, d'après la structure RMN obtenue du DBD de $ERR\beta$, ces lysines, hormis la Lys129 qui interagit avec le squelette de l'ADN, n'interagissent pas avec l'ADN, mais sont plutôt impliquées dans

des interactions intra-moléculaires. Cette régulation agirait donc plutôt sur la structure du domaine que sur la perte directe de son affinité avec l'ADN.

4. Les complexes impliquant les récepteurs nucléaires

Il n'y a, à ce jour, pas de complexes d'activation de la transcription entièrement caractérisé structurellement impliquant les ERR. Il y a néanmoins des complexes de récepteurs nucléaires entiers sur leurs éléments de réponse, ainsi que des complexes impliquant des coactivateurs qui peuvent nous servir de références dans la mise au point des expériences nécessaires à l'accomplissement des objectifs de cette thèse.

4.1. Les complexes de récepteurs nucléaires entiers impliquant l'ADN

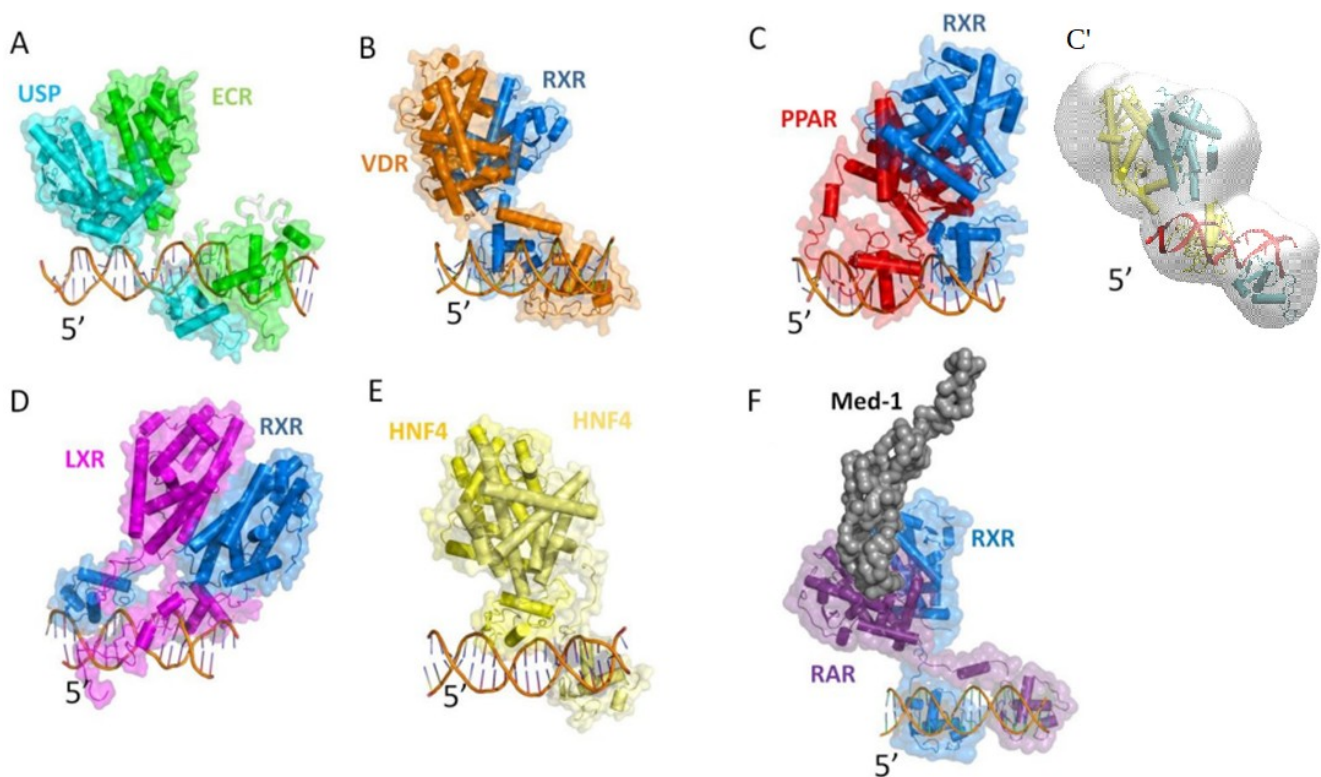


Figure 14: Récepteurs nucléaires sur leurs éléments de réponses

Légende :

A/ USP et ECR sur leur élément de réponse (carte obtenue par cryo-microscopie électronique (cryo-EM), EMDB : 4UMM) (Maletta et al., 2014). B/ VDR et RXR sur leur élément de réponse (cryo-EM) (Orlov et al., 2012). C/ PPAR et RXR sur leur élément de réponse. (PDB : 3DZU) (Chandra et al., 2008). C'/ Enveloppe de PPAR et RXR obtenue par l'analyse de diffusion de rayons X aux petits angles en solution (SAXS) (Osz et al., 2012). D/ LXR et RXR sur leur élément de réponse (PDB : 4NQA). E/ HNF-4 sur son élément de réponse (PDB : 4IQR) (Chandra et al., 2013). F/ RAR et RXR sous l'action du composé Med-1 sur leur élément de réponse (SAXS).

Globalement, les récepteurs nucléaires présentent une structure quaternaire ouverte et en forme de « L » sur leurs éléments de réponse (Chandra *et al.*, 2008 ; Rochel *et al.*, 2011 ; Osz *et al.*, 2012 ; Orlov *et al.*, 2012 ; Chandra *et al.*, 2013 ; Maletta *et al.*, 2014 ; Lou & al, 2014). Cette dernière a le plus souvent été observée avec les LBD décalés en 5' par rapport aux DBDs, bien qu'il existe quelques exceptions impliquant RXR, tel que l'hétérodimère RXR/LXR, lesquels adoptent une forme de « X » avec les domaines LBD orientés en 3' par rapport aux DBD (**Fig. 14**).

Les structures connues impliquant RXR ont permis de montrer l'adaptabilité de ce récepteur dont les complexes formés présentent différentes organisations topologiques asymétriques (Chandra *et al.*, 2008) et dont la différence de conformations est permise par sa région coudée (Chandra *et al.*, 2017). Ces structures ont également permis de mettre en lumière l'existence d'une hétérogénéité de conformation des domaines les uns par rapport aux autres qui peut conduire à des différences fonctionnelles (Orlov *et al.*, 2012). Cette hétérogénéité est suggérée dans le cas du récepteur aux androgènes (AR) (Wasmuth *et al.*, 2022).

Enfin, l'étude cristallographique de HNF-4 a mis en avant le rôle essentiel de sa région coudée et de ses LBDs dans son interaction avec l'ADN, dont des résidus clefs des domaines LBDs et DBDs interagissant entre eux pour maintenir une conformation fixe sur l'élément de réponse et assurer la communication entre domaines (Chandra *et al.*, 2013).

4.2. Les complexes impliquant des répresseurs

Les récepteurs nucléaires n'étant pas des oxo-stéroïdiens se retrouvent dans le noyau sous le contrôle de complexes corépresseurs en l'absence de ligands. Parmi ces corépresseurs, ceux de la famille NCOR et HDAC ont été mis en avant dans la répression de gènes sous le contrôle de RXR (Mangelsdorf DJ. *et al.*, 1995) (**Fig. 15**). Néanmoins, le désordre intrinsèque de ces protéines est l'un des facteurs expliquant l'absence de complexes corépresseurs entièrement caractérisés. Structuellement, les corépresseurs sont porteurs du motif « (L/I)XX(I/V)I » ou «(I/L)XXX(I/L)XXX(I/L)» dont la structuration nuit à la formation des ponts salins impliquant la lysine de l'hélice H3 et le glutamate de l'hélice H12 des récepteurs nucléaires (Hu X. *et al.*, 1999 ; Nagy L. &al., 1999 ; Perissi V. &al., 1999 ; Webb P. &al., 2000, Bourguet *et al.*, 2000).

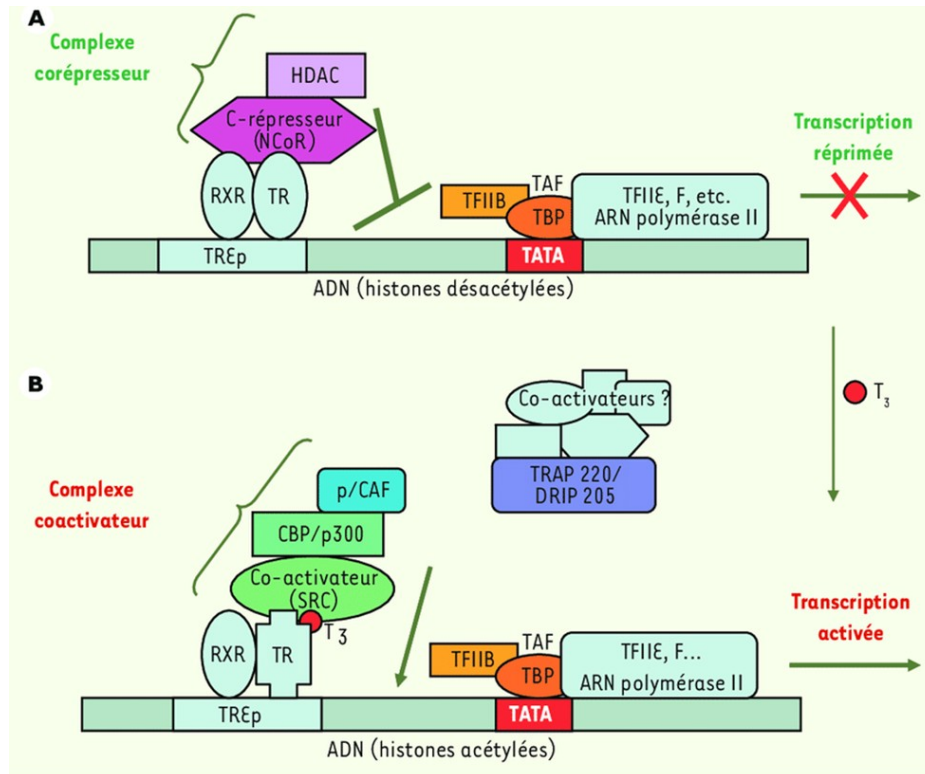


Figure 15: Les complexes de régulation de la transcription impliquant les récepteurs nucléaires
(Guissouma et al., 2005)

Légende :

A/ Le complexe corepresseur est formé par les corépresseurs NCOR et la protéine HDAC afin de maintenir les histones désacétylés et ainsi empêcher la transcription. L'ensemble des protéines formant ce(s) complexe(s) n'est pas totalement identifié aujourd'hui.

B/ Une fois le ligand fixé, les coactivateurs se fixent aux récepteurs, formant un complexe activateur permettant, notamment sous l'action de l'histone acetyltransferase p300 (p300) et/ou CBP, d'acétyler la queue des histones et d'activer la transcription. On parle ici d'un échange de corégulateurs.

4.3. Les complexes impliquant des coactivateurs

Des structures des LBDs des ERR ont pu être obtenues avec un peptide du coactivateur PGC-1 α (Greschik et al., 2008). Cependant, seule la structure du motif d'interaction connue, sous la forme d'une hélice alpha, a pu être obtenue à ce jour, y compris au sein d'autres récepteurs nucléaires. Cela est principalement dû au désordre intrinsèque de ces protéines coactivatrices dont la stabilisation dépend probablement de la présence d'autres protéines impliquées dans ce type de complexes.

Deux complexes impliquant le récepteur ER α et son coactivateur SRC ont été résolus à basse résolution par cryo-EM. Une fois les protéines SRC-3 fixées, elles agissent comme une plate-forme pour le recrutement de la protéine acétyltransférase des histones p300 (p300) dont l'une des fonctions est l'acétylation de la queue des histones et l'activation de la transcription. Un autre complexe, impliquant cette fois la protéine Arginine Methyl-transferase associée au coactivateur 1 (CARM1), a

également été caractérisée par la même équipe et montre une étape supplémentaire dans l'activation de la transcription en proposant que la protéine CARM1 va méthyler la queue des histones séquentiellement après la modification apportée par le complexe impliquant les deux unités de SRC-3 (**Fig. 16**). Néanmoins, la résolution de cette structure ne permet pas à ce jour d'identifier les mécanismes mis en jeu dans de telles interactions (Yi *et al.*, 2015 ; Yi *et al.*, 2017).

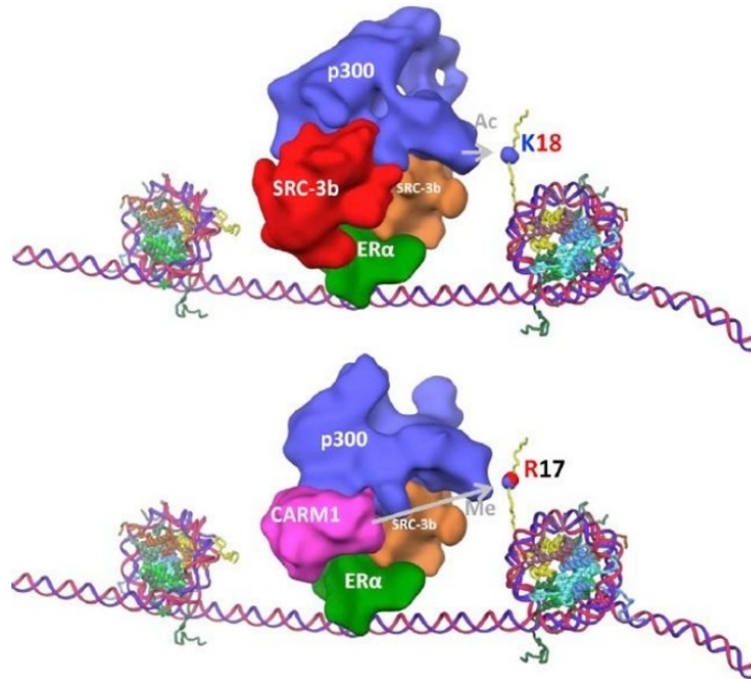


Figure 16: Complexes d'activation de la transcription formés par ER α
(Yi *et al.*, 2020)

Légende :

Le modèle proposé par cette équipe repose sur la faculté de ER α à se fixer sur une région accessible de l'ADN, recruter les protéines SRC-3, puis, par l'intermédiaire de p300, permettre l'acétylation de la K18 de la queue des histones, puis la méthylation de la R17 une fois l'un des SRC remplacé par la protéine CARM. Un tel mécanisme permet de déplier localement l'ADN et de le rendre accessible à la machinerie de transcription.

Plus récemment, l'étude de l'impact du coactivateur ERG sur le récepteur aux androgènes a été menée par une combinaison de l'étude structurale par cryo-microscopie électronique (cryo-EM) et par spectrométrie de masse (MS) après réticulation (Wasmuth *et al.*, 2022). Il y a notamment été montré l'hétérogénéité conformationnelle de l'interface de dimérisation entre les LBDs (**Fig. 17**).

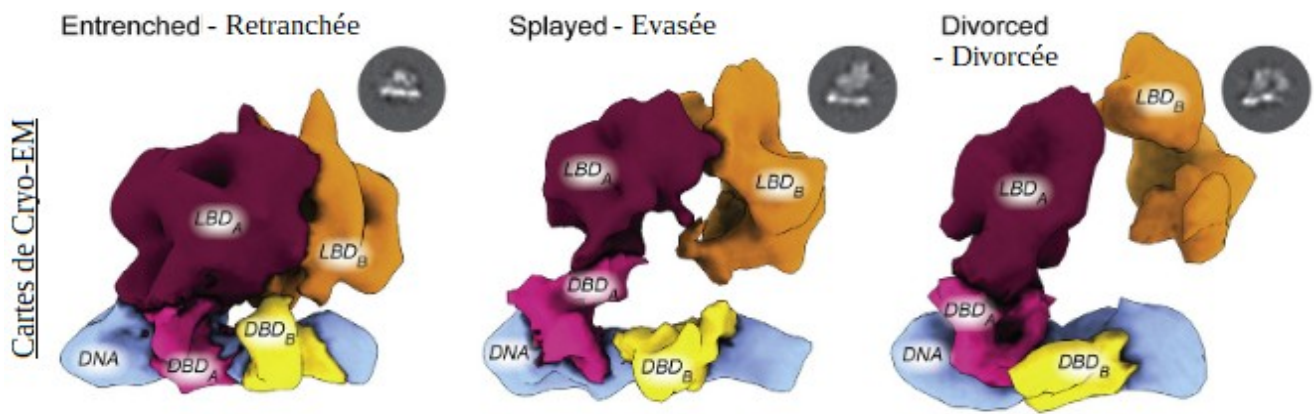


Figure 17: Cartes de cryo-microscopie du récepteur aux androgènes sur son élément de réponse
 (modifiée d'après Wasmuth et al., 2022)

Légende :

Cartes de cryo-EM du récepteur aux androgènes sur son élément de réponse. Trois conformations sont identifiées en fonction de l'interface de dimérisation des LBDs : la conformation retranchée, la conformation évasée et la conformation divorcée.

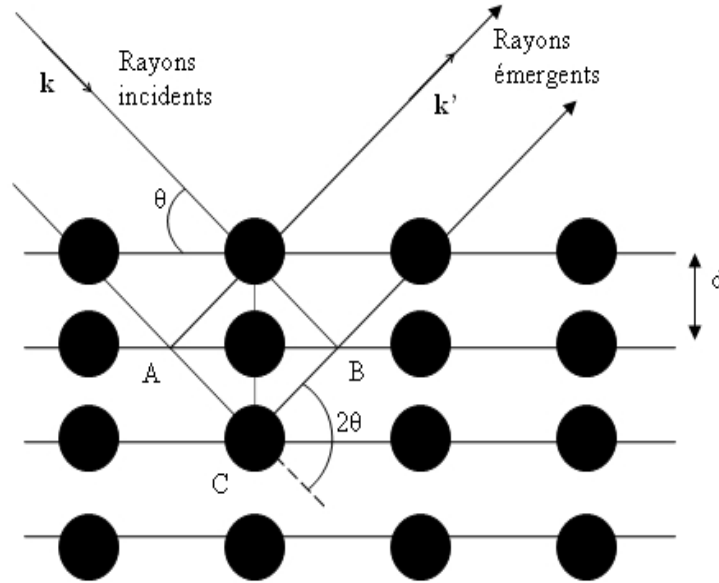
L'objectif de cette thèse étant de caractériser les mécanismes moléculaires régulant la liaison de ERR α sur ses éléments de réponse et l'action de son coactivateur PGC-1 α sur son activité, la partie suivante présentera les techniques de biologie structurale intégrative permettant d'obtenir ces informations.

II. Méthodes d'analyses par biologie structurale intégrative et biophysiques appliquée aux récepteurs nucléaires

Plusieurs techniques permettent d'étudier la structure et la fonction des protéines. Chacune peut présenter un certain nombre d'avantages et d'inconvénients en fonction de ce que l'on cherche à étudier si bien qu'il est devenu plus efficace de combiner ces différentes techniques. Cette démarche est qualifiée de biologie structurale intégrative.

1. L'étude par diffraction aux rayons-X

La diffraction aux rayons-X repose sur l'utilisation de rayons-X sur un échantillon cristallisé. Son principe est le suivant : un faisceau de rayons-X est envoyé sur l'échantillon dont les électrons de chaque atome vont en dévier le faisceau. De part la nature de l'état cristallin de l'échantillon, ces faisceaux sont diffractés de la même manière selon un plan réticulaire et interfèrent entre eux dans certaines directions spécifiques de façon constructive, conduisant à la production d'un signal intense dans certaines zones précises de l'espace qui dépendent de l'empilement cristallin des molécules (**Fig. 18**). Ce signal accumulé est collecté par un détecteur qui présente des pics dans des directions spécifiques de l'espace en fonction de la symétrie du cristal. La position de ces pics est la signature de la maille cristalline et leur intensité dépend du contenu de cette même maille. Un cliché de diffraction est enregistré pour chaque angle incident. Il contient la position des pics, leurs formes et leurs intensités. Avec ces informations et l'acquisition d'informations de phases par la diffusion anormale d'atomes lourds ou le remplacement moléculaire, il est possible de retrouver la structure de l'objet étudié sous la forme de densités électroniques.



Loi de Bragg : $2d \sin(\theta) = n \lambda$

Figure 18: Principe de la diffraction aux rayons-X

Légende :

Selon la loi de Bragg, les directions de l'espace qui vérifient la relation indiquée ci-dessus, où d est la distance inter-réticulaire, λ la longueur d'onde, θ le demi-angle formé par le faisceau incident et le détecteur et n l'ordre de diffraction, permettent l'obtention de pics d'intensité. Autrement dit, il y a interférence constructive lorsque la différence de marche correspond à un nombre entier de longueur d'onde.

Plusieurs structures des domaines DBD et LBD des récepteurs nucléaires ont pu être obtenus par l'utilisation de cette méthode, ainsi que certaines structures de récepteurs nucléaires entiers. Néanmoins, si cette technique permet effectivement d'obtenir des informations structurales à haute résolution, elle comprend aussi des limitations.

La première d'entre elles repose sur la nécessité d'obtenir des cristaux. La cristallogenèse repose sur l'optimisation de conditions physico-chimiques permettant l'empilement ordonné de la protéine sous forme de cristal (**Fig. 19**). Ces conditions sont parfois difficiles à obtenir pour certaines protéines, notamment lorsque ces dernières présentent des régions flexibles et/ou intrinsèquement désordonnées. Les récepteurs nucléaires et leurs coactivateurs sont particulièrement concernés par ce problème. En effet, les récepteurs nucléaires disposent tous d'un domaine N-terminal peu structuré en solution et dont la longueur varie et qui peut grandement nuire à l'empilement cristallin. Dans les cas où des cristaux ont pu être obtenus en le conservant, ce dernier n'est pas visible sur les cartes, ce qui confirme sa flexibilité et donc l'existence de multiples conformations qui se retrouvent moyennées dans le bruit de fond. De la même manière, la région coudée qui sépare les LBDs des DBDs n'est pas très bien connue structuralement. Prédite comme étant peu structurée pour la plupart des récepteurs nucléaires, elle semble pouvoir s'organiser localement sous la forme de tours d'hélices dans le cas de l'hétérodimère RXR-PPAR γ , bien que l'on observe une hétérogénéité de conformation à leur niveau en fonction de la structure considérée (Chandra *et al.*, 2008, codes PDB : 3dzu, 3dzy, 3e00). Dans le cas

du récepteur HNF-4, la densité de cette région existe, mais elle demeure relativement faible par comparaison avec les autres domaines (Chandra *et al.*, 2013).

La seconde limitation repose sur le risque de cristalliser une conformation induite par l'empilement cristallin et donc peu ou pas représentative de l'état physiologique de la protéine. C'est en effet ce que nous avons abordé dans l'étude des complexes formés par les récepteurs nucléaires avec leurs éléments de réponse. La structure cristallographique du complexe PPAR-RXR (Chandra *et al.*, 2008) présente une orientation atypique par comparaison aux autres récepteurs nucléaires connus, ainsi que des interactions entre le DBD de RXR et le LBD de PPAR. L'enveloppe SAXS, qui rend compte des structures possibles du complexe en solution, et donc dans un état proche des conditions physiologiques, indique un ensemble d'orientations plus communes, présentant la forme en « L » que la plupart des récepteurs nucléaires adoptent. Les enveloppes produites disposent certes d'une résolution plus basses en plus de n'être que de potentielles interprétations de la courbe obtenue par les expériences de SAXS, mais les contacts observés, au vu d'autres complexes comme RXR/VDR, y sont plus favorables que dans la structure cristallographique. Cela illustre parfaitement le fait que la cristallisation puisse piéger par l'exercice des contraintes imposées par la formation du cristal une conformation non pertinente biologiquement (Osz *et al.*, 2012) (**Fig. 14**).

Enfin, la cristallisation piège souvent une ou quelques conformations stables d'une protéine parmi d'autres qui favorisent l'état cristallin, ne fournissant pas toujours d'informations sur ses degrés de liberté et son éventuelle flexibilité. Cela va de paire avec le problème de la flexibilité structurale d'une protéine : il existe probablement un nombre déterminé de conformations séparées par des barrières énergétiques que nous ne pouvons discriminer.

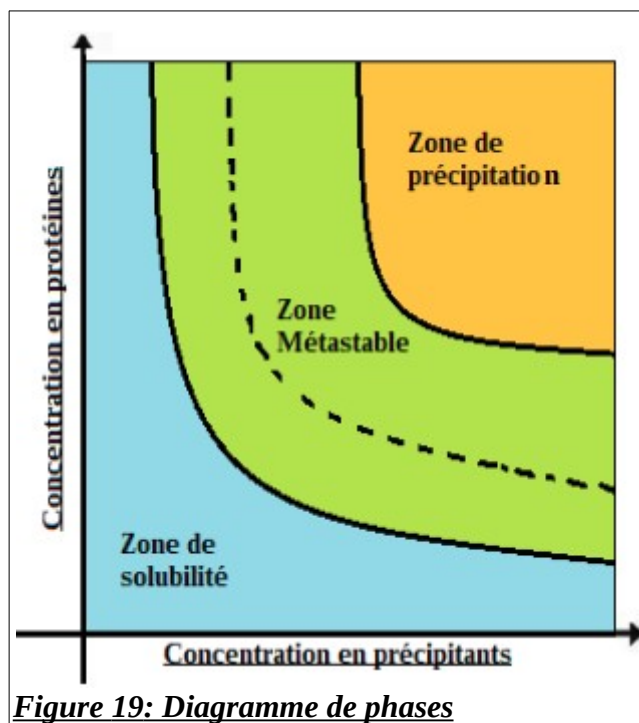


Figure 19: Diagramme de phases

Légende :

Zones attribuées à l'état de la protéine en fonction de sa concentration et de la concentration en agents précipitants. La zone de solubilité est indiquée en bleue (protéine à l'état soluble), la zone Métastable en vert (protéine dans des conditions favorables à son empilement cristallin) et la zone de précipitation est en orange (protéine précipitée de façon non cristalline).

2. L'étude d'échantillons biologiques par cryo-microscopie électronique à transmission

La cryo-microscopie électronique (cryo-EM) consiste en la reconstruction d'un objet en trois dimensions à partir d'images en deux dimensions d'un échantillon. Par comparaison à un microscope optique, un microscope électronique utilise des électrons comme source lumineuse. La longueur d'onde de ces derniers permet de s'affranchir de la limite de résolution dictée par le critère de Rayleigh. La faisabilité de cette méthode sur des échantillons biologiques a été montrée en 1975 par la détermination de la structure de la bacteriorhodopsine à 7 Å de résolution (Henderson & Unwin., 1975), puis par la détermination de la structure de virus (Zhang *et al.*, 2010) et l'obtention de structures à des résolutions atomiques (Yip *et al.*, 2020). La cryogénéisation de l'échantillon dans une fine couche de glace vitreuse le protège des dommages par irradiation par comparaison à des conditions à température ambiante, en plus de conserver l'échantillon sous forme hydraté en préservant la couche de solvatation l'entourant (Glaeser, 1971 ; Taylor & Glaeser., 1974 ; Dubochet *et al.*, 1988). Cela permet notamment de pouvoir acquérir les images à une plus haute dose d'électrons appliqués à l'échantillon. Il y a, communément, deux façons d'analyser un échantillon par cryo-EM. La première consiste à prendre des images d'un même objet pour ensuite le reconstruire en 3D à partir des différents angles de vues capturés. Ce procédé se nomme la cryo-tomographie, laquelle s'applique principalement à des études cellulaires (Schmid MF., 2011). La seconde, dont nous détaillerons l'usage dans cette thèse, est l'analyse de particules isolées (SPA). Elle consiste en la prise d'images de particules orientées aléatoirement, puis en l'attribution angulaire de ces particules et leur moyennisation pour reconstruire ensuite l'objet en 3D (Frank J., 2006).

2.1. La préparation de l'échantillon

La préparation d'un échantillon destiné à la cryo-microscopie repose sur la purification et la vitrification de ce dernier (Dubochet *et al.*, 1988). La vitrification consiste en la congélation rapide de l'échantillon dans une couche fine de tampon n'excédant pas quelques dizaines de nanomètres dans un bain d'éthane ou d'hélium, lui-même refroidit par un bain d'azote liquide afin de permettre un échange thermique rapide (Dubochet *et al.*, 1988 ; Bammes *et al.*, 2010). Le but est de capturer l'échantillon au sein d'une couche de glace vitreuse, laquelle dispose des propriétés d'un liquide qui la rende perméable aux électrons du faisceau (Dubochet *et al.*, 1988). Les conditions de tampon faisant suite à la purification sont importantes, une protéine pouvant présenter, dans de rares cas, des structures différentes en fonction du tampon dans lequel elle a été cryogénisée (Dang *et al.*, 2017; Paulino *et al.*, 2017; Feng *et al.*, 2019). Le tampon a également un impact sur la répartition des particules dans les trous et leur état d'agrégation. Il existe aujourd'hui des techniques de criblages de conditions destinées à la vitrification, notamment par la méthode « Protéoplex » qui repose sur le suivi de l'état de dénaturation d'une protéine marquée selon un gradient de température dans différentes conditions de tampons (Chari *et al.*, 2015).

Le support utilisé en cryo-EM est une grille perforée de trous réguliers dans lesquels l'échantillon va être déposé. Les grilles sont préalablement activées par une décharge lumineuse afin de les rendre plus hydrophiles. La durée de la décharge lumineuse, le temps d'incubation de l'échantillon sur les grilles avant la vitrification, ainsi que les matériaux qui composent le support sont des paramètres cruciaux dans l'optimisation d'un échantillon (Drulyte *et al.*, 2018 ; Weissenberger *et al.*, 2021) pour éviter qu'il ne s'adsorbe sur la grille ou ne se dénature en s'accumulant au niveau de l'interface air-eau

(Glaeser *et al.*, 2017 ; Glaeser., 2018 ; Klebl *et al.*, 2020). Dans le but de limiter ce phénomène, plusieurs types de grilles ont été développées. La grille Quantifoil est une grille commerciale qui peut disposer d'un support en carbone amorphe au sein des trous pour réduire une éventuelle adsorption sur le carbone en dehors des trous (Grassucci *et al.*, 2007 ; Bai *et al.*, 2013 ; Khatter *et al.*, 2015). Les grilles en or sont un type de grille aux propriétés hydrophiles qui permettent éventuellement une meilleure répartition de l'échantillon au sein des trous (Russo & Passmore., 2014). Ces grilles peuvent également être pegylées pour réduire la migration des protéines vers l'interface air-eau. Il existe des grilles disposant d'une couche de graphène oxydé afin d'augmenter le nombre de particules dans les trous, diminuant ainsi leur adsorption au niveau de l'interface air-eau et augmentant leur nombre pour une même concentration de départ (Boland *et al.*, 2017 ; Naydenova & al., 2019 ; Wang F. *et al.*, 2020). Enfin, des grilles d'affinité ont été développées dans le but de concentrer l'échantillon directement sur grille, notamment lorsque celui-ci est au sein d'un mélange complexe. Les grilles disposant d'un cristal bidimensionnel de molécules de streptavidine en sont un exemple (Wang *et al.*, 2008 ; Han *et al.*, 2016 ; Maréchal *et al.*, 2022).

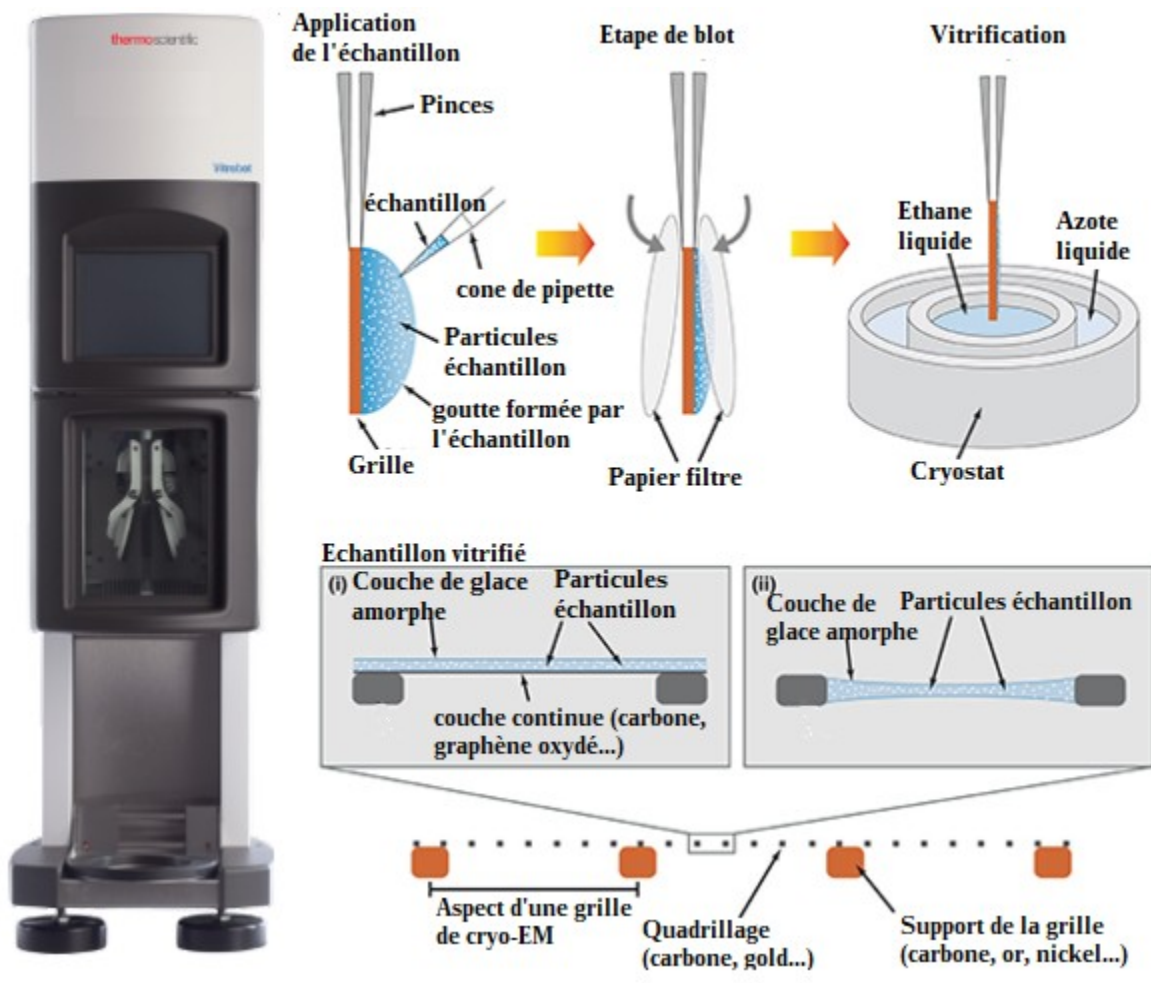


Figure 20: Principe de la vitrification d'un échantillon pour analyse par cryo-EM
(G. Sgro et D. Costa., 2018)

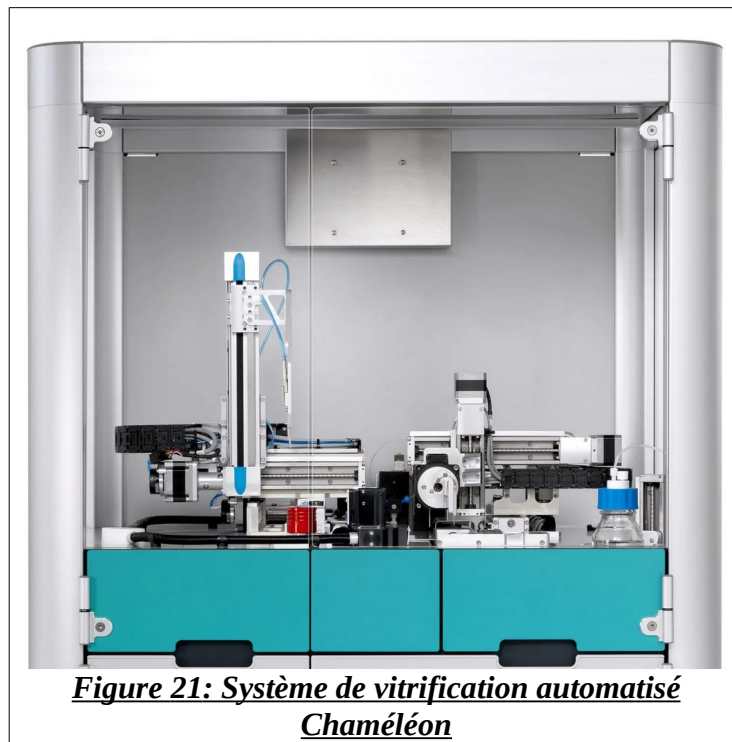
Légende :

Image d'un Vitrobot (à gauche).

Schéma de la vitrification d'un échantillon au vitrobot (à droite). L'échantillon est appliqué sur la grille de cryo-EM. L'excès est ensuite absorbé au papier filtre et la grille est plongée en éthane liquide.

Ces conditions permettent la formation d'une fine couche de glace amorphe contenant l'échantillon au sein des trous que contient la grille ou sur le film de carbone.

La méthode de vitrification la plus couramment utilisée consiste en l'usage du Vitrobot (Thermo Fisher), un outil qui maintient la grille dans une chambre humide à température constante. Il permet l'absorption de l'excès d'échantillon en exerçant une pression à l'aide d'un papier filtre durant un temps déterminé et une force contrôlée avant de plonger la grille dans l'éthane liquide en un temps s'approchant de la seconde (**Fig. 20**). D'autres systèmes existent et utilisent des sprays ou des systèmes microfluidiques pour répandre l'échantillon plus rapidement sur les grilles, évitant ainsi le contact avec le papier filtre qui peut retenir une part significative de l'échantillon et diminuant aussi le temps durant lequel l'échantillon peut se déplacer à l'interface air-eau tout en contrôlant l'épaisseur de la glace (Feng *et al.*, 2017 ; Khaledhonkar *et al.*, 2018 ; Maeots *et al.*, 2020). Le Chameleon (SPT LabTech) est l'un de ces systèmes. Le premier installé en Europe le fut au sein du centre de Biologie Structurale Intégrative de l'Institut de Génétique et de Biologie Moléculaire et Cellulaire où cette thèse a été menée (Darrow *et al.*, 2019) (**Fig. 21**). Il permet de déposer l'échantillon par l'intermédiaire d'un jet rapide dont la vitesse, la dispersion et la pression peuvent être définies et contrôlés par l'utilisateur. L'échantillon projeté sous forme de gouttelettes va être absorbé par une grille disposant de nanocristaux permettant sa répartition sur une échelle de quelques dizaines à centaines de nanosecondes et conduisant à la production d'une couche de glace fine (Wei *et al.*, 2018). Cette thèse nous donnera l'occasion de discuter de ses avantages et limites.



2.2. La prise d'images

Un microscope électronique à transmission (MET) utilise un faisceau d'électrons accélérés par une différence de potentiel élevée créée par l'utilisation d'une haute tension d'accélération. Son architecture

est assez semblable à celle d'un microscope optique, si ce n'est que la source lumineuse est un faisceau d'électrons cohérents. Ses lentilles électromagnétiques focalisent le faisceau d'électrons et ses diaphragmes au niveau des ouvertures des condenseurs permettent de bloquer ceux diffusés hors du faisceau voulu et d'en contrôler la cohérence (**Fig. 22**).

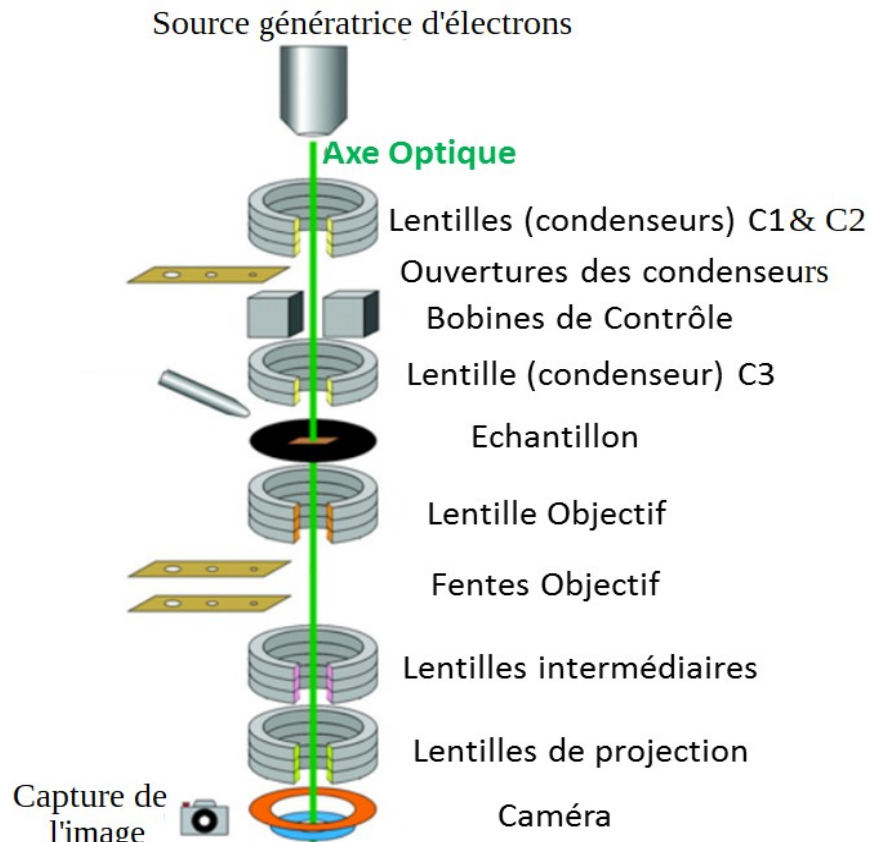


Figure 22: Anatomie d'un microscope électronique

La source génératrice d'électrons peut être un filament de tungstène en forme de V, un cristal taillé en pointe (hexaborure de Lanthane) ou, plus récemment et tel qu'utilisé dans cette thèse, un canon à émission de champ. Cette source est chauffée tout en étant soumise à un courant électrique. Les collisions entre électrons qui en résultent produisent, si ces derniers ont assez d'énergie, l'émission d'électrons sous forme d'un faisceau cohérent spatialement et temporellement à partir de la pointe appartenant à la source d'électrons. En biologie, la tension d'accélération utilisée est généralement comprise entre 100 et 300 kV afin de trouver le bon compromis entre résolution et contraste des images. Il existe également des sources d'électrons qui n'ont plus besoin d'être chauffées : les canons à émission de champs froids. Ces derniers permettent d'obtenir un faisceau plus cohérent en évitant les gradients de température locaux qui pouvaient induire des émissions d'électrons non cohérents avec le reste du faisceau.

Une fois émis, les électrons sont envoyés au travers de plusieurs lentilles électromagnétiques. Ces dernières sont des bobines de cuivre sous tension dont la fonction consiste à émettre un champ magnétique capable de modifier la trajectoire des électrons. Elles vont ainsi permettre la focalisation du faisceau d'électrons de sorte à ce que ce dernier soit parallèle à l'axe optique et conserve sa cohérence.

Les défecteurs jouent également ce rôle et interviennent donc dans l’alignement du microscope, étape nécessaire à toute acquisition.

Les électrons traversent ensuite l’échantillon et peuvent agir de trois façons distinctes lors de cette rencontre (**Fig. 23-A**). Dans le premier cas, l’électron passe à travers l’échantillon et n’interagit pas, ou très peu, avec ce dernier. Dans ce cas, l’électron conserve son énergie et sa direction. Dans le second cas, l’électron va subir une diffusion élastique lors de son interaction avec un atome de l’échantillon (**Fig. 23-B**). Il ne perd pas d’énergie, mais dévie de sa trajectoire initiale. C’est principalement ce type d’interaction qui contribue au contraste de l’image et à l’obtention de données structurales. On parle ici de contraste de phase, ce dernier étant principalement dû au décalage de phase entre les électrons ayant interagit et ceux n’ayant pas interagit avec l’échantillon. Dans le dernier cas, l’électron va subir une diffusion inélastique et perdre en énergie tout en déviant de sa trajectoire en interagissant avec l’échantillon (**Fig. 23-C**). Ce type d’interaction cause la formation de radicaux libres au sein de l’échantillon, causant sa dégradation. Cela correspond à ce que l’on peut qualifier de dommages par irradiation. Ces électrons sont souvent éliminés au cours de leur trajet dans la colonne du microscope par l’intermédiaire des lentilles et défecteurs filtrant les électrons en fonction de leur énergie par l’utilisation de champs magnétiques.

Il est également possible de filtrer les électrons en utilisant un filtre à énergie avant la caméra de type GIF. Ce système introduit un changement d’angle au cours du trajet des électrons et utilise celui-ci en combinaison avec des filtres physiques et des champs électromagnétiques afin de s’assurer du parallélisme du faisceau par rapport à la caméra, ainsi que de la vitesse et donc de l’énergie des électrons. Cela permet l’élimination des électrons diffusés inélastiquement.

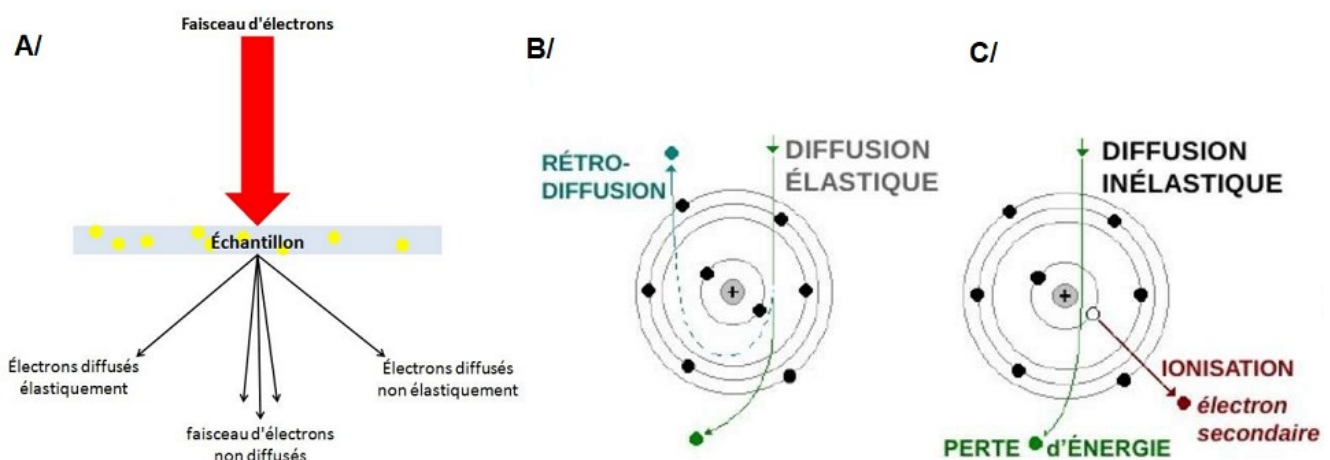


Figure 23: Interactions entre le faisceau d’électrons et l’échantillon

(selon la thèse « Origine et évolution des récepteurs nucléaires et étude structurale du premier stéroïdien, ERR » de Brice Beinsteiner, (2018))

Légende :

A/ Lorsque le faisceau rencontre l’échantillon, les électrons peuvent interagir avec les noyaux de l’échantillon ou ne pas être diffusés. B/ Diffusion élastique. Cette dernière contribue à l’image si elle n’est pas rétro-diffusée. Dans ce cas, l’électron conserve son énergie et sa vitesse. C/ Diffusion inélastique. L’électron interagit avec l’échantillon, causant sa ionisation. Il y a perte d’énergie de l’électron primaire, perte de vitesse, mais aussi émission d’un électron secondaire de moindre énergie.

Le fait qu'un électron interagisse ou non avec l'échantillon et le type d'interaction avec celui-ci dépend fortement de la tension d'accélération utilisée. En effet, la diminution du voltage tend à augmenter l'interaction des électrons avec l'échantillon, augmentant le contraste observé. Cependant, la contre-partie est généralement une diminution de la résolution causée, notamment, par l'augmentation des dommages par irradiation.

Une fois que les électrons ont traversé et potentiellement interagit avec l'échantillon, ils vont traverser une nouvelle lentille magnétique : la lentille objectif. Cette dernière va faire converger le faisceau d'électrons au niveau du plan focal, formant le spectre de diffraction de l'échantillon. Il est alors possible de décrire ce spectre mathématiquement par sa transformée de Fourier. Les électrons sont ensuite projetés sur le plan image pour former l'image de l'objet qui correspond donc à la transformée de Fourier inverse du spectre de diffraction de l'échantillon. Cette image est acquise au niveau de la caméra (Fig. 24).

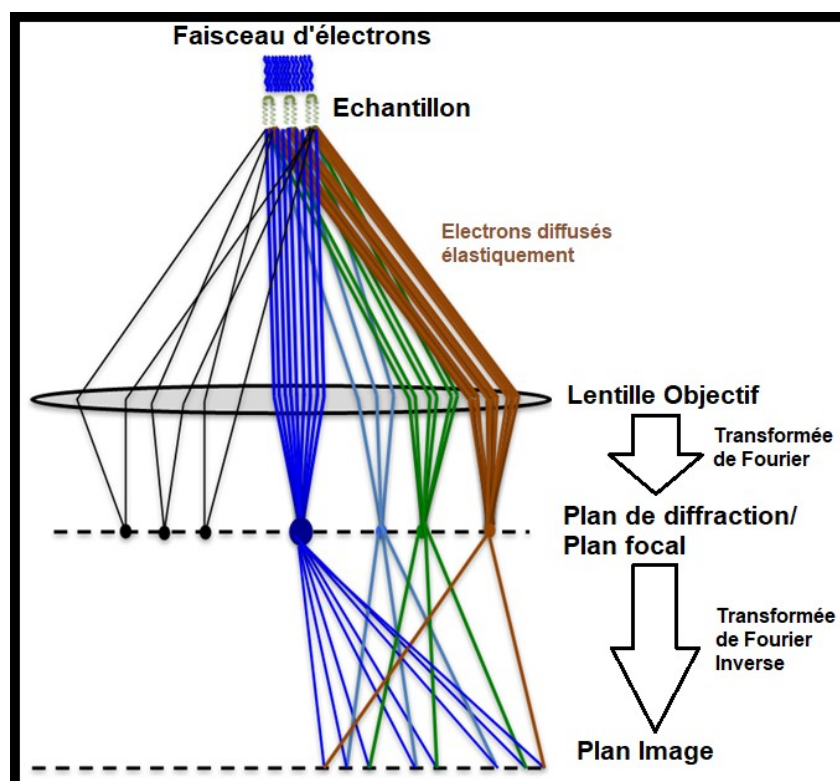


Figure 24: La formation de l'image

Légende :

Les électrons diffusés élastiquement sont focalisés par la lentille objectif au niveau du plan focal afin de former une image de diffraction. Cela correspond à effectuer une transformée de Fourier de l'image telle qu'arrivée au niveau de la lentille objectif. Puis, depuis le plan focal, les électrons diffusés et non diffusés vont interférer et produire le contraste de phase à l'origine de l'image au niveau du plan image. Cela correspond à effectuer la transformée de Fourier inverse du cliché de diffraction.

Les caméras utilisées avant 2011 étaient principalement des caméras à détection indirecte. Elles fonctionnaient en convertissant les électrons en photons à l'aide d'un scintillateur pour en augmenter le signal, puis les transférait vers un dispositif à transfert de charge servant de détecteur par

l'intermédiaire d'un réseau de fibres optiques plus petits que les pixels en eux-mêmes. Cette étape de conversion des électrons s'accompagnaient d'une dégradation du signal et donc du rapport signal/bruit. La plupart des caméras utilisées aujourd'hui sont des caméras à détection directe. Elles s'affranchissent de cette étape de conversion en utilisant les technologies CMOS (oxyde métallique semi-conducteur complémentaire) et détectent les électrons lorsque ces derniers entrent en contact avec la couche de silicium du détecteur. Cette amélioration permet d'augmenter l'efficacité quantique de détection (DQE), soit la quantité de signal détecté par rapport au bruit pour une fréquence considérée (McMullan *et al.*, 2009 ; Faruqi & McMullan., 2011 ; Ruskin *et al.*, 2013 ; McMullan *et al.*, 2014 ; Kuijper *et al.*, 2015 ; Spear *et al.*, 2015) (**Fig. 25**).

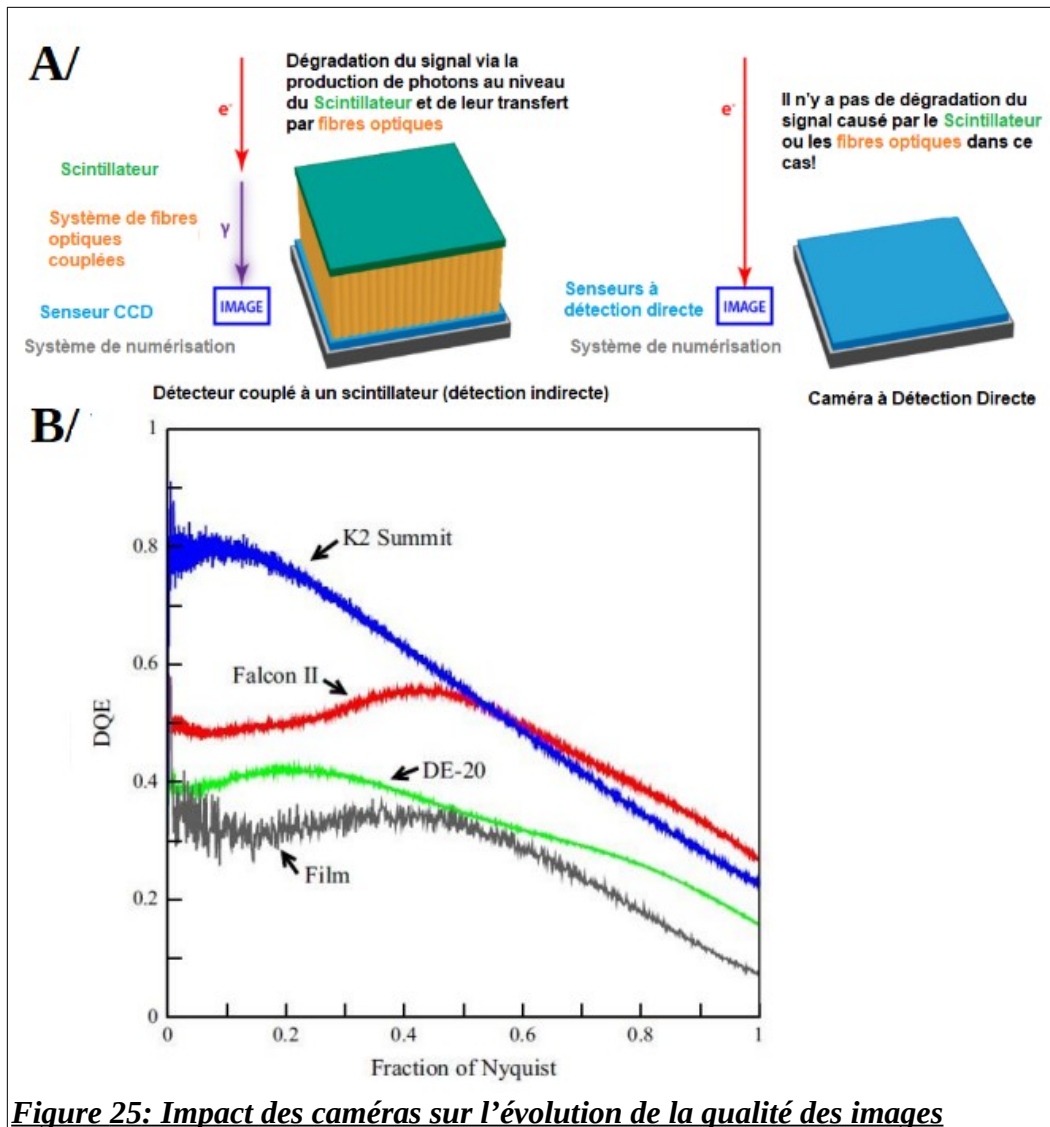


Figure 25: Impact des caméras sur l'évolution de la qualité des images
(inspiré par le site directelectron.com et McMullan *et al.*, 2014)

Légende :

A/ Comparaison des modes de détection des caméras à détection indirecte et directe.

B/ Comparaison des DQE en fonction du type de caméra. Le Film indiqué correspond au film photographique, DE-20 est une caméra à détection indirecte. La Falcon II est une caméra à détection

directe de 1ère génération et la K2 est une caméra à détection directe « récente » bénéficiant de multiples modes de collecte.

Sous l'action du faisceau d'électrons, l'échantillon subit des changements locaux de températures, des effets de charge et de la diffusion mécanique au sein de la glace vitreuse, ce qui peut engendrer le déplacement, la rotation et le changement des angles des particules. Afin de corriger ce déplacement au cours du temps, les images sont acquises sous forme de films composés d'images successives. Il est alors possible de corriger d'une part le mouvement de l'objet par l'alignement des images reposant sur des calculs de corrélation et, d'autre part, de mettre plus de poids sur les premières images pour lesquelles la dose appliquée à l'échantillon est plus faible afin de diminuer l'impact des dommages par irradiation (Kunath *et al.*, 1984 ; Brilot *et al.*, 2012 ; Campbell *et al.*, 2012; Li *et al.*, 2013; Veessler., 2013 ; Allegretti *et al.*, 2014 ; Scheres, 2014 ; Grant *et al.*, 2015). L'image finale résultante de ce procédé se nomme le « micrographe ».

Les échantillons biologiques sont des objets à faible contraste de phase car ils possèdent des atomes légers, disposant de noyaux avec peu de protons, ce qui conduit à ce qu'ils soient plus facilement dégradables par les dommages de radiation et qu'ils interagissent moins avec les électrons du faisceau, réduisant le contraste des images. Pour préserver l'échantillon, les acquisitions sont effectuées à faible dose. Les images ainsi obtenues sont de faible contraste rendant difficile la détection de l'échantillon pour les étapes ultérieures dans le traitement d'images.

L'image obtenue par microscopie n'est pas l'objet imagé en 3D, mais une projection 2D de celui-ci déformée par un certain nombre de paramètres physiques qui altèrent la qualité de l'image. La relation entre l'objet et son image est telle que :

$$o = \text{psf} * e * i + n \quad (1)$$

où « o » correspond à l'objet imagé, « psf » correspond à la fonction d'étalement des points, « e » correspond à la transformée de Fourier inverse de la fonction enveloppe, « i » correspond à l'image de l'objet obtenue et « n » correspond au bruit de fond. Le symbole « * » signifie la réalisation d'une opération de convolution entre ces fonctions.

Cette expression mathématique peut être mise dans l'espace de Fourier, transformant ainsi les positions dans l'espace en fréquences spatiales. On peut alors écrire :

$$O = \text{CTF} \times E \times I + N \quad (2)$$

« O » est ici la transformée de Fourier de l'objet imagé, « CTF » correspond à la fonction de transfert du contraste, « E » correspond à la fonction enveloppe, « I » à la transformée de Fourier de l'image de l'objet et « N » la transformée de Fourier du bruit de fond. Le symbole « x » désigne ici l'opération de multiplication.

Les fonctions de transfert du contraste (CTF) et enveloppe (E) sont en général propres au microscope utilisé, ainsi qu'à la méthode de collecte choisie. La fonction de transfert du contraste est une fonction complexe qui modélise mathématiquement la capacité du microscope à restituer le contraste dans l'image produite à partir de l'objet imagé dans un

système optique donné. Elle peut s'écrire de différentes façons et se compose d'une fonction de transfert de modulation (ou d'amplitude) et d'une fonction de transfert de phases.

Concernant la fonction enveloppe, elle traduit l'atténuation du signal pour les hautes fréquences qui limite grandement la résolution. Elle dépend du facteur β qui est propre au microscope et aux différentes parties le composant ainsi que des conditions d'acquisition. Elle a d'autant plus d'impact que le faisceau d'électrons perd en cohérence spatio-temporelle et que la défocalisation choisie est prononcée (Wade & Franck., 1977 ; Glaeser *et al.*, 2021).

Ces deux fonctions peuvent être définies mathématiquement ainsi :

$$\text{CTF}(\lambda, g, \Delta f, Cs, \Delta \phi) = -w_1 \times \sin[\Pi \lambda |g|^2 (\Delta f - 0,5 \lambda^2 |g|^2 Cs) + \Delta \phi] - w_2 \times \cos[\Pi \lambda |g|^2 (\Delta f - 0,5 \lambda^2 |g|^2 Cs) + \Delta \phi] \quad (3)$$

$$E(w) = e^{-2Bw^2} \quad (4)$$

Ces équations, telles que définies dans les travaux de Wade en 1992 (Wade., 1992 ; Mindell & Grigorieff., 2003 ; A.Rohou & N.Grigorieff., 2015), traduisent chacune les déformations subies par l'image obtenue par comparaison avec l'objet observée. Les paramètres qu'elles font intervenir sont décrits ci-après.

λ est la longueur d'onde utilisée, soit ici celle de l'électron. En imagerie, la longueur d'onde que l'on utilise dicte ce que nous sommes capable d'observer. En effet, la résolution se définit comme la distance (ou l'angle) minimale à partir de laquelle il est possible de distinguer deux objets distincts.

Dans le cas d'un instrument optique doté d'un diaphragme traversé par une onde monochromatique de longueur d'onde λ , l'image de l'objet obtenue est composée, non pas de points définis représentant l'objet, mais d'un ensemble de disques de diffusion (tache d'Airy) de diamètre D pour un angle d'ouverture du diaphragme θ . On peut écrire le diamètre D tel que :

$$D = (1,22 \times \lambda) / (n \times \sin(\theta)) \quad (5)$$

où n est l'indice de réfraction du milieu traversé.

Le critère de Rayleigh décrit mathématiquement la résolution comme étant la distance « d » pour laquelle deux pics (soit le cercle central d'une tache d'Airy) sont séparables pour une configuration où la première annulation de la tache d'Airy du premier pic (inversion du contraste) correspond au maximum du second pic.

$$D = (0,61 \times \lambda) / (n \times \sin(\theta)) \quad (6)$$

En cryo-EM, nous travaillons à petits angles et dans le vide ($n = 1$). Il est donc possible de décrire la distance donnant la valeur de résolution telle que :

$$d = \lambda / 2 \quad (7)$$

De plus, il est possible de décrire la relation entre le voltage utilisé et la longueur d'onde. Sachant que l'équation de Broglie donne :

$$\lambda = h / mv \text{ (8)}$$

Avec « h » la constante de Planck, « m » la masse du corpuscule et « v » sa vitesse.

Sachant que l'énergie cinétique d'un électron peut s'écrire telle que :

$$0,5 \times m \times v^2 = e \times V \text{ (9)}$$

avec « e » la charge de l'électron ($1,6 \times 10^{-19}$ Coulomb), « m » la masse de l'électron ($9,11 \times 10^{-28}$ g) et « V » le voltage rendant compte de la différence de potentiel en volts qui accélère le corpuscule.

D'après la relation donnée par Louis de Broglie, la longueur d'onde s'écrit alors :

$$\lambda = (h/m) \times (1/\sqrt{2e \times V/m}) \text{ (10)}$$

Sachant que 1 joule correspond à 10^7 g.cm².s⁻², alors :

$$\lambda = \sqrt{(150 / V) \times 10^{-8}}, \text{ en cm (11)}$$

$$\lambda = 12,3 / \sqrt{V}, \text{ en nm (12)}$$

Cependant, l'électron étant relativiste, sa vitesse n'étant pas négligeable par rapport à celle de la lumière, ce qui affecte le calcul de son énergie et donc de sa longueur d'onde. On peut donc écrire, d'après l'hypothèse formulée par Louis de Broglie :

$$\lambda = (12,3 / \sqrt{V}) (1 / \sqrt{1+(eV / 2mc^2)}), \text{ en nm (13)}$$

Ainsi, sachant que pour une tension d'accélération de 300 kV, la longueur d'onde des électrons émis est d'environ 0,019 Å d'après l'équation 13, la résolution théorique vaut 0,0095 Å d'après l'équation 7. Cependant, une telle résolution ne peut pas être atteinte du fait de l'intervention des autres paramètres dans la fonction CTF et dans la fonction Enveloppe. La longueur d'onde n'est donc pas limitante puisqu'elle est environ 200 fois plus petite que la distance entre atomes.

Le paramètre Δf correspond au defocus utilisé. Les images produites en microscopie électronique étant des images de faible contraste, celui-ci peut être augmenté en défocalisant l'échantillon (**Fig. 26 A-B**). En effet, si dans le cas d'une lentille parfaite, toutes les ondes électromagnétiques provenant d'un échantillon sont focalisées en un point unique dans le plan image où elles arrivent parfaitement en phase, dans le cas d'un échantillon défocalisé, au dessus ou au dessous du plan image, les décalages de phase produisent des interférences, induisant des variations de contraste. Cependant, la défocalisation a également un effet néfaste sur la qualité de l'image obtenue. Plus elle est importante, plus la fonction de transfert du contraste voit sa période se raccourcir. Cela induit plus de passages de la fonction par le zéro et donc une perte substantielle d'informations, la différence d'intensité entre les électrons déviés par l'échantillon et le faisceau devenant nulle en ces points. De plus, la forte défocalisation nuit également à la fonction enveloppe en affectant l'amplitude des fréquences, ce qui cause une perte des informations de hautes résolutions (**Fig. 26 C-D**). Enfin, l'estimation des paramètres de la fonction de transfert du contraste devient plus délicate. A l'inverse, être trop proche de la focale nuit, comme décrit précédemment, au contraste de l'objet.

Intervient ensuite le paramètre C_s , lequel correspond aux aberrations sphériques (C_s) induites par les lentilles électromagnétiques. Les lentilles induisent que la déviation des électrons sera plus importante proportionnellement à leur éloignement de l'axe optique. La correction de ce paramètre par un correcteur d'aberrations sphériques permet l'obtention de meilleures résolutions (Khatter *et al.*, 2015).

$\Delta\phi$ correspond à la variation de phase, notamment dans le cas de l'utilisation d'une Volta phase plate et g correspond au vecteur des fréquences spatiales de coordonnées w_1 et w_2 en deux dimensions.

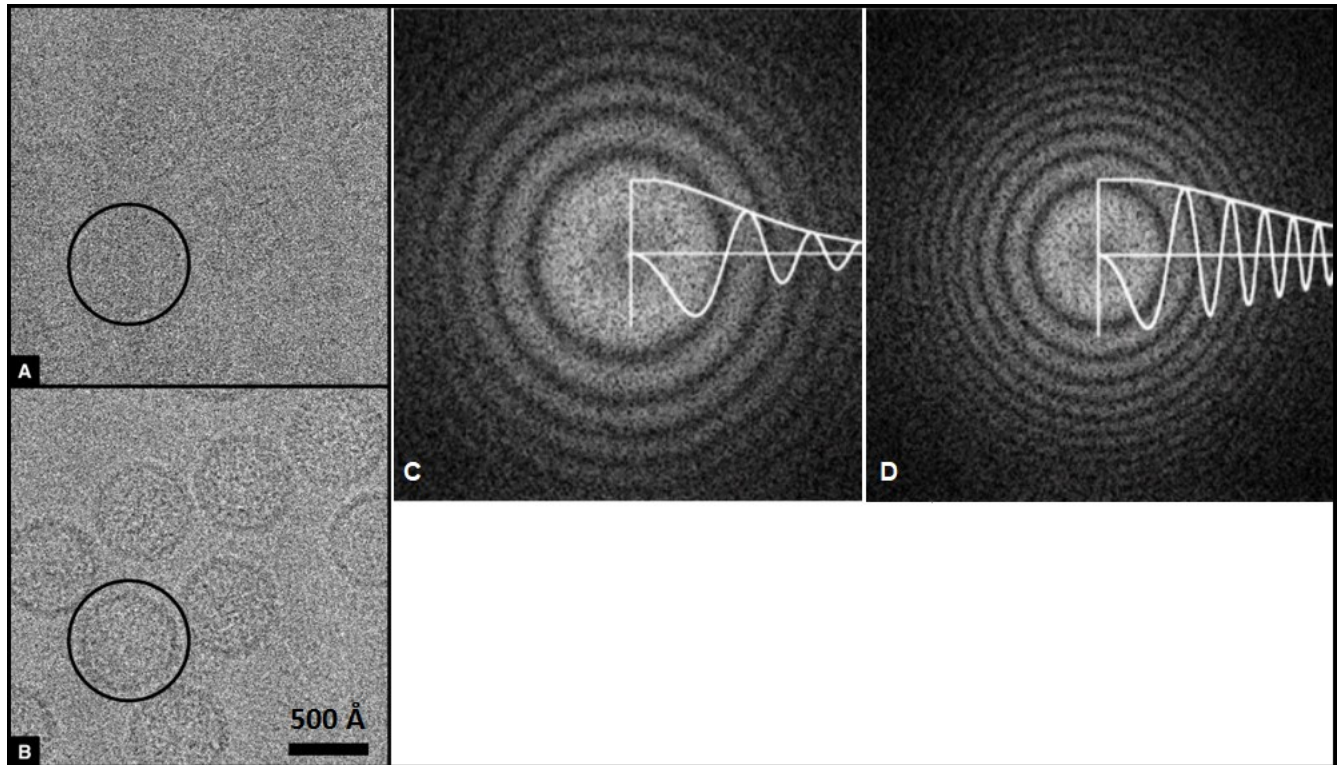


Figure 26: Impact du defocus sur le contraste et la CTF
(Orlova et Saibil, 2011)

Légende :

A/ Image prise avec un defocus de $-0,97 \mu\text{m}$. Les informations de hautes résolutions dominent. **B/** Image prise avec un defocus de $-2,7 \mu\text{m}$. Les informations de basses résolutions sont plus importantes et l'on observe un gain significatif de contraste. **C/** Spectre de puissance avec superposition de l'estimation de la CTF et de la fonction enveloppe en 1D d'un micrographe pris à un defocus de $-0,5 \mu\text{m}$. **D/** Spectre de puissance avec superposition de l'estimation de la CTF et de la fonction enveloppe en 1D d'un micrographe à un defocus de $1 \mu\text{m}$. Le premier pic est plus proche de l'origine et les pics s'enchaînent plus rapidement : l'augmentation du defocus a diminué la périodicité. De plus, l'amplitude des anneaux de Thon est moins forte : les informations de plus hautes résolutions sont perdues.

La CTF ressemble en une dimension à une fonction sinusoïdale variant entre -1 et 1 (**Fig. 27 A**). Les valeurs négatives de la CTF correspondent à une information de signal qui est perdue ce qui conduit à devoir considérer la CTF comme n'étant constituée que de valeurs positives en inversant le signe des régions négatives, permettant de récupérer le signal pour les fréquences concernées (**Fig. 27 B**). Ensuite, la valeur précédent le premier zéro varie en fonction de la valeur de defocus à laquelle

l'image a été réalisée, ne nous permettant pas d'avoir une information couvrant l'entière des fréquences spatiales d'intérêt (**Fig. 26 C-D**). Pour remédier à ce problème, il faut accumuler des images à plusieurs valeurs de défocalisation pour que les spectres de CTF cumulés permettent d'obtenir une meilleure couverture du signal pour les différentes fréquences spatiales. Enfin, le signal est perdu pour les fréquences spatiales où la CTF passe par zéro. Là aussi, l'accumulation d'images à différents défocus permet de pallier à ce problème et de récupérer le signal perdu.

Cette fonction CTF va diminuer au fur et à mesure que l'on s'approche des hautes fréquences de par la fonction enveloppe évoquée précédemment dont le facteur β explique les défauts liés aux aberrations causées par les lentilles et la stabilité du microscope.

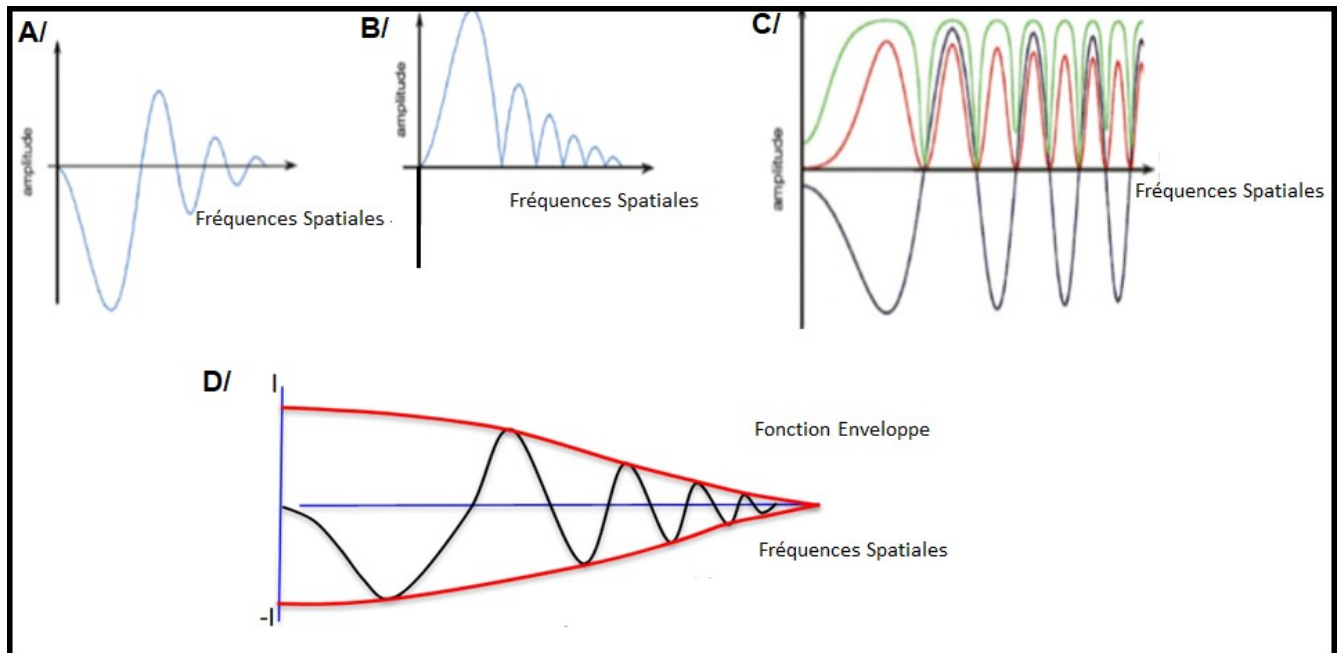


Figure 27: La correction de la CTF

Légende :

A/ CTF non corrigée. **B/** CTF après inversion des phases négatives (phase flipping). **C/** CTF originale représentée en noir, puis après inversion des phases négatives en orange et après amplification des amplitudes en vert. **D/** Impact de la fonction enveloppe sur la CTF et son atténuation aux hautes fréquences.

Aujourd'hui, les logiciels sont de plus en plus fins dans l'estimation des paramètres de la CTF, la prise en compte de l'astigmatisme et la détection du faible signal pour les hautes fréquences. Cela permet, en plus des améliorations précédemment citées, l'obtention de structures à plus haute résolution. Par exemple, CTFFIND ou encore gCTF servent à estimer les paramètres de la CTF à partir du spectre de puissance du micrographe (Rohou et Griegorieff., 2015 ; Zhang., 2016).

2.3. La reconstruction 3D de l'objet imagé

Une fois les paramètres de la CTF estimés et les micrographes triés afin de ne conserver que ceux répondant à des critères de qualités suffisant vient la sélection des particules. L'augmentation des

capacités de calcul et la rapidité des acquisitions permet aujourd'hui de collecter des milliers d'images et de sélectionner, par l'utilisation de logiciels d'apprentissage à partir d'une sélection par l'utilisateur ou par « deep-learning », des millions de particules-objets (Redmon *et al.*, 2016 ; Bepler *et al.*, 2018 ; Wagner *et al.*, 2019). Parmi ces logiciels, nous pouvons, entre autre, citer crYOLO ou TOPAZ qui permettent, à partir d'une sélection de particules objets définis par l'utilisateur, d'apprendre et de sélectionner les particules sur l'ensemble des micrographes.

Ces particules sont ensuite extraites des micrographes. Il est possible, à cette étape, de réaliser un ré-échantillonnage des particules, c'est-à-dire de réduire le nombre de pixels décrivant l'objet en les fusionnant de proches en proches. Ainsi, un ré-échantillonnage d'ordre deux permet de réduire une boîte de 256 par 256 pixels en une nouvelle boîte de 128 par 128 pixels. Se faisant, la résolution maximale atteignable, dépendante de la fréquence de Nyquist, soit deux fois la taille des pixels, change également d'un facteur 2. Cependant, lors des premières étapes de traitement de l'image, un tel procédé permet de réduire considérablement les temps de calculs puisque le nombre total de pixels et donc de données à traiter diminue d'un facteur 4 dans notre exemple. De plus, cela peut également permettre un meilleur alignement des objets. En effet, la fusion des pixels augmente le signal et donc le contraste de l'image ce qui induit un meilleur rapport signal sur bruit.

Une fois extraites, les particules sont classifiées en 2D afin de regrouper les particules avec une orientation similaire par maximum de vraisemblance (Sigworth., 1998). Cela permet de séparer les mauvaises particules des bonnes particules sur la base de la taille de l'objet et de sa forme. Il est possible d'utiliser des masques 2D circulaires à cette étape afin de permettre le meilleur alignement des particules. Une fois les bonnes classes sélectionnées par l'utilisateur, elles sont utilisées pour reconstruire en 3D l'objet image à basse résolution.

Différentes méthodes mathématiques permettent d'aboutir à la reconstruction 3D *ab initio*. Nous nous concentrerons ici sur le théorème de la section centrale. Ce dernier stipule que la transformée de Fourier 2D de chaque image correspond à une section de la transformée de Fourier en 3D de l'objet considéré. Le croisement des plans 2D dans l'espace de Fourier se fait selon une section commune qui indique l'origine de la transformée de Fourier 3D. C'est-à-dire que si l'on considère 2 projections 2D différentes d'un même objet, alors chacune correspond à une section de la transformée de Fourier en 3D de ce même objet et qu'il y a donc une ligne commune correspondant à l'intersection de ces deux sections. Cette méthode permet ainsi d'assigner pour chaque particule une orientation dans l'espace via l'attribution de trois angles que l'on qualifie d'angles d'Euler (van Heel., 1987). Trois plans sont nécessaires afin de déterminer une orientation initiale. En effet, les images obtenues étant des projections de l'objet, deux plans ne permettraient pas de déterminer la chiralité de l'objet. Ce problème, lors de la reconstruction de l'objet, peut survenir et est nommé le problème de la main. Il consiste en l'obtention d'une structure qui est la molécule chirale de l'échantillon. De plus, cette étape étant très sensible au bruit, elle ne prend généralement en compte que les informations de basses résolutions, c'est-à-dire plus de 12 Å (Crowther et al, 1970) (**Fig. 28**). Plus récemment, une approche par descente stochastique de gradient a vu le jour et repose sur le fait de diviser l'ensemble des particules à traiter en sous-ensembles traités séparément afin d'affiner la carte en réduisant les risques que la sélection des particules ne conduisent à l'obtention d'une carte *ab initio* biaisée, correspondant à un minimum énergétique local que de mauvaises particules ou de mauvaises assignations d'angles auraient produites (Punjani *et al.*, 2017).

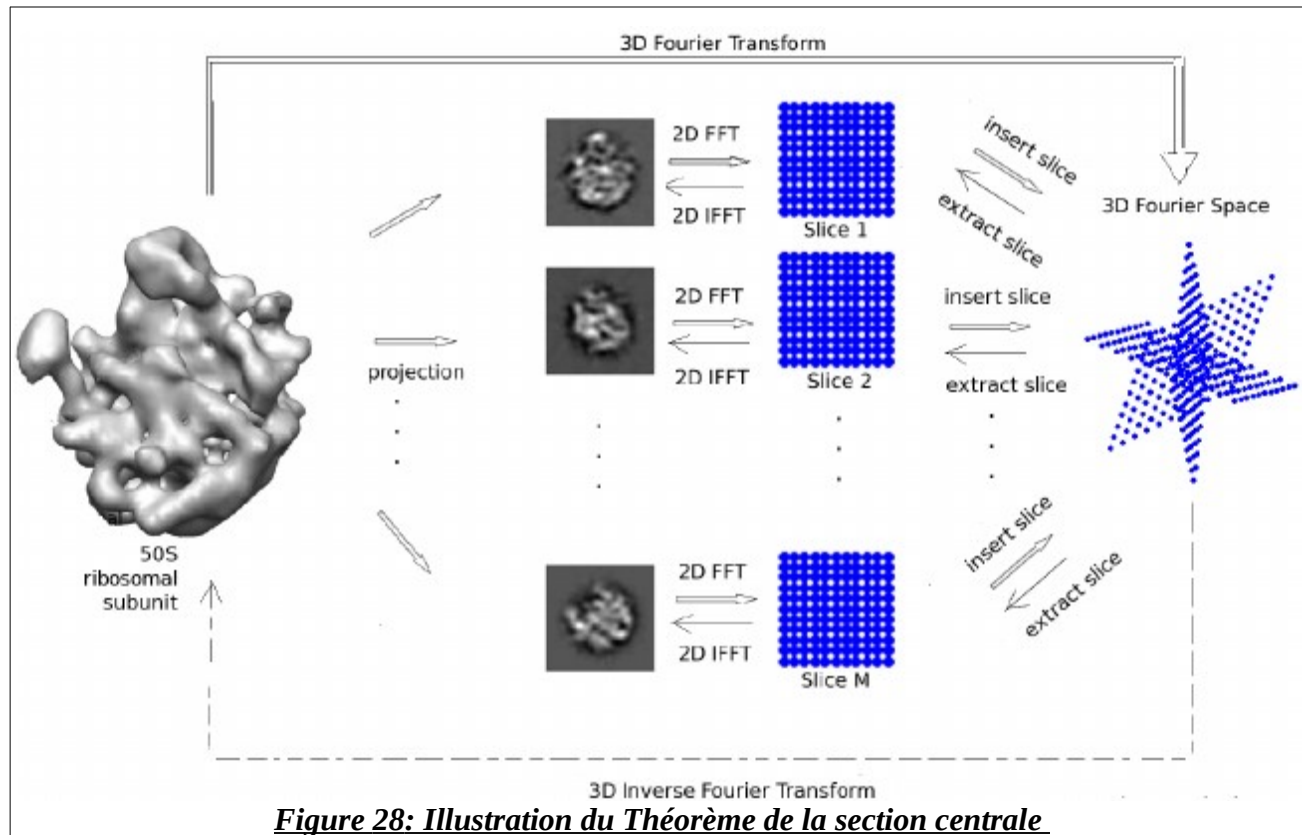


Figure 28: Illustration du Théorème de la section centrale

(Wang et al., 2013)

Légende :

Les projections de l'objet (aussi appelée particules) sont des images 2D dont la transformée de Fourier est une coupe en 2D de la transformée de Fourier en 3D de l'objet initial. Ainsi, en sommant les projections avec des angles de vues différents, la transformée de Fourier de l'objet imagé en 3D est reconstruite. La transformée de Fourier inverse permet de reconstruire l'objet imagé.

Une fois une bonne structure *ab initio* obtenue, il est courant d'affiner la carte initiale pour atteindre les hautes résolutions ou alors de classifier les particules obtenues afin de séparer plusieurs conformations distinctes de l'objet au cours d'une opération que l'on nomme la classification 3D. La classification 3D peut se baser sur des cartes de références connues et trier les particules en fonction de leurs paramètres d'alignements 3D obtenus après la reconstruction *ab initio*. Elle peut également ne pas utiliser de références et projeter les particules pour reconstruire de nouvelles cartes en fonction de leurs paramètres d'alignements 3D. Sinon, la classification peut se faire en séparant les particules en un nombre d'ensembles demandés par l'utilisateur suivi d'une décomposition par une analyse par composante principale de l'espace 3D pour ensuite utiliser les volumes produit comme point de départ pour classifier les particules. Pour sa part, l'affinement repose de plus en plus sur des algorithmes de maximum de vraisemblance et d'analyses Bayésienne (Sigworth et al., 1998 ; Scheres et al., 2005, dans *Bioinformatics* ; Scheres et al., 2005 dans *J.Mol.Biol*) par lesquels une particule n'est plus associée strictement à une classe, mais à un ensemble de classes 3D avec une probabilité d'attribution plus ou moins forte suivant la classe considérée. Ces approches sont notamment employées dans les logiciels RELION et cryoSPARC. Ces logiciels sont des interfaces graphiques qui regroupent un ensemble de logiciels de traitement de l'image et permettent un confort d'utilisation plus important pour l'utilisateur

ainsi qu'une optimisation des temps de calculs nécessaires à chacune des tâches du traitement d'image (Scheres *et al.*, 2012 ; Punjani *et al.*, 2017)

La résolution de la carte affinée obtenue en fin de traitement est estimée par le calcul de la corrélation des amplitudes en espace Fourier de deux cartes affinées indépendamment en séparant le jeu de données en deux (van Heel, 1987). Le critère communément utilisé par la communauté scientifique est de 0.143. Ce procédé sert à éviter, au moins partiellement, le phénomène de « overfitting » qui consiste en l'affinement du bruit en lieu et place de l'objet. Cela peut cependant arriver même en affinant deux cartes indépendamment (Scheres & Chen, 2012).

Certains logiciels servent également à prendre en compte la flexibilité dont font preuve certains complexes ou peuvent rechercher l'existence de conformations différentes directement au sein des particules plutôt que par la reconstruction de classes 3D indépendantes. C'est notamment le cas d'un certain nombre d'opérations dans le logiciel cryoSPARC, mais aussi au sein de RELION. Au sein de RELION, le « multi-body refinement » permet d'affiner un complexe ou une protéine multi-domaines en sélectionnant la position de ces derniers au sein d'une carte à basse résolution et en utilisant des masques pour les affiner un par un en conservant les distances relatives entre eux. Cela sert à pallier le mouvement de certaines régions (Nakane & Scheres, 2020). Dans cryoSPARC, l'opération « 3D variability » permet de chercher, dans un ensemble de particules pré-alignées, les changements conformationnels directement depuis les images. Cela permet notamment d'identifier les régions capables de mouvements ou présentant des conformations que l'on ne peut isoler par l'intermédiaire de la moyennation que réalise la classification 3D (Punjani & Fleet, 2021). Enfin, l'opération « d'affinement non-uniforme » permet d'affiner au sein d'une carte les régions rigides malgré la présence de régions mobiles. Ce procédé fonctionne par le changement de l'une des étapes de l'affinement afin de prendre en compte le désordre intrinsèque d'une protéine en prévenant les risques de sur-affinement, aussi appelé overfitting, par un système de régulation croisée adaptative qui identifie les régions moins bien résolues et les filtre différemment du reste de la carte plutôt que d'appliquer un filtre basé sur l'enveloppe de Fourier de façon uniforme. En effet, lors de la séparation du jeu de données en deux demi-cartes, chacune d'entre elles est affinée de façon indépendante et se retrouve avec sa propre valeur d'enveloppe de Fourier plutôt que de dépendre d'une valeur partagée avec l'autre demi-carte (Punjani *et al.*, 2020).

3. La titration calorimétrique isotherme

La titration calorimétrique isotherme (ITC) est une méthode d'analyse biophysique permettant de mesurer les paramètres thermodynamiques d'une interaction en solution. La méthode consiste à déposer un échantillon à une concentration connue dans une cuve adiabatique à température et pression constantes, puis d'ajouter graduellement une autre solution contenant la molécule censée interagir avec l'échantillon à une concentration connue par le biais de plusieurs injections suffisamment espacées pour que le système revienne à l'équilibre thermique (**Fig. 29**). En effet, la réaction entre les molécules testées va induire un transfert de chaleur qui va être mesurée par comparaison à une autre cuve contenant l'échantillon ou le tampon seul, lui-même maintenu à la même température que l'échantillon. Les pics de variations de température correspondant à chacune des injections vont ensuite être intégrés pour produire l'isotherme. Cette courbe permet de déduire directement la constante de dissociation et l'enthalpie de la réaction, ainsi que la stœchiométrie du complexe formé. Il est alors possible d'obtenir le profil thermodynamique complet de l'interaction étudiée à la température donnée grâce à la relation suivante :

$$\Delta G = - RT \times \ln(Kd) = \Delta H - T\Delta S \quad (13)$$

Si cette technique dispose de l'avantage de ne pas avoir besoin de marquer les molécules étudiées, les mesures de transfert de chaleur sont généralement difficiles à détecter, ce qui contraint l'utilisateur à devoir utiliser une concentration et une quantité importante d'échantillon et limite l'application de cet outil à certains projets.

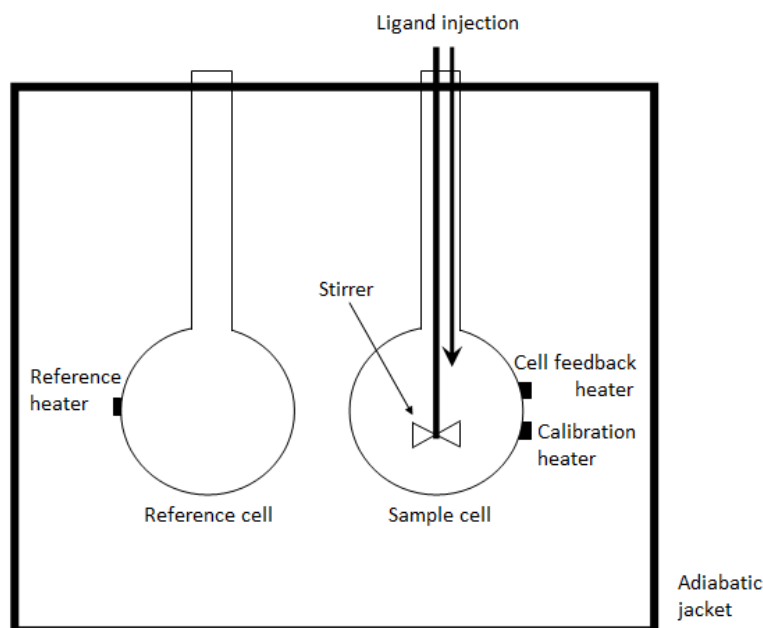


Figure 29: Principe de la titration calorimétrique isotherme

Légende :

L'enceinte adiabatique est représentée par le rectangle dans lequel la cellule de référence (à gauche) et la cellule expérimentale (à droite), sont représentées. L'agitateur (stirrer) maintient l'homogénéité de la solution présente au sein de la cellule expérimentale lors de l'ajout du ligand par plusieurs injections espacées. Entre chaque injection, la chaleur dégagée ou absorbée par la réaction est analysée par la mesure de la température de la solution et cette dernière est ensuite compensée. Cela permet d'obtenir la variation de la chaleur liée à la réaction.

4. La méthode d'analyse Proteoplex

La méthode d'analyse Protéoplex est une méthode d'analyse de la stabilité d'un échantillon dans un tampon donné selon un gradient de température. Cette méthode est une mesure de fluorimétrie à balayage différentiel (DSF), mais elle a été conçue et optimisée pour l'analyse de complexes afin de trouver les conditions de tampon les plus stabilisantes et/ou les plus propices à leur analyse structurale. La méthode Protéoplex va suivre la fluorescence d'un échantillon selon un gradient de température. La mesure s'effectue dans une large gamme de tampon en faisant varier les conditions de pH, de salinité ou la présence d'éventuels adjuvants. Le fluorophore utilisé est souvent une molécule capable de lier non spécifiquement les régions hydrophobes des protéines, par exemple le SYPRO Orange (Niesen *et al.*, 2007). Celui-ci voit sa fluorescence éteinte par l'eau contenue dans le tampon. Puis, lors de l'augmentation de la température, la protéine qui se dénature expose des régions hydrophobes à sa surface. Le SYPRO Orange va alors s'y fixer et la fluorescence va pouvoir être observée, témoignant du passage de l'échantillon à un état dénaturé. La méthode Protéoplex se concentre sur la découverte de conditions de tampons dans lesquelles une seule à deux phases de transitions sont observées pour des protéines composées de plusieurs domaines et/ou de complexes multi-protéiques. Ces dernières rendent compte d'une dénaturation coopérative entre les protéines impliquées, et donc une condition de tampon dans laquelle le complexe est stable. De part l'existence de transitions non coopératives et coopératives au sein d'un complexe, un nouveau modèle a été développé par rapport à une mesure de DSF standard. Il a notamment été montré que des conditions dans lesquelles une seule transition de dénaturation est observée sont favorables à la bonne dispersion du complexe sur grille de microscopie électronique et correspondent donc à des conditions de tampons optimisées (Chari *et al.*, 2015) (**Fig. 30**).

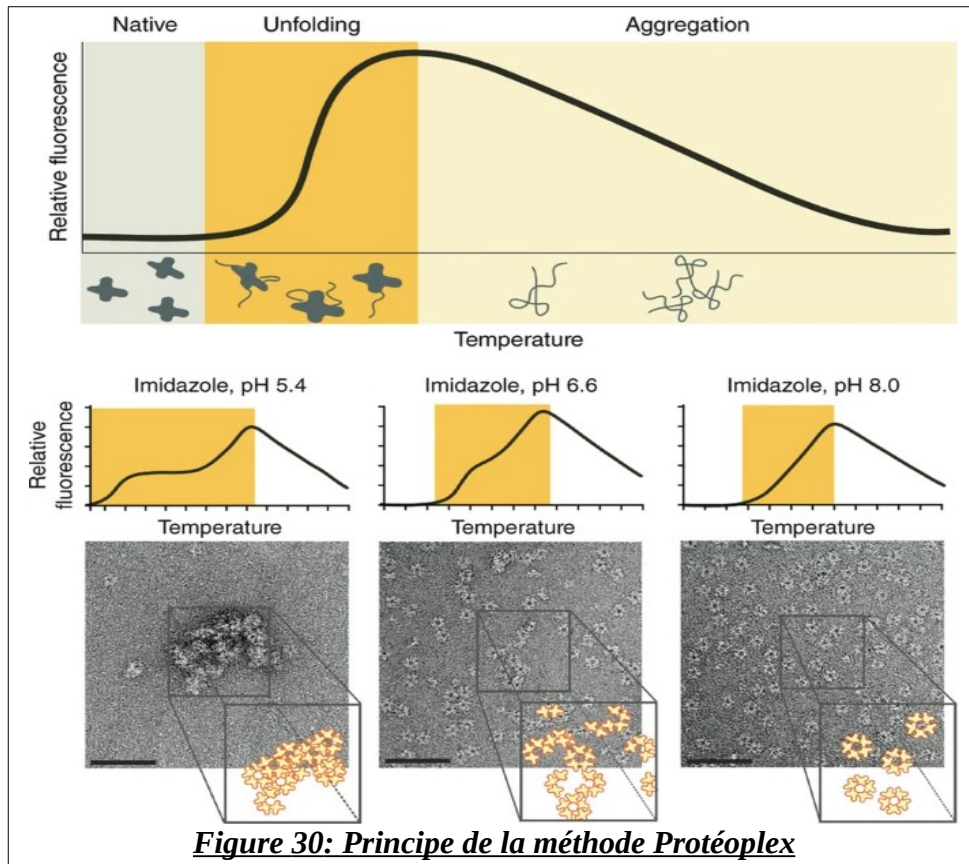


Figure 30: Principe de la méthode Protéoplex

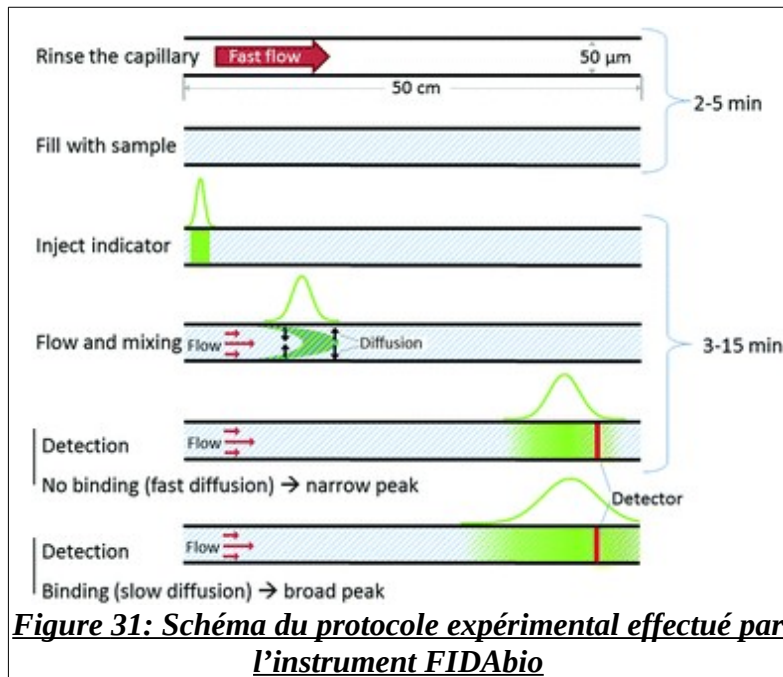
(Chari et al., 2015)

Légende :

Les protéines marquées sont dénaturées par l'application d'un gradient de température causant leur dépliement et l'exposition des résidus hydrophobes au solvant ce qui engendre un signal de fluorescence suivi par le détecteur. En fonction de la nature du tampon, ici de l'imidazole à différents pH, il est possible de voir une dénaturation sous la forme d'une seule et unique étape, montrant une coopération des protéines, et donc une complexation, lors de la dénaturation thermique. Cela se traduit en cryo-EM par la meilleure répartition des particules dans ce type de tampon.

5. FIDAbio : Analyse de dispersion par flux induit

Cette méthode consiste à caractériser une interaction entre deux molécules en suivant l'évolution du rayon hydrodynamique de l'un des partenaires marqué par un fluorophore. Elle permet notamment la mesure de constantes de dissociation de l'ordre du nanomolaire, voir du picomolaire pour une faible dépense d'échantillon : quelques nanolitres à une concentration de quelques nanomolaires. Son principe repose sur la diffusion de Taylor et sur le changement des propriétés de diffusion de l'échantillon lorsque ce dernier interagit avec la molécule d'intérêt. D'un point de vue expérimental, cette méthode requiert la préparation de plusieurs tubes contenant une concentration fixe de la molécule marquée et une concentration croissante du partenaire.



Un schéma d'expérience classique se décompose généralement comme suit : dans un premier temps, le capillaire est rincé à la soude, sauf dans le cas où celui-ci est « enrobé » d'une couche de molécule hydrophiles, ce qui peut être nécessaire si l'échantillon tend à interagir avec le capillaire. Puis, le capillaire est rempli de tampon. L'analyte est ensuite injecté, suivi de l'indicateur, puis de l'analyte à nouveau. L'indicateur désigne ici l'échantillon, soit la protéine marquée et/ou le complexe. L'analyte désigne le ligand testé à la même concentration que la protéine d'intérêt. Au niveau du détecteur, la fluorescence est mesurée et le pic expérimental obtenu est comparé à un modèle en forme de gaussienne. Le flux induit et le temps nécessaire à la détection ainsi que l'étalement du pic obtenu décrivent l'évolution des populations de protéines seules et de complexes. Il est ainsi possible de tracer la courbe rendant compte de l'évolution du rayon hydrodynamique en fonction de la concentration et d'obtenir la constante d'affinité de l'échantillon (**Fig. 31**). Pour une mesure précise, il faut impérativement prêter attention à la viscosité du tampon, ainsi qu'à la température utilisée et à la pression choisie pour le flux.

6. L'échange hydrogène-deutérium

L'échange hydrogène-deutérium (HDX) correspond au fait de réaliser le marquage au deutérium d'une protéine en condition native afin que seules les hydrogènes exposés au solvant soit échangés. Le marquage est réalisé avec différents temps d'incubation afin de prendre en compte les éventuels changements conformationnels au cours du temps ou les zones difficiles d'accès au solvant qu'il peut y avoir dans une molécule multi-domaines. S'en suit l'arrêt du marquage par neutralisation acide et la digestion de la protéine afin d'identifier les fragments obtenus par spectrométrie de masse. L'identification de deutérium permet de mettre en avant l'exposition d'une région au solvant. Il s'agit donc d'un outil comparatif puissant dans le cadre de l'étude de l'effet d'un ligand ou d'un coactivateur afin de mettre en évidence les régions dont la conformation ou la dynamique évolue (**Fig. 32**). Cependant, dans le cas d'un même objet multimérique, il est difficile, voir impossible, de discriminer l'un des monomères de ses voisins si ces derniers ne sont pas équivalents suite à la fixation d'un ligand. Il s'agit, dans notre cas, d'une limite attendue à l'interprétation des résultats. De la même manière, une protéine peu structurée aura tendance à être marquée trop rapidement par le deutérium. Si la liaison d'un coactivateur ou d'un ligand ne change pas de façon importante sa conformation, ou que la conformation change en restant hautement flexible, cette expérience n'en captera pas les variations. Cette limite est attendue dans le cadre de l'étude du coactivateur PGC-1 α .

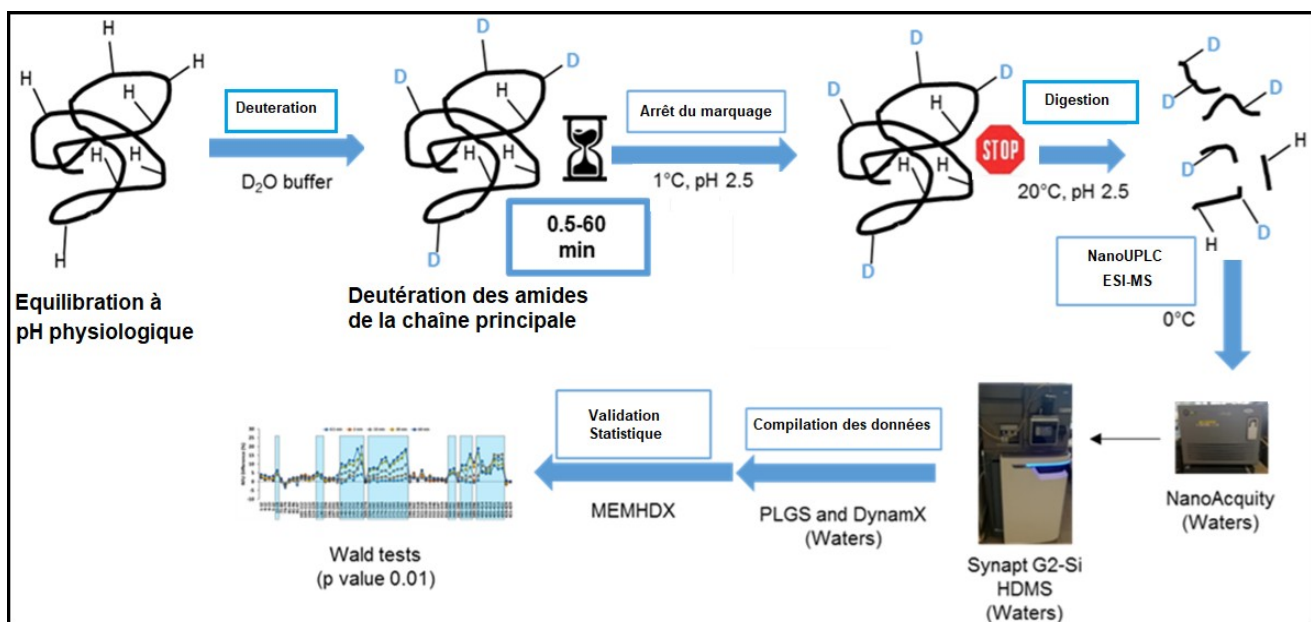


Figure 32: Protocole suivi pour les expériences d'échanges hydrogènes-deutériums

Légende :

Les hydrogènes exposés au solvant sont échangés par des deutérium durant un temps variant de 30 secondes à 1 h au cours de l'étape de marquage. L'échantillon est placé en tampon acide pour arrêter la réaction, puis l'échantillon est digéré par protéolyse ménagée. Les produits de dégradation sont analysés après passage sur colonne afin de les isoler en fonction de leurs tailles, puis ils sont analysés par spectrométrie de masse et comparés aux banques de données et à la séquence des protéines mises en jeu.

7. La réticulation couplée à la spectrométrie de masse (XL-MS)

La XL-MS correspond au fait de réticuler au cours de plusieurs temps donnés un complexe avant de le digérer afin de détecter par spectrométrie de masse les peptides des protéines physiquement proches les unes des autres en solution. La molécule de réticulation choisie est critique pour déterminer les distances d'interactions et les types de liaisons sans dénaturer ou agréger la protéine (**Fig. 33**). De plus, dans le cas de protéines peu structurées, il est possible que certaines interactions soient artéfactuelles, d'où le besoin de comparer et répliquer les expériences et de comparer les résultats avec d'autres techniques. Dans le cas des expériences réalisées au cours de cette thèse, une proximité structurale au niveau de la région connue de PGC-1 α pour son interaction avec l'hélice H12 de ERR α est attendue afin de valider l'expérience. De plus, une limitation attendue aux expériences menées est de ne pas pouvoir discriminer les interactions qu'un seul monomère de ERR α pourrait avoir avec PGC-1 α d'interactions qui impliqueraient les deux monomères du récepteur. Il sera cependant possible de comparer les régions d'interactions observées aux cartes de cryo-EM pour les valider.

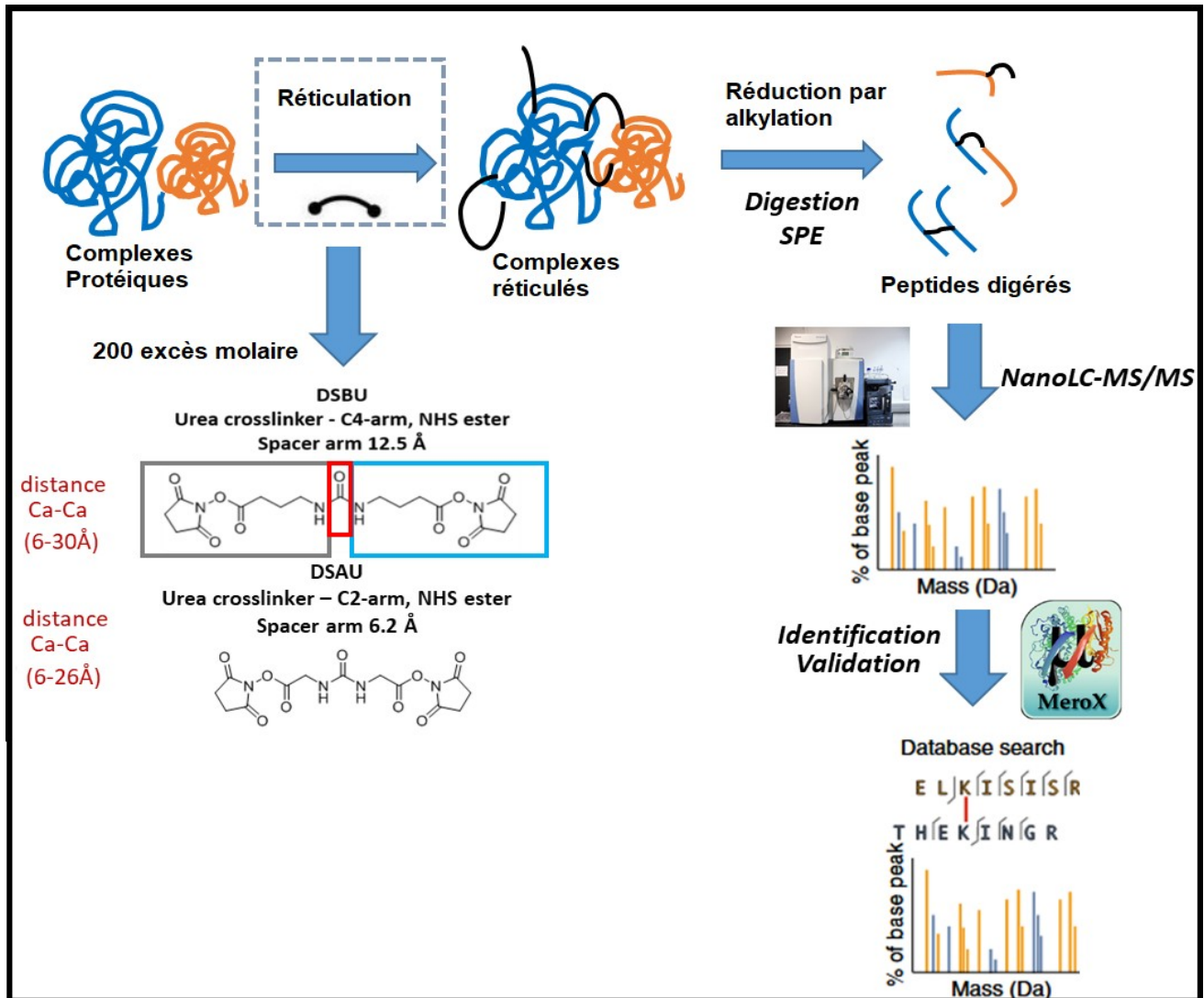


Figure 33: Protocole suivi pour les expériences de réticulation

Légende :

Le complexe protéique est réticulé avec l'une ou l'autre des molécules DSBU ou DSAU. Les échantillons sont réduits par alkylolation, puis une protéolyse ménagée est effectuée sur le complexe. Ces peptides sont isolés par un passage sur une colonne NanoLC, puis analysés par comparaison avec les données fournies de la protéine et contre plusieurs banques de données.

III. Les objectifs de la thèse

Les objectifs de cette thèse ont été d'apporter des réponses sur les mécanismes de reconnaissance et de sélection de l'ADN par le récepteur nucléaire apparenté au récepteur aux oestrogènes $ERR\alpha$. Il fut aussi question de comprendre la topologie du récepteur nucléaire entier sur son élément de réponse et l'impact que son coactivateur PGC-1 α peut avoir sur sa structure et sur son activité.

Afin de comprendre les mécanismes régulant la dimérisation des DBDs sur l'élément de réponse composite et le rôle de la séquence d'ADN dans la dimérisation, l'étude structurale du domaine de liaison à l'ADN sur un élément composite naturels a été réalisée. Cela a permis l'obtention d'une structure à haute résolution décrivant les mécanismes moléculaires mis en jeu dans la dimérisation et le rôle de la séquence d'ADN dans la fixation du second DBD. L'étude biochimique portant sur l'extension « TNA » du premier demi-site de fixation en 5' et sur la sélection des ERRE par les DBDs a également été menée. Cela a permis de déterminer le rôle de la base « N » dans la sélectivité des DBDs pour les ERRE. Le rôle de la coopération entre domaines dans la mise en place de la dimérisation du DBD et de son accroche sur le second demi-site fut également étudié biophysiquement. Cela a confirmé la spécificité de sélection des séquences par le récepteur, ainsi que la singularité de l'élément composite par comparaison aux autres ERRE naturels. Une analyse évolutive a été conduite afin de comprendre les mécanismes spécifiques aux ERR qui ont pu être perdus ou acquis durant l'évolution.

Afin de comprendre le fonctionnement du récepteur nucléaire $ERR\alpha$ et les mécanismes qui régissent son activité fonctionnelle, l'analyse biophysique du récepteur entier en présence de son coactivateur PGC-1 α a été conduite. L'étude du récepteur en présence de différents types d'éléments de réponse et en présence ou non de PGC-1 α a été conduite par l'utilisation du FIDABio et d'expériences de SAXS afin de décrire l'action de PGC-1 α sur la capacité de $ERR\alpha$ à se lier à l'ADN en fonction du type d'élément de réponse considéré. Sachant que la plupart des ERRE ne sont pas des séquences composites, il s'agissait de déterminer si le coactivateur permettait la fixation du second DBD sur l'élément de réponse, mais aussi sous quelles conditions par comparaison à l'élément composite.

Ensuite, afin d'avoir des réponses quant aux régions de $ERR\alpha$ que PGC-1 α est capable de lier en solution, différentes approches par spectrométrie de masse ont été menées sur le récepteur nucléaire seul ou en présence du coactivateur. Le coactivateur étant une protéine intrinsèquement désordonnée, cette méthode fournit davantage de régions d'interactions que celles que l'on aurait pu obtenir par une méthode d'analyse structurale standard. Ces expériences ont également été l'occasion de comparer l'impact de l'ADN et l'impact du coactivateur séparément et en concomitance sur le récepteur nucléaire.

Dans l'optique de comprendre comment s'effectue la formation et l'ordre d'assemblage des complexes d'activation de la transcription, l'expression, la production et la purification des protéines PCAF et p300 ont été réalisées.

Enfin, l'étude structurale du complexe a été effectuée d'une part par diffraction aux rayons-X et d'autre part par cryo-microscopie électronique. Les protéines $ERR\alpha$ et PGC-1 α disposent de régions flexibles et le complexe qu'elles forment fait intervenir un troisième partenaire sous la forme d'un ADN qui peut se dissocier. Plusieurs formes peuvent exister en équilibre thermodynamiques en même temps en solution. De plus, les forces de tractions, les contacts hydrophobes et les tensions de surface induites par l'interface air-eau intervenant dans le processus de vitrification peuvent conduire à une

déstabilisation du complexe. Ces éléments laissent entrevoir un certain nombre de difficultés lors des expériences de détermination structurale (déstabilisation de l'empilement cristallin par la flexibilité des protéines, existence de multiples conformations discrètes, difficulté de sélection et d'alignement des particules isolées, faible contraste, complexes dissociés ou partiellement dissociés, agrégation...etc). Les cartes obtenues et affinées au cours du traitement d'images permettent de replacer les structures individuelles des domaines LBDs et DBDs. En combinaison avec les résultats précédents de HDX, de XL-MS, mais aussi en remplaçant les domaines isolés obtenus par cristallographie dans les cartes, il est possible d'étudier l'architecture de ce complexe, ce qui constitue un défi majeur scientifiquement et techniquement.

Partie I – L’ADN dicte la liaison asymétrique de ERR α

Résumé de l’article

Le récepteur ERR α est un récepteur stéroïdien basal régulant le métabolisme énergétique des cellules. Il se trouve à une position clef de l’évolution pour comprendre comment les récepteurs stéroïdiens sont passés de la reconnaissance d’un demi-site étendu en 5’ à la reconnaissance de séquences palindromiques espacés de trois paires de bases. Parmi les ERRE que reconnaît ERR α , il fut identifié un élément de réponse composite composé d’une demi-site étendu compris au sein d’un élément de réponse palindromique (Deblois *et al.*, 2009). Notre équipe a montré que les DBDs de ERR α sont capable de s’y lier sous forme dimérique (Mohideen-Abdul *et al.*, 2017).

Cette première partie va donc avoir pour objectif d’expliquer les mécanismes mis en jeu dans la reconnaissance de ces éléments de réponses distincts par le DBD. L’article aborde l’expression et la purification du DBD de ERR α ainsi que la formation de complexes impliquant différents ADNs naturels. Il y est notamment comparé par titration calorimétrique isotherme des éléments de type IR3, soit des séquences répétées inverses séparées de trois paires de bases comportant un ERE, les ERRE « classiques » tels que la séquence Tr α ERRE, mais aussi l’élément composite sous le nom d’élément embERRE/IR3 au sein de l’article. Le rôle de la seconde base de l’extension en 5’ du demi-site canonique des ERR est étudiée par migration sur gel natif selon un gradient de sel afin de montrer la sélectivité exercée par cette région du petit sillon. Enfin, la structure cristallographique du DBD sur l’élément de réponse composite obtenue à 2,48 Å nous permet de commenter l’arrangement du domaine et ses particularités dans la reconnaissance des séquences cibles, ainsi que son interface de dimerisation.

Conclusions

En conclusion, nous avons obtenu la structure du DBD de ERR α sous forme dimérique à 2,48 Å sur son élément de réponse composite. L’étude de son interface de dimerisation révèle l’existence d’une boîte fonctionnelle propre aux ERR : la boîte « KRRRK ». Cette dernière permet la dimerisation asymétrique des DBDs en augmentant la taille de l’interface de dimérisation.

Concernant l’interaction avec l’élément de réponse, la séquence et les paramètres de structure de l’ADN vont être les éléments conditionnant la fixation du DBD sous forme dimérique ou monomérique. Le profil thermodynamique des différents ADN testés indique en effet que la fixation des DBDs s’effectue de façon coopérative si la séquence répond à certains critères au niveau du second demi-site, notamment au niveau de la forme de l’ADN au niveau de la position +5. La formation de la liaison à l’ADN a été suivie qualitativement par migration sur gel natifs afin de mettre en avant la sélectivité exercée par l’extension « TNA » en 5’ sur l’interaction du DBD avec différents éléments de réponse. Il en ressort que la base « N » en seconde position permet une interaction plus favorable avec une guanine ou une cytosine.

Enfin, l’analyse phylogénétique suggère que l’interface de dimérisation asymétrique de ERR α est un élément ancestral nécessaire à la formation d’une interface de dimérisation forte qui a été perdu au cours de l’évolution par les récepteurs stéroïdiens.

Néanmoins, les ERRE étant reconnus *in vivo* par ERR α et ce dernier étant sous forme de dimère obligatoire, cela signifie qu’il doit exister un mécanisme permettant la fixation des DBD en dépit de la

forme de la séquence seule. Cela peut être la résultante d'une communication entre les domaines de $ERR\alpha$, mais aussi l'impact de l'action d'un coactivateur.

Contribution

Ma contribution à cet article repose sur la purification des protéines et la réalisation des mesures de titration calorimétrique isotherme en collaboration avec l'Institut de Biologie Moléculaire et Cellulaire de Strasbourg. J'ai également participé à la production de certaines figures et à la correction de l'article.

Article – Résultat (version actuelle du Manuscrit)

DNA is the allosteric driver for the asymmetric binding of homodimer estrogen-related receptor

Kareem Mohideen Abdul Patel^{1,2,3,4}, Pierre Vilela^{1,2,3,4}, Tajith Baba Shaik^{1,2,3,4}, Alastair McEwen^{1,2,3,4}, Isabelle Hazemann^{1,2,3,4}, Karl Brillet⁵, Eric Ennifar⁵, Ali Hamiche^{1,2,3,4}, Gabriel V. Markov⁶, Vincent Laudet⁷, Dino Moras^{1,2,3,4}, Bruno P. Klaholz^{1,2,3,4} and Isabelle M.L. Billas^{1,2,3,4,*}

¹ IGBMC (Institute of Genetics and of Molecular and Cellular Biology), Illkirch, France

² Université de Strasbourg, Unistra, Strasbourg, France

³ Institut National de la Santé et de la Recherche Médicale (INSERM) U1258, Illkirch, France

⁴ Centre National de la Recherche Scientifique (CNRS) UMR 7104, Illkirch, France

⁵ Architecture Et Réactivité de L'ARN, CNRS UPR 9002, Institut de Biologie Moléculaire et Cellulaire, Université de Strasbourg, 67000, Strasbourg, France.

⁶ Sorbonne Université, CNRS, UMR 8227, Integrative Biology of Marine Models, (LBI2M, UMR8227), Station Biologique de Roscoff (SBR), 29680 Roscoff, France

⁷ Marine Eco-Evo-Devo Unit. Okinawa Institute of Science and Technology. 1919-1 Tancha, Onna-son, 904-0495 Okinawa, Japan

* Corresponding author: billas@igbmc.fr (I.M.L. Billas)

Keywords: ERR α ; DNA binding domain, X-ray crystallography; Steroid Nuclear Receptors; Protein-DNA complex; DNA allostery

Running Title: Crystal structure of the homodimeric ERR α DNA binding domain

Abstract

The estrogen-related receptors (ERRs, NR3B), orphan members of the steroid hormone receptor (SR) subfamily (NR3), are crucial for the transcriptional control of cellular energy metabolism. As a basal member of NR3 subfamily, ERRs are key to understand the binding of SRs to DNA from monomeric to dimeric palindrome sites. To unravel the DNA selection by SRs, we combined structural, biophysical, and phylogenetic studies. Our results showed the dimerization interface imprinted in the protein itself, where the DNA act as an allosteric driver by forming a novel extended asymmetric dimerization region (KR-box). Phylogenetic analyses suggest the dimerization asymmetry used by ERRs is an ancestral feature necessary for establishing a strong overall dimerization interface, which was progressively lost in the course evolution by other SRs.

Introduction

Nuclear receptors (NRs) are ligand-dependent transcription factors essential for the proper regulation of genes expression involved in major developmental, physiological and metabolic pathways (Billas and Moras, 2013; Gronemeyer et al., 2004). Importantly, dysregulation in their mode of action is often linked to diseases including cancer, metabolic syndrome, reproductive failure, among others, thus making NRs as promising drug targets (Lazar, 2017; Zhao et al., 2019). NRs share a common modular structure composed of a highly conserved DNA-binding domain (DBD) and a less conserved ligand-binding domain (LBD). These two domains are connected by a flexible hinge comprises the C-

terminal extension (CTE), located right after the DBD which plays an important role in the recognition of DNA binding sites. The latter are regulatory regions of target genes onto which NRs bind as homodimers, heterodimers or, more rarely, as monomers. These specific genomic sequences, known as response elements (REs), are located in the promoter-proximal locations or in enhancer-distal regions of target genes (Billas and Moras, 2013; Gronemeyer et al., 2004; Helsen and Claessens, 2014; Helsen et al., 2011).

The estrogen-related receptors (ERRs, NR3B) belong to the steroid receptor (SR) subfamily (NR3) that contains vertebrate steroid receptors (ER and the oxosteroid receptors (oxoSRs) AR, PR, GR and MR) and play a crucial role in the transcriptional control of cellular energy metabolism. ERRs are orphan receptors, but they react to various external factors, when high energy level is needed by the organism (Audet-Walsh and Giguere, 2015).

The ERRs are encoded by three paralogous genes (ERR α /ESRR1, ERR β and ERR γ) in vertebrates and only one in the cephalochordate amphioxus and ascidians which are the closest living relative to vertebrates (Bardet et al., 2006; Bertrand et al., 2010; Gomes et al., 2019). Furthermore, the ERRs represent the most ancient subfamily (NR3B1) of NR3 receptors and are present in all bilaterians such as arthropods, mollusks and vertebrates. (Bannister et al., 2013; Tennessen et al., 2011). This led to the initial naming of early metazoan NR3 as ERR (Baker, 2008), due to the NR3 members belongs to early branches of metazoan which are precursors of ERRs, ERs and oxoSRs (Baker, 2008; Hochberg et al., 2020; Khalturin et al., 2018). In fact, identification of NR3 in hemichordate at the base of the NR3A (ER) and NR3C (oxoSRs) clade, it is now clearly established that there was a late diversification of SRs among chordates. The SR ancestor is AncSR1, which was predicted based on the phylogenetic tree (McKeown et al., 2014).

ERR α was identified based on the high similarity of its DBD sequence with that of the estrogen receptor (ER α , ESR1, NR3A1) (Giguère et al., 1988). Like ER α , ERR α (as well as ERR β and ERR γ) can bind in vitro to the estrogen response elements (EREs) composed of a palindromic PuGGTCA DNA motif with a three base-pair spacer (inverted repeat RE, IR3 RE) (Tremblay and Giguère, 2007). However, in the cellular context, genome-wide studies revealed that ERR α and ER α recognize distinct DNA binding sites, ruling out the initial hypothesis of ERR α /ER α regulatory cross-talks, where ERR α would compete with ER α by binding to EREs. This suggests that ERR α and ER α differ with regards to transcriptional activity at the genomic, functional, and mechanistic levels (Deblois et al., 2009). In fact, these studies and others (Deblois et al., 2009; Dufour et al., 2007; Tremblay et al., 2010) clearly established that ERR α preferentially recognizes DNA sequences referred as ERR response elements (ERREs) that contain a single half-site of sequence PuGGTCA extended at its 5'-end by a 3 base pair TNA sequence, where 'C' is the most frequent nucleotide (N) (Dufour et al., 2007; Sladek et al., 1997; Tremblay and Giguère, 2007; Xia et al., 2019). In addition, it was shown that a small, albeit functionally important, fraction of ERR binding sites contains 18-pb ERRE/ERE hybrid binding motifs composed of the extended ERRE half-site embedded into an inverted IR3 RE similar to the EREs (Vanacker et al., 1999). These composite ERR α /ER α binding sites named as embedded ERRE/IR3 (embERRE/IR3) REs are found less than ~6 % of the promoters/enhancers target genes of ERR α /ER α . These embERRE/IR3 sites are significant to the biology of breast cancer including the genes located within ERBB2 amplicon and GATA3 (Deblois et al., 2009).

Due to the monomeric nature of ERRs binding sites, it was speculated that the ERRs function as monomers. However, functional studies clearly demonstrated ERR act as a homodimer during cofactor

recruitment and transcriptional activation (Barry and Giguere, 2005; Dufour et al., 2007; Horard et al., 2004; Huppunen and Aarnisalo, 2004; Vanacker et al., 1999; Zhang and Teng, 2000). In addition, structural data showed homodimerization of ERR α LBD similar to ER α LBD with a stable dimerization interface (Greschik et al., 2008; Kallen et al., 2007). The apparent contradiction between the existence of ERR α dimerization and the monomeric ERREs raised the fundamental questions on how the ERR α homodimer recognizes the DNA and the mechanisms that favors its DBD homodimerization on DNA. These questions cannot be easily answered with the available single NMR structural data that showed a monomer of ERR β DBD binding to its consensus ERRE (5'-TCAAGGTCA-3') (Gearhart et al., 2003) and also with other NR DBD structures, in particular those of ER α (Schwabe et al., 1993a; Schwabe et al., 1993b), AncSR1, whose sequence was reconstituted from phylogenetic analyses (Thornton et al., 2003) and monomeric NRs, such as SF1 and FTZ-F1/LRH-1 (Little et al., 2006; Solomon et al., 2005). Indeed, the molecular mechanisms for the stabilization of an ERR DBD homodimer on DNA cannot be fully understood from simple comparison with the existing DBD structures. In particular, the NMR structure of ERR β DBD is reported on a short 13 bp DNA fragment, much too short to address the dimerization.

Here, we report the first crystal structure of a ERR DBD homodimer bound to a long DNA fragment (26 bp) comprising a natural emberRE/IR3 RE. We experimentally confirmed our initial hypothesis that the DNA sequence and its associated shape features are the key driving factors in the homodimerization of ERR α on DNA (Mohideen-Abdul et al., 2017). Our structure provides a first-hand opportunity to gain insight into the molecular mechanisms of DNA recognition by ERR α which revealed a novel unexpected features associated with DBD dimerization on DNA. In particular, an extended asymmetric dimerization interface encompassing not only the D box motif, but also the region upstream of helix H2 which includes a conserved KR-box motif. Our structural analyses of ERR α DBD reveal that the DNA is the allosteric driver for the DBD dimer formation on DNA which brings a new information on the oligomerization behavior of ERR receptor. Thus, it shed the light on the evolutionary crossroad for a deeper understanding of the SRs homodimerization.

MATERIAL AND METHODS

Cloning, protein expression, and purification

Mouse ERR α DBD (70-170 amino acids), whose DBD sequence is identical to that human ERR α , was cloned in the in-house expression vector pNEAtH containing a His6 affinity tag coding sequence. The vector was transformed in to E. coli BL21 (DE3) pRARE2 strain. Cultures were grown at 37°C and protein expression was induced at OD_{600nm}= 0.6 with 1 mM IPTG for 3 hours at 25°C. The cell pellet was suspended in binding buffer (20 mM Tris pH=8.0, 400 mM NaCl, 10 % glycerol, 4mM CHAPS) and lysed by sonication. The crude extract was centrifuged at 45,000 g for 1 hour at 4°C. The lysate was loaded on a Ni affinity HisTrap FF crude column (GE Healthcare, Inc.) and the protein was eluted at a concentration of 250 mM imidazole. The hexa-histidine tag was cleaved overnight using thrombin protease. ERR α DBD was subjected to a Heparin column purification using an increasing salt gradient and followed by a size-exclusion chromatography (Superdex S75 16/60 column, GE Healthcare) using a SEC buffer (50 mM BisTris pH=7.0, 120 mM KCl, 0.5 mM CHAPS, 4 mM MgCl₂). Complexes were formed by mixing DNA to the purified ERR α DBD to achieve desired DNA: protein molar ratios and incubated overnight at 4°C before performing the crystallization assays.

DNA annealing

HPLC purified single-stranded DNAs were ordered from Sigma-Aldrich, Inc and Euromedex, France. Oligonucleotide DNA sequences are:

embERRE/IR3 (sense): 5'-ATGTCAAGGTCACCGTGACCTTTACG-3' and (anti-sense): 5'-TCGTAAAGGTCACGGTGACCTTGACA-3';

tff1 IR3/ERE (sense): 5'-GGATTAAGGTCAGGTTGGAGGAGACT-3' and (anti-sense): 5'-CAGTCTCCTCCAACCTGACCTTAATC;

traERRE (sense): 5'-TTTGTCAAGGTCACACAGTGAGTTAC-3' and (anti-sense): 5'-CATTGAGTGACACACTGGAAGTGT-3'.

The single-stranded DNA oligonucleotide and its respective anti-sense DNA fragment were mixed in a ratio of 1:1 and heated in the DNA annealing buffer (10 mM Tris pH=8.0, 100 mM NaCl, 1% DMSO, 0.1 mM EDTA) at 95° C and gradually cooled down to 4° C using a BIO-RAD PCR thermocycler. Finally, 4 mM MgCl₂ was added to the annealed DNA sample.

Crystallization and data collection

The ERR α DBD protein and the embedded DNA were mixed in a ratio of 2:1 and the final concentration of the complex (5 μ g/ μ l) was adjusted with SEC buffer. ERR α DBD/DNA complex crystals were obtained by mixing the complex with the reservoir solution in a ratio of 1:1 at 20°C using the screen Morpheus I from Molecular dimensions (Gorrec, 2009). Crystals appeared within 3 days. Crystals were cryo-cooled through flash-freezing into liquid nitrogen directly from the reservoir solution, which is already a cryo-protectant. Data were collected at the microfocus beamline PX2 of the SOLEIL synchrotron (Gif-sur-Yvette, France), using a an Eiger X 9M detector.

Structure determination and analysis

The data showed anisotropic diffraction and were processed with the autoPROC suite to handle this properly (Vonrhein et al., 2011). The space group of the crystal is P21212 ($a = 168.11 \text{ \AA}$, $b = 70.82 \text{ \AA}$, $c = 94.84 \text{ \AA}$, $\alpha, \beta, \gamma = 90^\circ$), with two molecules in the asymmetric unit and a solvent content of 57.7%. The structure was determined by molecular replacement using ER α DBD/IR3 structure (PDB code 1HCQ) (Schwabe et al., 1993b), as an initial model using the Phaser software (McCoy, 2007) from the Phenix suite (Adams et al., 2010). A primary search was carried out using the entire DNA/protein model. The structure was refined iteratively using Phenix software through model building with COOT (Emsley et al., 2010) and refined with Buster (Bricogne et al., 2017), which allows considering all reflections for refinement regardless of the crystal orientation related to the anisotropic diffraction. Overall map features were improved using the Feature-enhance map program (FEM) (Afonine et al., 2015). In addition, a few key residues were refined individually with the Polder map (Liebschner et al., 2017), where the Polder map represents an omit map that excludes the bulk solvent around the omitted region and hence allows to visualize weak densities close to the bulk solvent. FEM and Polder map analysis procedures are part of the Phenix suite (Adams et al., 2010). The DNA geometric parameters were calculated using 3DNA (Lu and Olson, 2003). All the structural figures were prepared using PyMol (Delano, 2002).

Electromobility shift assay (EMSA)

Native polyacrylamide gels were prepared to the final concentration of 5 % using the Acrylamide:Bisacrylmide solution with 10 mM HEPES-KOH (pH7.5) and several concentrations of MgCl₂. Gels were pre-run at a constant voltage of 120 for 45-60 minutes. The different DNA oligo-protein complexes were reconstituted and incubated for at least 20 minutes on ice before running at a constant voltage of 150 for 2-3 hours, followed by staining with ethidium bromide.

Isothermal titration calorimetry (ITC)

The purified DBD and the annealed DNA fragment were dialyzed separately overnight against the SEC buffer using a dialysis unit with a 10 kDa cutoff membrane. Samples were diluted to the working concentration and degassed prior to the experiments. ITC data were obtained on a Malvern Panalytical PEAQ-ITC microcalorimeter. Measurements were performed by titrating the protein (cell) with DNA (syringe). The heat of dilution for the protein and DNA were obtained by titrating them separately against the buffer. Experiments were performed at 10°C, 15°C, 20°C, 25°C and 30°C for emberRE/IR3 DNA in order to evaluate the temperature-dependency of the protein/DNA interaction. Data were processed with MicroCal PEAQ-ITC Analysis software and with AFFINImeter v 1.2.3 software using a two set of sites binding model (Piñeiro et al., 2019). This model is consistent with a specific binding site for a dimer of ERR α DBD onto the DNA and a second unspecific binding site, in agreement with native gel electrophoresis experiments. ITC data were also collected for the binding of ERR α DBD to IR3/ERE DNA (at 15°C in order to optimize the signal for this DNA sequence) and to the ERRE extended half-site DNA (at 20°C).

Sequence and Evolutionary Analysis

ERR DBD sequence was used to perform a blast search against protein data bank. Multiple sequence alignment was carried out using MAFFT program. Sequence alignments were checked and manually edited using Jalview. Final figures were prepared using ESPript.

Collected NR sequences were aligned using Clustal Omega (Sievers and Higgins, 2014). Alignments were checked manually and edited with Seaview (Gouy et al., 2009). Phylogenetic trees were built using PHYML (Guindon and Gascuel, 2003) by the LG model (Le and Gascuel, 2010) with a gamma law. The reliability of nodes was assessed by likelihood-ratio test (Anisimova and Gascuel, 2006). Ancestral character reconstruction and stochastic mapping (Huelsenbeck et al., 2003) were performed under R version 4.1.2 (R-Core-Team, 2021) using the make.simmap function as implemented in the phytools package version 1.0-1 (Revell, 2012). Character evolution was inferred using a model of symmetrical transition rates between the character states (SYM). 10,000-character histories were sampled to allow the incorporation of the uncertainty associated with the transition between different states. Inferred state frequencies for ancestral nodes were plotted using the describe.simmap function.

RESULTS

The overall architecture of ERR α DBD/DNA crystal structure

To gain insight into the molecular mechanisms of DNA recognition by ERR α DBD homodimer, we reconstituted, characterized, and crystallized the complex between the protein (mmERR α -70-170, Figures 1A and 1B) and a 26 base pair DNA containing the emberRE/IR3 RE (Figure 1C) showed to stabilize the homodimeric form of ERR α DBD (Mohideen-Abdul et al., 2017) (Table S1). The asymmetric unit contains two equivalent DNA-protein complexes (RMSD 0.83 Å) arranged in an upside-down manner, with subunits DBD1 and DBD2, forming a head-to-head homodimer on DNA. Besides, an additional subunit (DBD3) on the opposite side of the DNA aide in the crystal packing without affecting the homodimer (Figures S1A and S1B). The structure of one DNA bound with ERR α DBD homodimer is shown in Figure 1D, and the main protein-DNA interactions depicted in Figure S1C. The DBD1 is bound to the TCA 5'-extended half-site and the DBD2 is bound to the second half-site containing a 5'-TAA extended half-site, which offers the possibility to compare the roles of different 5'-extensions in the same complex (Figure 1C).

Asymmetry of the dimerization interface

The DNA sequence was key for the stabilization of the homodimer of ERR α DBD (Mohideen-Abdul et al., 2017). The embERRE/IR3 RE is composed of two palindromic half-sites are flanked by a different 5'-extension (Figure 1C). Thus, given the asymmetry in the DNA binding sites, we asked whether the asymmetry was also found in the protein dimer. At first look, ERR α subunits DBD1 and DBD2 are organized in a manner similar to the homodimers of other SRs bound to their respective IR3 REs (ER, GR, AR, PR and MR) (Hudson et al., 2014; Luisi et al., 1991; Meijnsing et al., 2009; Roemer et al., 2006; Schwabe et al., 1993a; Schwabe et al., 1993b; Shaffer et al., 2004). The two DBDs contact each other through their D-box motif (Pro116 to Glu120 in Figure 1B) via a classical “hand-shake” mode of interaction (Figure 2A). The dimerization interface comprising the two D-boxes is formed by hydrophobic packing between the methyl groups of Ala117 in each DBD and from the van der Waals contacts of Pro116 pyrrole ring which underlies Ala117 methyl group. Besides, polar contacts are also formed with Pro116 carbonyl group of DBD1 by forming a hydrogen bond with the amide group of Glu122 of DBD2 and vice versa, with relatively symmetric bonding schemes and equivalent interaction distances (3.1 Å and 2.9 Å, Figure 2A).

However, conformational asymmetry is seen in the region following D-box and prior to helix H2 (Glu122-Ala130 in Figure 1B). In DBD1, this region comprises a loop between Glu122-Thr124 and a 310 helix formed by residues of the KRRRK motif from Lys125-Lys129 (here after we referred as ‘KR-box’; Figures 1B, 1D and 2B) with a topology similar to other DBD structures, e.g. the NMR structure of ERR β /ERRE (PDB ID: 1LO1 (Gearhart et al., 2003), and crystal structures of ER α DBD (PDB ID : 1HCQ (Schwabe et al., 1993b), and AncSR1 DBD (PDB ID: 4OLN (McKeown et al., 2014). In stark contrast, in DBD2 a conformational difference is observed in the KR-box region, which folds as a loop, thereby extending the dimerization region and making it asymmetric (Figure 2B). The KR-box residues are strictly conserved for all the three ERR subtypes highlighting its functional importance (Figure 2C and S2A).

The asymmetric dimerization interface of ERR α DBD results in an increased dimerization area which arises from the interactions between residues of the KR box motif of DBD2 and DBD1 (ASA: 486 vs 468 Å², for the KR-box of DBD1 and DBD2 respectively). The side chains of Arg128 and Lys129 of DBD1 interact with the DNA phosphate backbone and Arg127 contact with Ser114 of DBD2 (Figure 2D and 2F). In DBD2, Arg128 side again participate in protein-protein interactions with DBD1 by forming an extended-loop with the help of Lys129. Besides, the intra-hydrogen bond formed between Lys129 and Ala130, where the latter residue is unique to ERR α that results in a differential electrostatic potential in the ‘KR-box’ regions. Furthermore, Arg127 of DBD2 make an additional contact with Ser114 and Pro116 (Figures 2E and S2B-D). Altogether, the structure of the ERR α DBD homodimer suggests that the conformational adaptation of DBD2 D-box and KR-box regions leads to increase in dimerization asymmetry and which eventually results in stabilization of dimer on its cognate DNA response element.

Evolution of the KR box in the NR3 subfamily

To better understand the evolutionary pressure acting on the KR box, we studied its evolution within the NR3 family. We categorized 6 main types of possible KR boxes that are shown in Figure S2E-F, allowing substitutions only within amino acid residues having equivalent physicochemical properties: (i) The relatively weakly constrained motif that contains two positions able to vary and which is found as KQRRK in the NR3Fs sequence of the basal bilaterians in the *Trichoplax* genus, as well as in the ERR of *Priapulius caudatus* (in the Protostomia clade), as RQ/NRRK in some annelid NR3Ds, and as K/RNRRK motif in chordate ERs; (ii) the relaxed K/RMSRK motif found in the cnidarian specific NR3E (Khalturin et al., 2018); (iii) The highly conserved ERR KR box with the

motif KRRRK that is present in all ERRs known of bilaterians, except that of *Priapulus caudatus*, (iv) the KHRRK motif, which is the most abundant in bilateria before the divergence of chordates at the AncSR1 node (v) the unique RRLKK motif found in amphioxus SR and (vi) the KI/FRRK motif from vertebrate SRs.

We then mapped the occurrence of these 6 types on a simplified version of the updated phylogeny as shown in Figure S2F, which is fully consistent in its topology with previous publications based on a similar dataset (Hochberg et al., 2020). The ancestral state reconstruction for every node of the phylogeny illustrates successive changes in constraints with a basal KQRRK motif found in the basal bilaterian *Trichoplax*, followed by an episode of high divergence in Cnidarians, leading to loss of the ancestral NR3 in some species (anthozoans) and retention of a diverging K/RMSRK motif in medusozoan cnidarians NR3Es, where the otherwise almost universally conserved R in the third position switched to a S. From the observations that the ancestral K/RxRRK motif (in blue in Figure S2F) is present in three distant parts of the tree (SRs/NR3Ds, *Priapulus* ERR and *Trichoplax* NR3F), we conclude that it must have been the one present at the basal nodes, including AncSR1, but we cannot formally exclude a convergent appearance of the K/RxRRK in those three branches. However, even removing the basal *Trichoplax* sequences does not change the inference of the ancestral state, given the long branch in cnidarians NR3Es. The ERR type appears then as a specific subtype of the ancestral K/RxRRK motif that has remained remarkably well constrained during bilaterian evolution. Overall our evolutionary analysis highlights the structural and functional importance of the ERR KR box and suggests that it is a constrained subtype of the ancestral motif K/RxRRK present in vertebrate steroid receptors.

Differential interactions of DBD1 and DBD2 on ERRE half sites

The sequence of the two half-sites are identical, we question whether the interactions seen between the core of DBD and the DNA major grooves composing the half-site are same. The recognition helix (H1) of ERR α interacts with DNA major groove half-site via base-specific interactions with its residues Glu97, Lys100, Lys104, and Arg105 (Figure 3A). These interactions are conserved in ER α and other PuGGTCA-binding NRs, such as RXR, RAR, or VDR (Helsen and Claessens, 2014). In particular, Glu97 has a key role in discriminating between AGGTCA from AGAACA half-site sequences which are also recognized by oxoSRs (Gronemeyer et al., 2004; Helsen et al., 2011). A similar binding pattern is observed in the two DBD subunits, but with some important qualitative differences in the electron density map of residues interacting with the bases in DNA major grooves. For DBD1, clear H-bonds interactions are observed between the side-chains of Glu97, Lys100, Lys104, and Arg105 and DNA bases of the major groove (Figure 3A and S3A). Instead, in DBD2, the side chains of Glu97 and Arg105 are clearly seen in the electron density map oriented towards C19 and G17 in chain A (Figure 3B), but Lys100 and Lys104 cannot be defined from the weak electron density map (Figure S3B). This shows the stabilization of DBD2 on the DNA is not as strong as compared to DBD1. In concordance, the experimental temperature factor of these residues, which is a measure of the disorder of each atom position, is higher in DBD2 than in DBD1, demonstrating a more dynamic conformation for DBD2. Moreover, the core of ERR α DBD is innately stabilized by a patch of hydrophobic residues from helix H1 and H2 which is observed for all subunits, including DBD3 and this unique feature is absent in ER α (Figure 3C).

Taken together, the analysis of the interactions between ERR-specific residues and the DNA major grooves are conserved, but DBD1 and DBD2, are not bound with the same strength to their cognate half-site. It further emphasizes a unique asymmetry not only at the level of the dimer interface, but also

for the interactions of the two individual subunits with the DNA major groove of each respective half-site.

The 5' DNA extension as a key determinant of the A-box conformation

Our structure offers a unique opportunity to study in atomic details the interactions between the DBD and a characteristic 5' TCA (for DBD1 and DBD3) or a less frequent 5' TAA (for DBD2) extended half-site. The CTE region contains T- and A-boxes (Figure 1B), which are visible in the electron density until the main chain of Arg158 for DBD1 and DBD2. The T-box consists of one-turn of a 310 helix (Lys146-Gly148) with Val149 packing against Phe103 and Ile107 of helix H1 before crossing the phosphate backbone and making a β -turn formed by residues Arg150 to Arg153 (Figure 4A), similar to the NMR structure of ERR β DBD on ERRE (Gearhart et al., 2003). The A-box region (Arg155-Lys162) comprises an AT-hook (Arg155, Gly156, Gly157, and Arg158) and represents the main motif for the interaction with the 5' DNA extension by insertion of this loop into the minor groove. The interaction pattern of the AT-hook motif is similar to that seen in monomeric ERR β DBD, with two glycine residues actively participating in the interactions with the 5' TCA extension (Figure 4B) Since the electron density of the side chain of Arg158 was rather weak, we performed a polder (omit) map analysis to improve the local electron density, which exclude bulk solvent around the omitted region (Liebschner et al., 2017). This analysis suggests that Arg158 is inserted into the minor groove and forms a H-bond with the O4 atom of the ribose of C25 on the complementary strand (Figures S4A and S4B). On the other hand, compared to the ERR β DBD solution structure, additional specific interactions that stabilize the AT-hook motif are seen for DBD1. In particular, salt bridge interactions between the side chains of two residues of the T-box, Arg150 and Asp152 help maintaining a stable conformation of the AT-hook which interact with the DNA backbone (Figure 4A). These observations contrast with those seen for DBD2 and DBD3, where the side chain of Arg150 interacts with the carbonyl moieties of Val154, Gly156 (A-box) and Gly157 (Figure 4C). In stark contrast, in DBD2, the AT hook glycine residues Gly156 and Gly157 do not interact with the minor groove of the 5' TAA extension, leading to loss of specific interactions with the second base (Figure 4C). Improvement of the electron density with the polder map analysis for Arg158 suggests a H-bond with the N3 position of A5 in chain B (Figure S4B). However, the strong H-bonds seen in the case of TCA 5'-extension are lost in TAA 5' extension. These observations unravel for the first time, the influence of atomic details on the 5' extension of the DBD stabilization on the DNA which emphasize the importance of the central base C/G in the TNA 5'-extension. Lastly, our structure clarifies the functional role of Tyr161 residue of the A-box. In fact, it does not directly stabilize the core of the protein as reported in the solution structure (Gearhart et al., 2003), instead, Tyr161 plays a role in the stabilization of the CTE through interaction with the phosphate backbone of the 5' extension (Figure S4D).

Our structural observations of the TCA/TAA differential interactions are supported by electrophoretic migration shift assays (EMSA) of the full-length receptor ERR α bound to ERRE response elements containing different types of 5' extensions, including TCA, TTA, TGA, TAA (Figure 4D). In EMSA, the concentration of MgCl₂ was progressively increased in order to determine the conditions for the best binding specificity. Simultaneously, we examined the binding of ERR α to a negative control DNA sequence for increasing MgCl₂ concentration. Above 50 μ M MgCl₂, we lose the non-specific binding of the protein to the negative control sequence. For higher concentration of MgCl₂ (100 μ M MgCl₂ (gel III) and 250 μ M MgCl₂ (gel IV)), the binding of ERR α to the TTA (lane 4) and TAA (lane 6) containing ERRE is progressively lost, in contrast to the TCA (lane 3) and TGA (lane 5) containing ERRE. This suggests that the binding to the TCA/TGA 5' extensions is stronger than

to a TTA/TAA extension. Together with the structural analysis, our data unveil the molecular mechanisms for the preferential binding of the ERR α CTE to TCA and TGA 5' extended half-sites.

Asymmetry in the positioning of ERR α DBD homodimer on DNA as compared to related NRs

To gain a global understanding of the binding mode of the ERR α DBD homodimer on DNA, we compared the structure of ERR α DBD with ER α DBD (Schwabe et al., 1993a; Schwabe et al., 1993b) and the reconstructed ancestral steroid receptor AncSR1 DBD (McKeown et al., 2014). The superimposition study with ER α DBD (DBD1 subunit as a reference) show that the corresponding second subunit (DBD2) of the homodimers is not placed in the same manner (Figure 5A). In fact, ERR α DBD2 is rotated counter-clockwise when looking down the DNA axis compared to the symmetrical ER α DBD (Figure 5B). This movement impacts the position of the two major helices, H1 and H2, and affects the positioning of D-box loops and the extended dimerization region including the KR-box. When the dimer interface is visualized along the pseudo-symmetry axis, the entire dimerization region of the second subunit is strongly displaced (Figure 5B).

We further noticed that AncSR1 DBD homodimer bound to IR3/ERE presents an asymmetric conformation (Figure 5C and 5D). The KR-box of one AncSR1 DBD subunit is folded as a 310 helix with the last two residues interacting non-specifically with DNA, while the KR-box of the other subunit unfolds as a loop and interacts with the neighboring subunit (Figure 5D). However, the loop of AncSR1 is kinked inward and forms a tiny dimerization interface. Whereas, in ERR α DBD dimer the KR box loop extend outwards forming a larger interface which contributes to an increased dimerization surface (Figure 5D).

The similarity between ERR α and AncSR1 suggests that the asymmetry of the dimerization interface is an ancestral feature of receptor homodimerization which is necessary for establishing strong contacts between the two D-box regions and the extended dimerization region and thus increasing overall dimerization interface.

DNA is the allosteric driver for cooperative ERR α DBD dimerization

To gain biophysical insights in dimerization mechanism of ERR α DBD on emberRE/IR3, isothermal titration calorimetry (ITC) was used. Our previous work with different types of DNA REs showed that ERR α DBD behaves as a monomer (Mohideen-Abdul et al., 2017). The analysis of the ITC data indicates a dimer of ERR α DBD binds to emberRE/IR3 DNA in nanomolar range ($K_d = 4\text{--}25$ nM) (Figure S5A-E and Table S2). The event is strongly enthalpy driven from 10-30°C range with a large unfavorable binding entropy. A temperature-dependence study shows that the heat capacity change $^{\circ}C_p$ for emberRE DNA binding is highly negative ($^{\circ}C_p = -473$ cal/mol/K, Figure S5F), which is indicative of a significant local folding coupled to site-specific DNA binding. An initial non-specific binding event is observed at lower molar ratios, when the protein is in surplus over DNA, with unfavorable binding enthalpy and positive entropic contribution. These unspecific protein/DNA interactions can be attributed to the avidity of DBD for DNA at high protein concentration, as well as to ion pairing formation (Oda et al., 1998; Peters et al., 2004; Privalov et al., 2010). In contrast, ITC data show that only one monomer of ERR α DBD binds specifically to tra ERRE (Figure S5G) and to tff1 IR3/ERE (Figure S5H) in agreement with previous biophysical results (Mohideen-Abdul et al., 2017). However, the ITC data analysis indicates significantly different thermodynamic profiles for these two types of REs. Regarding the tff1 IR3/ERE, one monomer binds specifically (stoichiometry $n=1.0$) in the low nanomolar range ($K_d = 28.2$ nM) and this interaction is highly enthalpy-driven ($^{\circ}H_1 = -11.7$ kcal/mol) with an unfavorable entropic component ($-T^{\circ}S_1 = 1.7$ kcal/mol) (Suppl. Fig. 5H and Suppl

Table 2). A second unspecific binding event, likely of a second monomer ($n=0.7$), is observed on this DNA sequence and is resulting from a strongly entropy-driven interaction ($^{\circ}H_2=0.9$ kcal/mol, $-T^{\circ}S_2=-12.3$ kcal/mol). Regarding binding of ERR α DBD to the tra ERRE extended half-site, a monomer binds in the low nanomolar range (stoichiometry $n=1.0$, $K_d = 1.1$ nM) with an enthalpy-driven interaction ($^{\circ}H_1=-12.0$ kcal/mol, $-T^{\circ}S_1=0.0$ kcal/mol), as well as a second, likely unspecific monomer, with a rather similar binding profile ($n=1.3$, $K_d=29.7$ nM, $^{\circ}H_2=-16.1$ kcal/mol, $-T^{\circ}S_2=6.0$ kcal/mol). The ITC data are furthermore consistent with native gel electrophoresis analysis of the protein/DNA complexes, where a first monomer and then a dimer are seen for the embERRE/IR3 RE (Figure 5I), but only a monomer is observed for tff1 IR3/ERE DNA (Figure 5J).

ITC analyses indicate the dimerization of ERR α DBD on embERRE/IR3 is a cooperative mechanism, which is driven by DNA. It is well established that in addition to the base-specific recognition mechanism, the DNA shape plays a key role in the specific recognition of DNA by proteins. The DNA shape parameters refer to the geometrical and physical properties of DNA, such as minor and major groove widths, twist and roll parameters can be predicted based from the mere DNA sequence (DNAShape software (Rohs et al., 2009; Zhou et al., 2013)). When applied to the various ERR binding sequences, the analysis allowed us to suggest that the sequence-dependent DNA shape is a key structural component promoting dimer formation (Mohideen-Abdul et al., 2017). Our crystal structure thus gave us the possibility to calculate the experimental DNA structural parameters and to discuss them in light of our initial hypotheses (Lu and Olson, 2003) (Figures 6A-D). The experimental minor groove width (MGW) shows two strong minima in the middle of each half-site (Figure 6A) as predicted, confirming our hypothesis that the second DBD requires a narrow minor groove for dimerization (Mohideen-Abdul et al., 2017). However, the minima of the experimental and predicted MGW profiles are not identical. In the prediction, the two minima are located at position ± 6 (corresponding to the second nucleotide of the half-site). This is true for the first minimum of the experimental MGW profile (Figure 6A), but not for the second one, which is found at position +5, and thus displaced by one nucleotide with respect to the predicted symmetrical value. Hence, the experimental MGW profile is asymmetric despite an equivalent half-site composition. This suggests that the binding of the dimer is conditioned by a narrow DNA MGW, but it affects reciprocally the DNA shape in the second half-site. An asymmetry can also be seen for the experimental major groove width and roll parameter (Figure 6B and 6C). In particular, the width of the major groove at position +2 (T/A) in the second half-site adopts a strong minimum, displaced compared to its position in the first half-site, leading to a strong asymmetry. This observation can be rationalized in terms of local base-pair deformation, in particular visible from the alternation of positive and negative roll values along with the spacer and the second half-site sequence. The detailed analysis suggests that Lys100 in H1 interacts with two consecutive G nucleotides (at positions +5 and +6), leading to local DNA deformation. These observations strongly contrast with the experimental DNA shape parameters of ERE/IR3 as calculated from the structure of ER α DBD on this cognate sequence (Figures 6E-6H). In the latter case, the two half-sites are perfectly symmetrical with respect to the central base pair (position 0), both in composition and DNA shape. The formation of these molecular interactions may be triggered by differences in the dimerization interface of the DBD protein. This is not observed for the roll values of ER α binding to ERE/IR3 which are always positive. Similarly, in the structure of AncSR1 bound to ERE/IR3, no major variations of the MGW, roll and twist parameters are observed along the DNA sequence (Figures S6A-S6D). Altogether, our observations suggest that the binding of the second ERR α DBD subunit to its cognate DNA is favored by a narrow minor groove. However, the binding of the protein leads concomitantly to severe deformation of the DNA shape for adequate positioning of the two DBD subunits. Hence, the observed conformation of the ERR α DBD dimer differs from that of

ER α DBD or AncSR1 bound to palindromic ERE/IR3 sequences, highlighting the peculiarity of the molecular mechanisms that lead to the asymmetric homodimerization of ERR α DBD on DNA.

DISCUSSION

Cooperative homo- or heterodimerization is a mechanism used by NRs to strengthen their binding to its respective REs and also increase their DNA binding site repertoire in the genome. However, the ERRs are stark contrast with this general view. Because, most of their DNA targets are single extended half-site elements, suggesting originally that the ERRs act as monomers (Sladek et al., 1997). But, this view is challenged by a large set of functional data which proved that the ERRs are obligate homodimers for the binding of coactivators and transcriptional activation (Barry and Giguere, 2005; Dufour et al., 2007; Horard et al., 2004; Huppunen and Aarnisalo, 2004; Vanacker et al., 1999). Thus, the mode of ERR binding to its cognate DNA remained elusive and controversial. To address this question, we focused on the isolated ERR α DBD bound to emberRE/IR3 RE, where we showed that a ERR α DBD forms a stable dimer (Mohideen-Abdul et al., 2017). The emberRE/IR3 REs are natural ERR α binding sites found in the human genome and shown to carry important biological functions in the etiology of breast cancer (Dufour et al., 2007). We further hypothesized that the sequence-dependent DNA shape of the ERR α binding sites was pivotal for the cooperative binding of ERR α homodimer to DNA. Therefore, we solved the first crystal structure of the ERR α DBD dimer on a natural emberRE/IR3 RE and unveiled the molecular determinants for the stabilization of the ERR α DBD homodimer on DNA.

The ERR α homodimer is globally organized in a similar manner to other steroids NRs in which a strong cooperative homodimerization is observed on their cognate IR3 REs (Denayer et al., 2010; Helsen and Claessens, 2014; Helsen et al., 2011; Luisi et al., 1991; Meijsing et al., 2009; Schwabe et al., 1993a; Schwabe et al., 1993b). Furthermore, the recognition of the core half-site sequence by each DBD subunit is identical to what is observed for ER α DBD and AncSR1 DBD. The base-specific interactions between residues of the recognition helix H1 and the AGGTCA half-site nucleotides of the major groove are preserved (Figure 4A), suggesting an evolutionary conserved recognition mechanism. These molecular mechanisms of specific DNA recognition evolved for oxoSRs through permissive substitutions into a different type of binding, both in terms of protein residues and DNA half-site sequence (AGAACA), but remained identical for the ER and ERR subfamilies (Ortlund et al., 2007). However, the ERR α DBD homodimer differs from the other SR DBD structures, by forming an asymmetric assembly. The asymmetric dimerization region covers the interface formed by the D-box of each subunit and a unique extended dimerization interface formed between the KR-box of DBD2 and the N-terminal part of the D-box of DBD1 (Figure 2E). In fact, the KR box of DBD1 features a 310 helix makes interactions with DNA (Figure 2D), whereas the KR-box of DBD2 forms an extended loop and makes additional interactions with the adjacent DBD1, leading to an increased dimerization surface area and eventually the stabilization of the dimer (Figures 2B & 2E). This asymmetric dimerization interface in a homodimer complex is highly unusual and highlights the flexibility and adaptability of the NR DBD architecture.

Our evolutionary analysis suggests that the role of the KR-box in the dimerization process has been overlooked for other SRs. The residues of the KR-box (KRRRK) are highly conserved in all ERR sequences which represents a constrained version of the box found in ERs (KNRRK) and oxoSRs (KF/IRRK) as well as in the reconstituted ancestral receptor AncSR1 (KHRRK) (Figure 2C). In fact, a single amino acid change separates the ERR KR box from other oxoSRs box (K/RxRRK) motifs

(Figure 2C). In ER α DBD or oxoSR DBD bound to their corresponding IR3 REs, no significant change in the conformation of the KR box is seen, despite a slight asymmetry of the distances involved in the interaction network formed by the KR-box residues (Figure 5). In contrast, the ancestral steroid receptor AncSR1 DBD binds to an AGGTCA IR3 RE in an asymmetric manner. However, since the sequence of AncSR1 was predicted from phylogenetical analysis (McKeown et al., 2014), it cannot be considered as a real receptor isolated from a living organism. Besides, the conformational change of the KR box motif to a loop for the second subunit leads to no additional interactions, a negligible increase of the dimerization interface and consequently no further stabilization of the dimer. Along with these observations is a minor deformation seen for the IR3 DNA fragment bound by AncSR1 upon binding of AncSR1 DBD dimer (Figure S6A). In contrast, in the case of ERR α DBD, the DNA shape of embERRE/IR3 -characterized by physical properties such as minor and major groove widths, twist and roll parameters- suggests that the binding of the protein leads to the deformation of the DNA to adequately position the ERR α DBD dimer. This DNA deformation reinforces the overall interaction network and the stabilization of the dimeric complex. Hence, the DNA sequence is the allosteric driver for the formation of the asymmetric dimerization of ERR α DBD. Thus, the asymmetric binding of the ERR α DBD homodimer represents the most extreme case of adaptation in the binding of two, sequence identical, protein subunits to a relatively symmetrical IR3 binding site.

Remarkably, the experimental values of DNA geometrical parameters obtained from our crystal structure are extremely consistent with the ones we predicted earlier (Mohideen-Abdul et al., 2017) based on the DNA sequence (Rohs et al., 2009; Zhou et al., 2013). This demonstrates that the sequence-dependent DNA shape is crucial in the initial recognition process by the receptor, especially a narrow minor groove in the flanking region of the second half-site. In fact, the DNA shape favors the positioning of the dimer on the DNA, even though DBD2 does not make the most appropriate interactions with DNA. But, this is largely compensated by the conformational asymmetry of the protein interface, which is a key element in the stabilization of the dimer on DNA. This observation is supported by the weak electron density of some protein side chains of DBD2 that make base-specific interactions and by the corresponding higher B-factors.

Our structure sheds further important highlights on crucial molecular determinants for the binding of ERR α DBD to different 5'-extended sites. DBD1 binds to a TCA extended half-site, whereas DBD2 binds to a TAA extended half-site, which is a less favorable extension in terms of stability and transcriptional activation (Barry et al., 2006; Vanacker et al., 1999). Our structure emphasizes the strong intricate H-bond network formed between the receptor CTE and the minor groove TCA extension, whereas interactions are much fewer for the TAA extension, which is in agreement with our EMSA assays and functional data. We further clarified the structural role of the functionally important residue Tyr161, the last residue of the A-box, in the stabilization of the ERR-DNA complex, through interactions with the DNA backbone (Figure S4D), rather than with the hydrophobic protein core as suggested in the solution structure.

In our work, we considered the special case of the embERRE/IR3 binding site. However, most of the ERR binding sites were identified as single extended half-sites (ERREs) (Deblois et al., 2009; Dufour et al., 2007; Sladek et al., 1997). Irrespective, the ERRs function as dimers to activate transcription and must therefore bind to dimeric DNA binding sites, composed of the specific extended half-site and of a 3' DNA sequence for the 'pseudo' second half-site with no consensus. Our structure highlights the structural principles allowing the binding of the ERR homodimer to ERRE binding sites (Figure 7). Specifically, it suggests that the molecular mechanisms of dimerization are imprinted in the

protein itself through the formation of an extended dimerization interface arising from the KR-box region. Since the ERRs regulate gene transcription upon activation by protein cofactors, rather than by small ligands, we propose that, given the imprinted asymmetric behavior, the homodimer will be further fully stabilized onto DNA upon cofactor binding and thus fully functional for gene regulation. Work from our lab is in progress to confirm this model.

Finally, our results provide insights and new perspectives into the evolution of the DNA recognition and binding specificity by the steroid nuclear receptors. The ERRs represent one of the most ancient and still existing group of SRs with a remarkable sequence conservation throughout evolution (Bannister et al., 2013; Tennessen et al., 2011). We hypothesize that such a sequence conservation is linked to the preservation of biological functions of ERRs and may be related to the crucial and most ancestral processes in the regulation of energy metabolism. Indeed, in *Drosophila* it has been shown that ERR is instrumental in the metabolic program that coordinate the conversion of nutrients into biomass to support larval growth by upregulating the transcription of metabolic genes implicated in glycolysis, pentose phosphate pathway and lactate production (Tennessen et al., 2011; Tennessen and Thummel, 2011). Although considered mostly with the angle of tumor growth, the observation that in mammals ERR direct a transcriptional switch that supports oxidative metabolic pathways is going in the same direction (Tremblay and Giguère, 2007). This suggest that the coordination between growth and maturation, a feature often neglected, but central for metazoan life cycle which may be the ancestral function of ERRs.

This ancestral function may be a critical functional element for the evolutionary conserved DNA recognition mechanisms by the ERRs. Our crystal structure of ERR α DBD suggests that asymmetry is used for function, allowing the better stabilization of a functional dimer on the cognate DNA sequence, in the presence of coactivators. Given the ancestral character of the ERRs, the ancestor of the ER/ERR/SR clade may have similar function. Evolutionary diversification in bilaterians further led to the emergence of the common ancestor of the ERRs and the common ancestor of the bilaterian NR3 group containing the steroid receptors, which may have been an AncSR1-like molecule. Our structure and the comparative analysis with the structure of the AncSR1 DBD suggest that the asymmetry still remained in AncSR1 as a structural remnant, but is likely not used for the homodimer stabilization and hence for its function. Therefore, our crystal structure could serve as an incomparable model for the understanding of the structural evolution in the DNA binding recognition by the steroid nuclear receptor family that are of medical and therapeutic outmost importance.

DATA AVAILABILITY

Protein Data Bank accession number: The crystal structure of ERR α DBD has been deposited to the Protein data Bank (www.pdb.org) under the accession number xxxx.

ACKNOWLEDGEMENTS

The authors acknowledge the support and the use of the French Infrastructure for Integrated Structural Biology FRISBI ANR-10-INSB-05-01 and Instruct-ERIC, a Landmark ESFRI project, the French Proteomic Infrastructure ProFI ANR-10-INBS-08-03. The authors thank the Structural Biology and Genomics platform of IGBMC, the SOLEIL synchrotron and William Shepard for assistance in using beamline.

FUNDING

This work was funded by the Agence Nationale de la Recherche (ANR) grants (xxxxxx) and by CNRS, INSERM, The Université de Strasbourg and l'Alsace contre le Cancer. This study was supported by the grant ANR-10-LABX-0030-INRT, a French State fund managed by the Agence Nationale de la Recherche under the frame program Investissements d'Avenir ANR-10-IDEX-0002-02. P. V acknowledges the 'Ligue Contre le Cancer' and the 'Association pour la Recherche sur le Cancer' (ARC) for financial support.

CONFLICT OF INTEREST

The authors declare to have no conflict of interest.

Figures

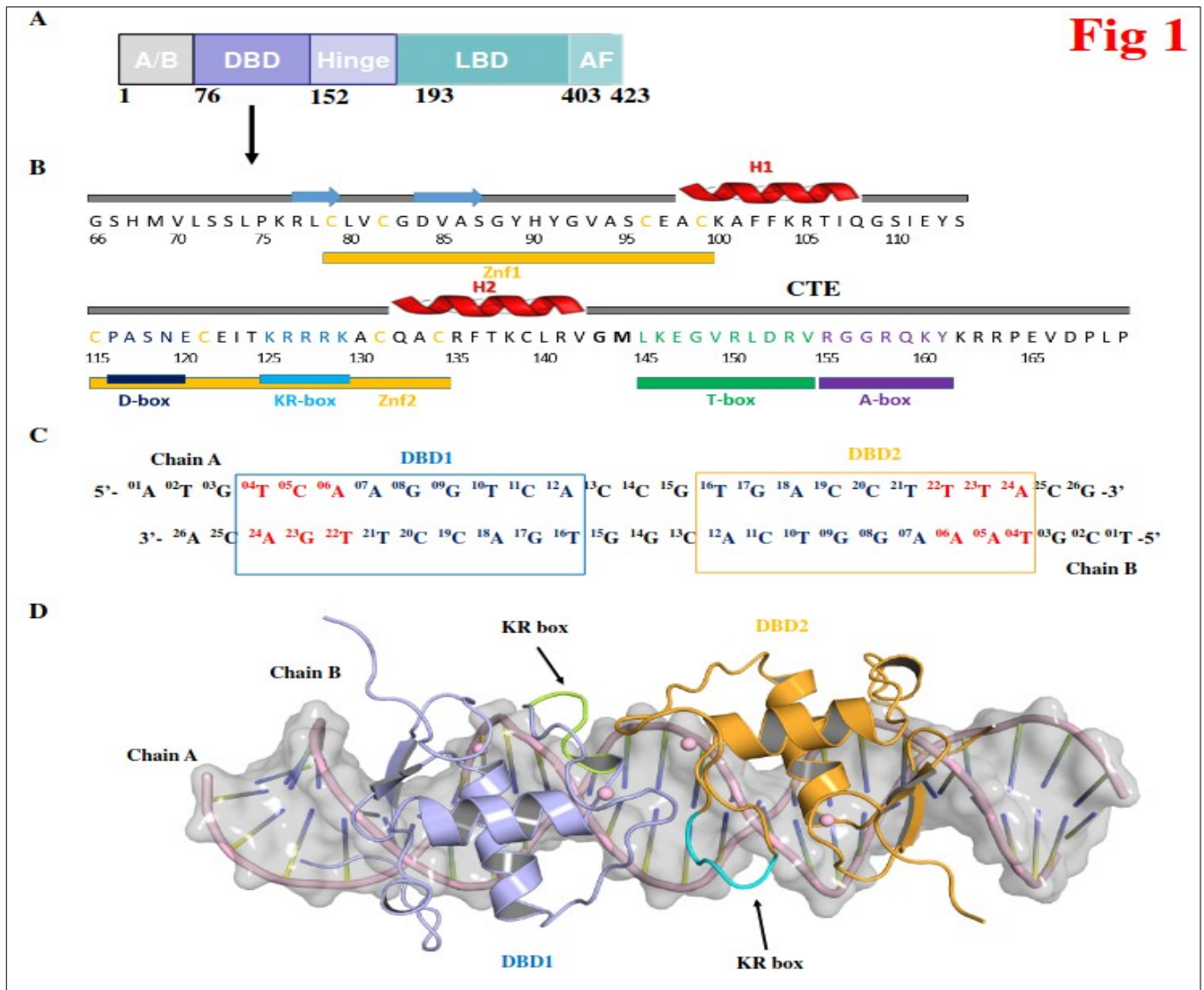


Figure 1

A: Domain organization of ERR α . **B:** ERR α DBD protein sequence and schematic secondary structure representation. **C:** Sequence of the 26 bp DNA fragment, emberERRE/IR3, used for crystallization where the two binding sites for DBD1 and DBD2 are outlined with a box. **D:** Cartoon representation of the homodimer of ERR α DBD bound to the emberERRE/IR3 response element, with DBD1 in light blue and DBD2 in orange color. The KR-box indicated in **Fig. 1B** is highlighted on the structure.

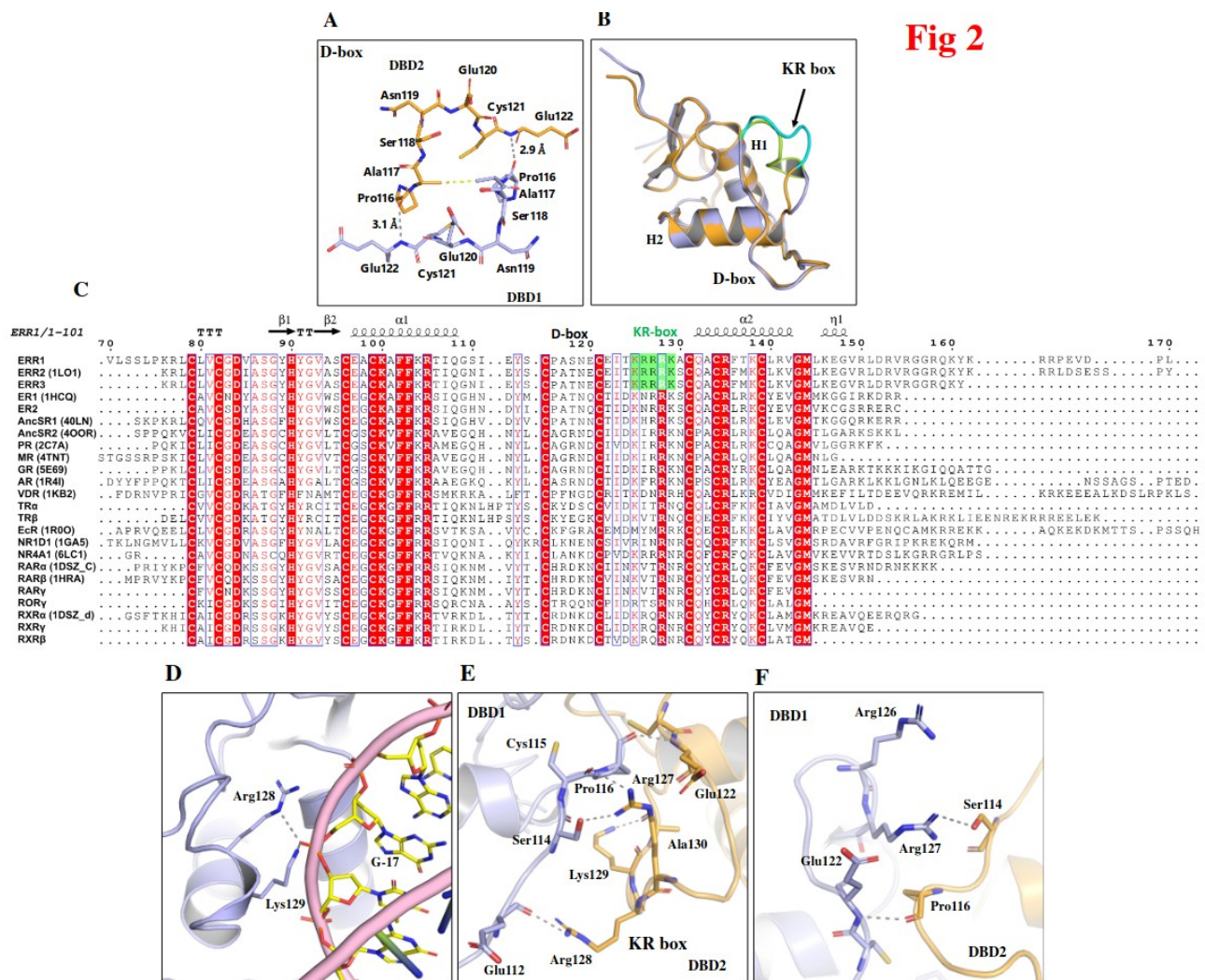


Figure 2 : The asymmetry of the dimerization interface involves a conserved KR-box motif

A. Enlarged view of the residues belonging to each of two D-boxes (for DBD1 and DBD2) that form the canonical dimerization interface. Residues are shown with a stick representation, with carbon atom in light blue for DBD1 and orange for DBD2, red for oxygen, blue for nitrogen and yellow for sulfur atom. **B.** Cartoon representation of the KR-box region for the subunits superimposed, with DBD1 in light blue and DBD2 in orange color. **C.** Multiple sequence alignment of the human ERR sequences with those of the other steroid NRs (NR3 subfamily (AncSR1, ER, AncSR2, AR, GR, MR, PR) and NRs of the subgroups NR1 (TR, RAR, PPAR, ROR, VDR), NR2 (RXR and HNF4), NR4 (NGF-IB, SF1), showing xxx . The sequence numbering corresponds to the HsERR α sequence and the colors relate to the percentage of sequence identity. The top panel represents the secondary structure elements of ERR α . **D.** Stick representation of the interactions between the residues Arg128 and Lys129 of DBD1 with the DNA phosphate backbone at the level of G17 of the complementary strand (chain B). **E-F.** Enlarged view of the dimerization interface interactions between DBD1 and DBD2 looking at the region of the KR-box of (**E**) DBD2 and (**F**) DBD1. Residues are shown with a stick representation, with carbon atom in light blue for DBD1 and orange for DBD2, red for oxygen, blue for nitrogen and yellow for sulfur atom.

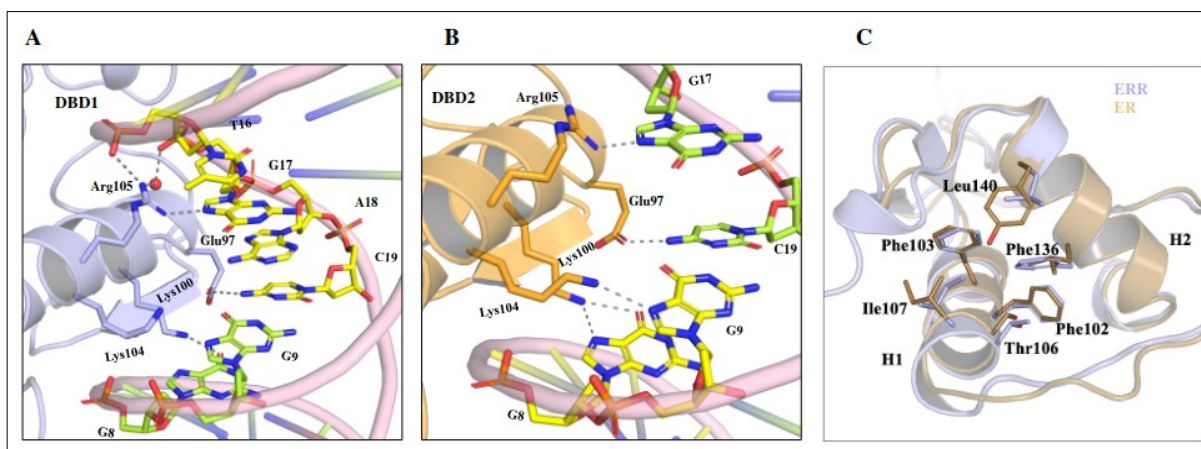


Figure 3 :Major groove interactions in DBD1 and DBD2

A-B. The residues Glu97, Lys100, Lys104, and Arg105 in H1 interact with the major groove half-site via base-specific interactions, in a similar manner for DBD1 (**A**) and DBD2 (**B**), but the interaction network is much stronger in DBD1 than in DBD2, as indicated by the structural water molecules seen in the electron density of DBD1, but not of DBD2. In addition, the side chains of residues Lys100 and Lys104 are only well defined in DBD1, but in DBD2 (see **Suppl. Fig. 3A-B**). **C.** Enlarged view of the superimposition of one subunit of ERR α DBD (DBD1 in light blue) with the corresponding one in ER α DBD in the region encompassing H1 and H2. The core of ERR α DBD is stabilized by a pocket of hydrophobic residues, Phe102, Phe103, Thr106, Ile107, Phe136, and Leu140. The ERR α residues, Phe136, Leu140, and Thr106 are replaced in ER α , by Ile, Tyr, and Ser, respectively, weakening the contribution of hydrophobic contacts seen for ERR α . In ERR α , Ile107 forms hydrophobic contacts with Phe103 (H1) and Phe136 (H2), whereas in ER α Ile interacts only with the Phe residue of H1. The second interaction is not possible due to the Phe136 residue in ERR α is replaced by Leu that pushed away from H2 by a Tyr residue in H2. Phe136 further strengthens the core interactions through an intricate network of π - π stacking interactions.

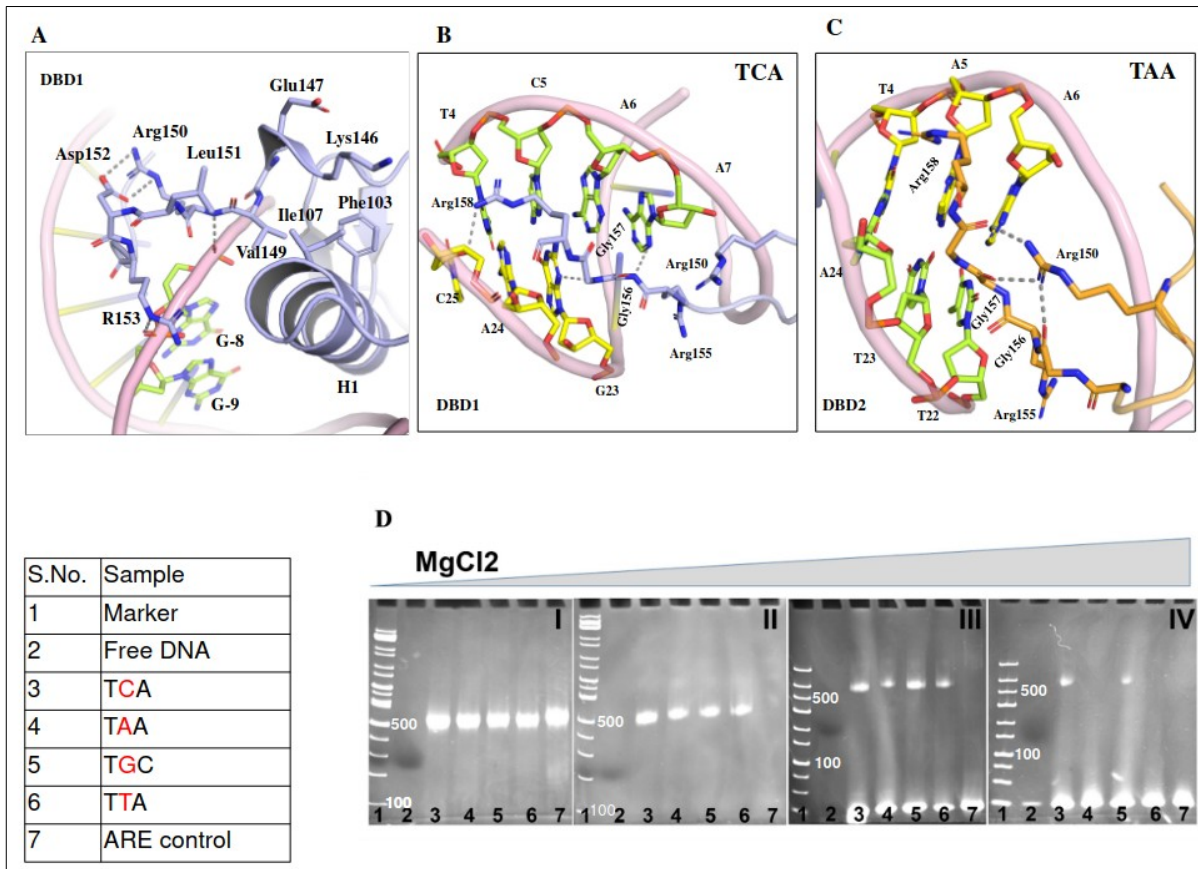


Figure 4 : TCA and TAA 5'extensions are key in the conformation of the A-box

A. The conformation of the T-box (residues 145-154) adjacent to the A-box (residues 155-161) is well defined and helps the A-box loop to penetrate into the minor groove at the level of the 5' extension, shown here for DBD1, but conserved for DBD2 and in solution. **B.** Interaction pattern of the CTE residues of DBD1 with the TCA 5'extension. **C.** Interaction pattern of the CTE residues of DBD2 with the TAA 5'extension. An almost complete loss of H-bond interactions is observed, demonstrating the weakening of the interaction xxx. **D.** EMSA shows the significant role of 5'- extension TNA nucleotides. EMSA gel for different concentration of MgCl₂ 25, 50 100, 250 μM (gels I to IV); lane 1: marker, lane 2; free DNA control; lane 3 TCA ERRE; lane 4: TTA ERRE; lane 5: TGA ERRE.

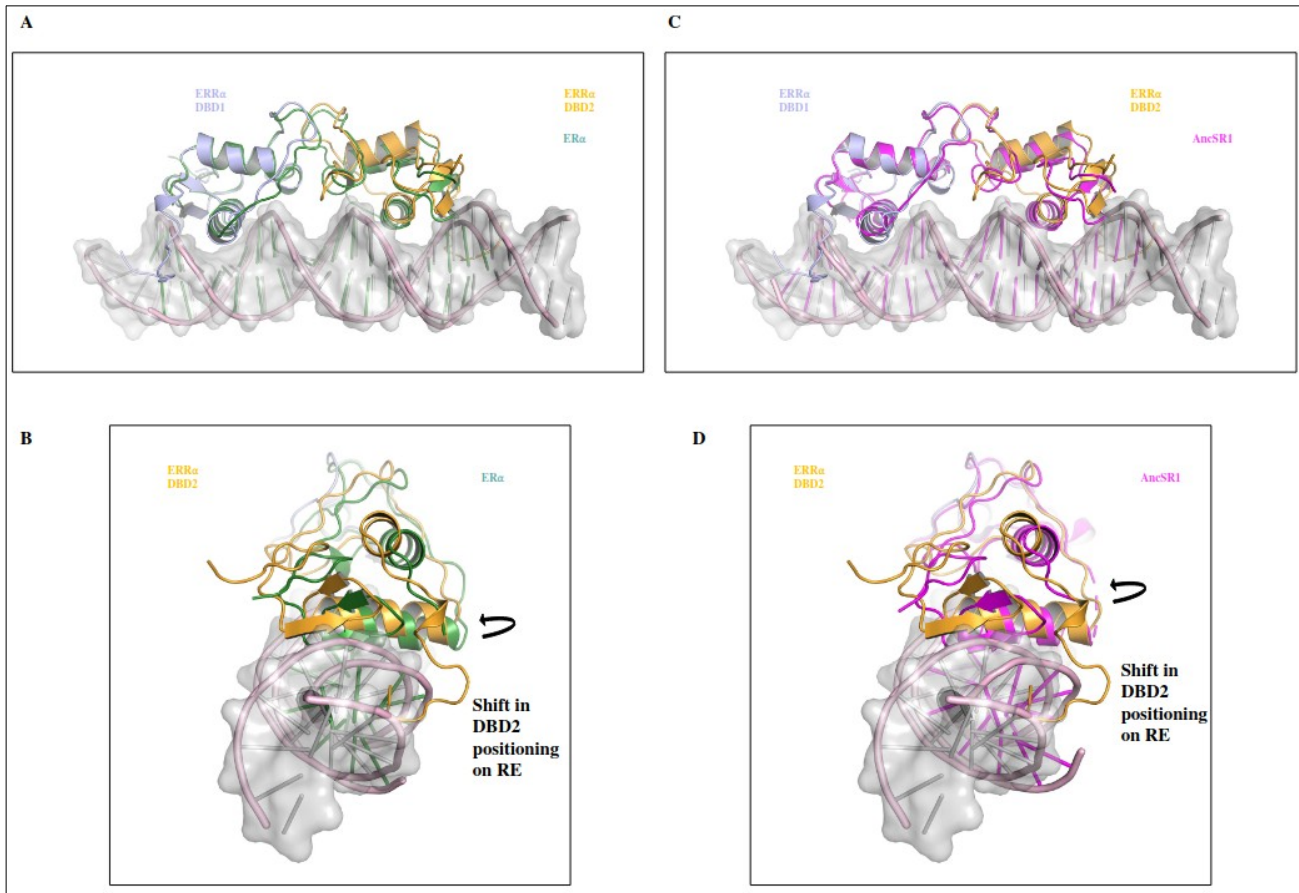


Figure 5 : The KR-box motif is an ancestral feature crucial for the asymmetric organization of ERR α DBD on DNA compared to ER α DBD and AncSR1 DBD

A-B. Superimposition of ERR α DBD and ER α DBD on DNA (2 orientations). **C-D.** Superimposition of ERR α DBD and AncSR1 DBD on DNA, with (C) a global view and (D) an enlarged view in the KR-box region of DBD2.

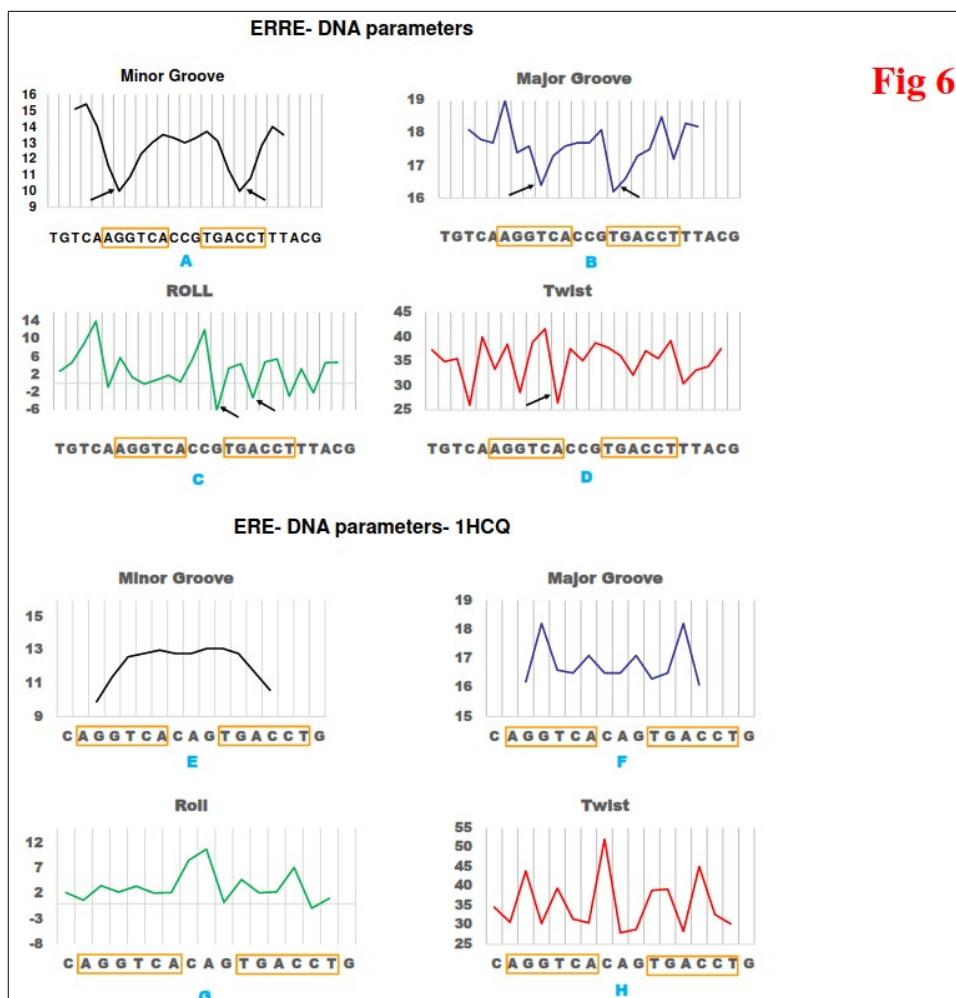


Fig 6

Figure 6
DNA geometrical parameters of ERRE (A-minor groove, B-major groove, C-Roll, and D-Twist), ER α (E-minor groove, F-major groove, G-Roll, and H-Twist).

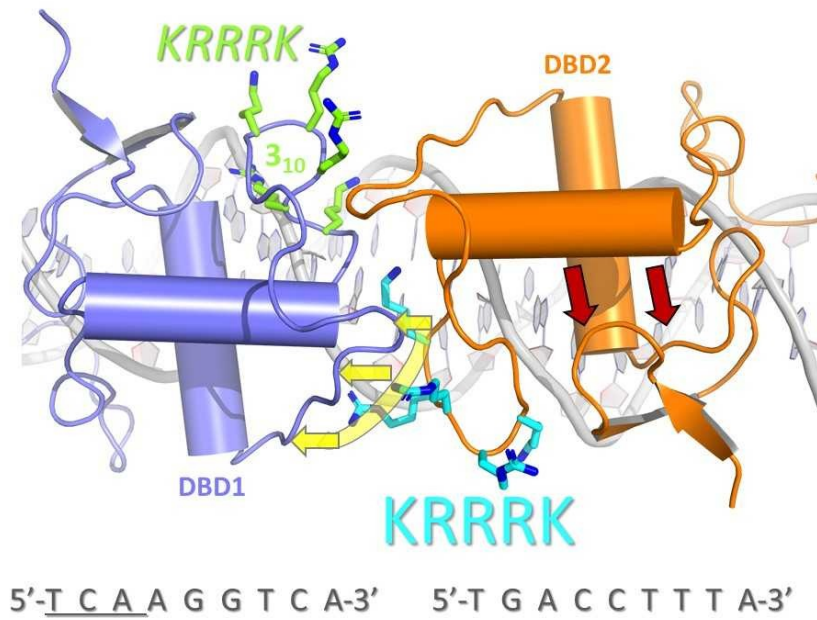
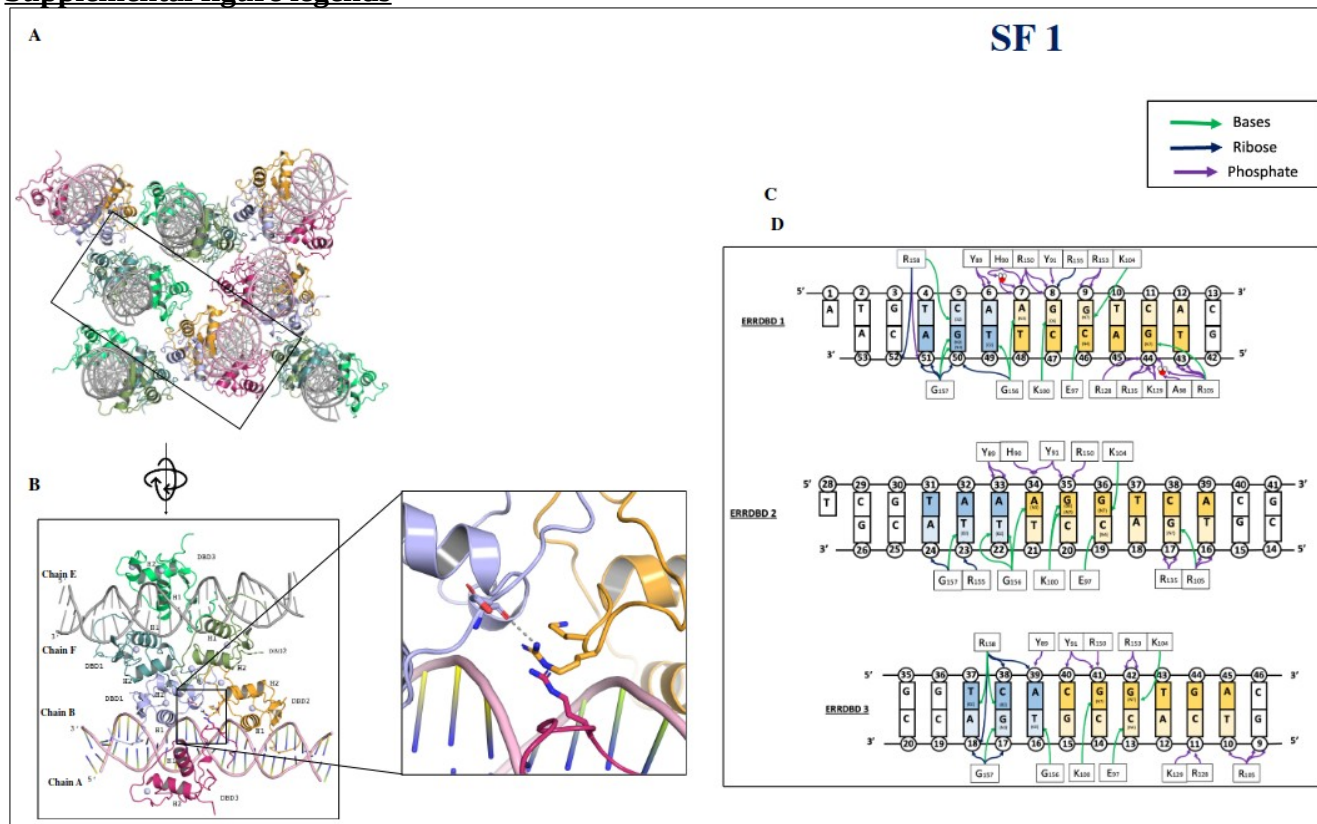


Figure 7: Schematic drawing of the proposed model for the binding of the ERR DBD homodimer to its cognate DNA response element.

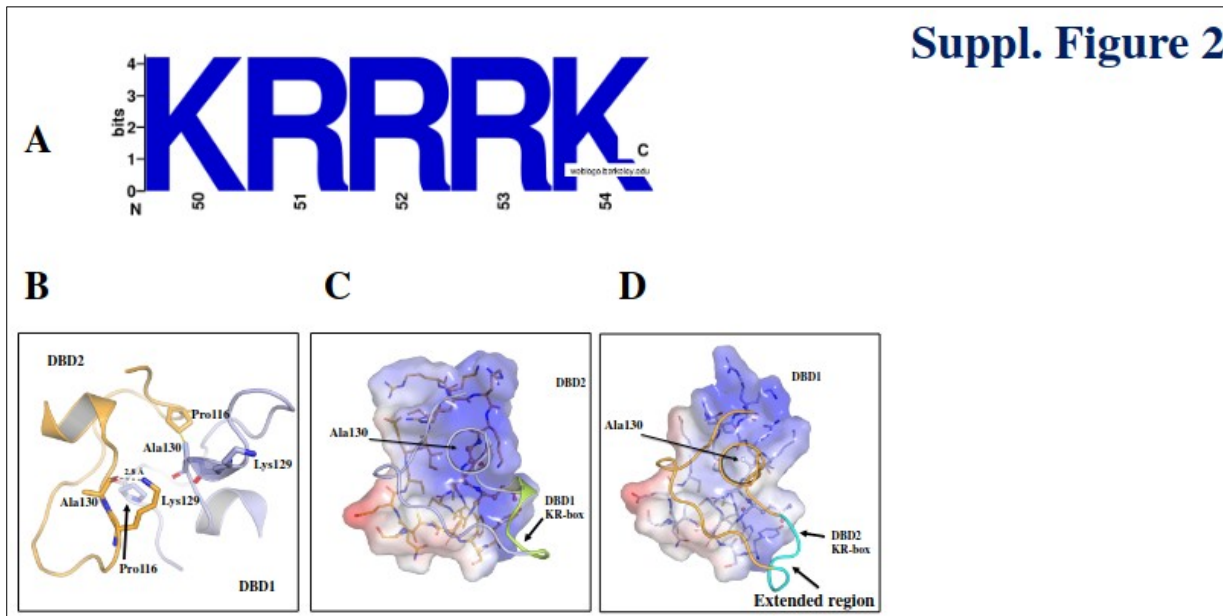
In addition to D-box/D-box interactions, the dimerization interface is reinforced by the presence of an extended interaction area arising from the KR-box region of DBD2 that unfolds as a loop and interacts with the dimerization region of DBD1, as indicated by yellow arrows. Notice that the dimerization interface is asymmetric with different KR-box conformations for the two subunits (highlighted by the green and yellow colors) and necessary for proper stabilization of the DBD homodimer on DNA (as suggested by the red arrows). DBD1, shown in blue, is the subunit bound to the extended half-site and DBD2 is shown in orange.

Supplemental figure legends



Suppl. Figure 1:

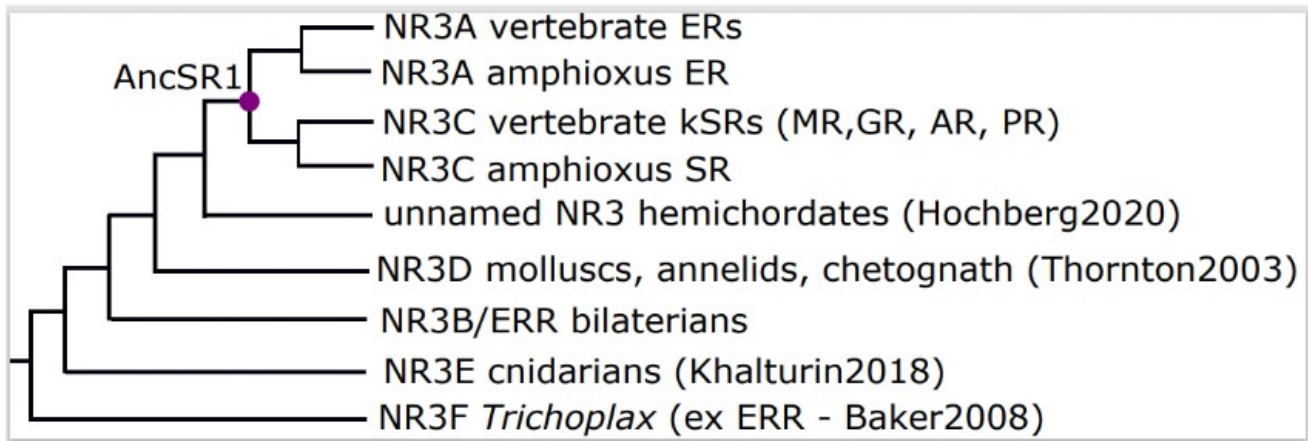
A. Cartoon representation showing the overall arrangement of the molecules found in the asymmetric unit which comprises two DNA molecules, each bound by the biological relevant ERR α DBD dimers (DBD1 in light blue and dark blue and DBD2 in orange and forest green, respectively) and a third monomeric DBD (DBD3 in pink and green, respectively) bound on the opposite side of the DNA. **B.** Crystal packing showing that the protein-protein contacts between the molecules found in the asymmetric unit are mainly driven by DBD3. **C.** Enlarged view of the interactions made between DBD3 and the homodimer on the opposite side of the DNA, showing that the DBD3 molecule barely interacts with DBD1 and DBD2. **D-F.** Interactions schemes of **(D)** DBD1, **(E)** DBD2 and **(F)** DBD3 with DNA, where the interactions between the protein residues and DNA are depicted for in green for bases, black for ribose and magenta for phosphate.



Suppl. Figure 2: Asymmetry and Conservation of the KR-box

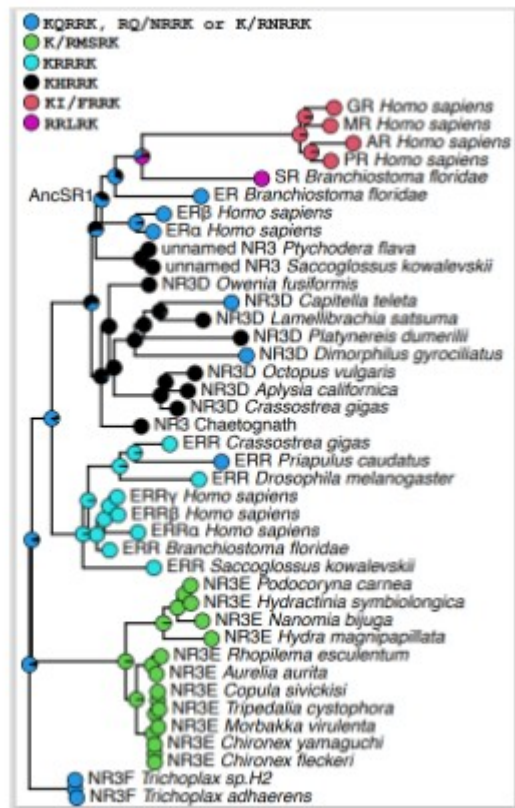
A. Multiple sequence alignment of the ERR sequences found for different organisms along the evolutionary tree. **B.** Conservation logo obtained from the MSA of the ERRs indicating the strict conservation of the sequence KRRRK motif in the ERRs. **C.** Enlarged view of the region of Ala130 in both subunits DBD1 (in light blue) and DBD2 (in orange), that is seen to contribute to the asymmetric dimerization interface. Ala130 is conserved in the ERR α sequences and replaced by the small serine residue in ERR β , ERR γ , and AncSR1. Other NRs have a large residue at this location, such as Arg, His or Asn. In DBD1, it is involved in hydrophobic contacts with Pro116, whereas in DBD2, the flexibility of the extended loop conformation of the KR-box allows the Ala130 carbonyl oxygen to interact strongly with the amino group of Lys129 (2.8 Å). The latter interaction is not possible for DBD1 Ala130, because of the dissimilar conformation of the KR boxes. **D-E.** The asymmetric interactions lead to differences in the distribution of the electrostatic potential in the region (amino acids 113-136) of ERR α DBD, including the D-box and KR-box of the extended region. In DBD2, **(D)** this region displays a highly positively charged electrostatic potential surface. In contrast in DBD1 **(E)**, the region exhibits a dispersed positively charged electrostatic potential surface.

E

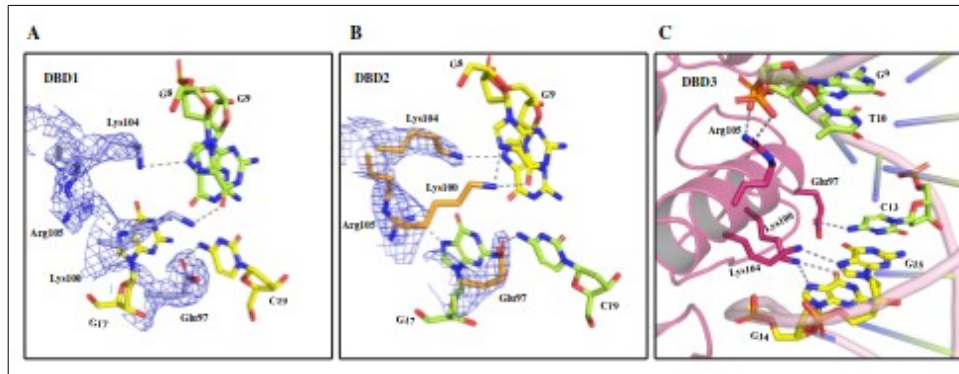


Suppl. Fig. 2E

F

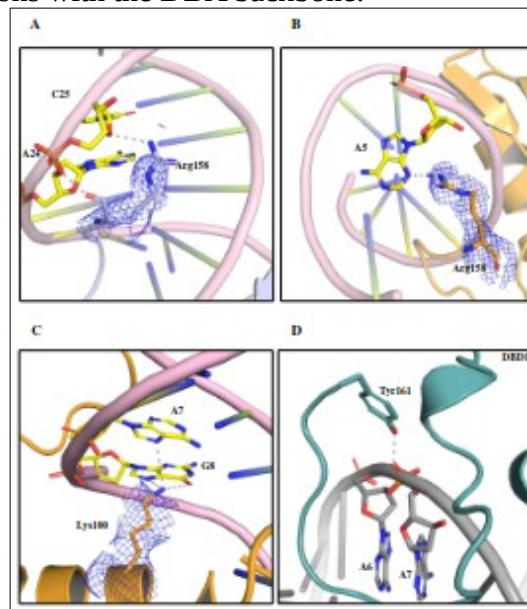


Suppl. Fig. 2F : Evolution of the sequence motifs corresponding to the KR-box in the NR3 family. source files and script will be provided in dryad with a doi.



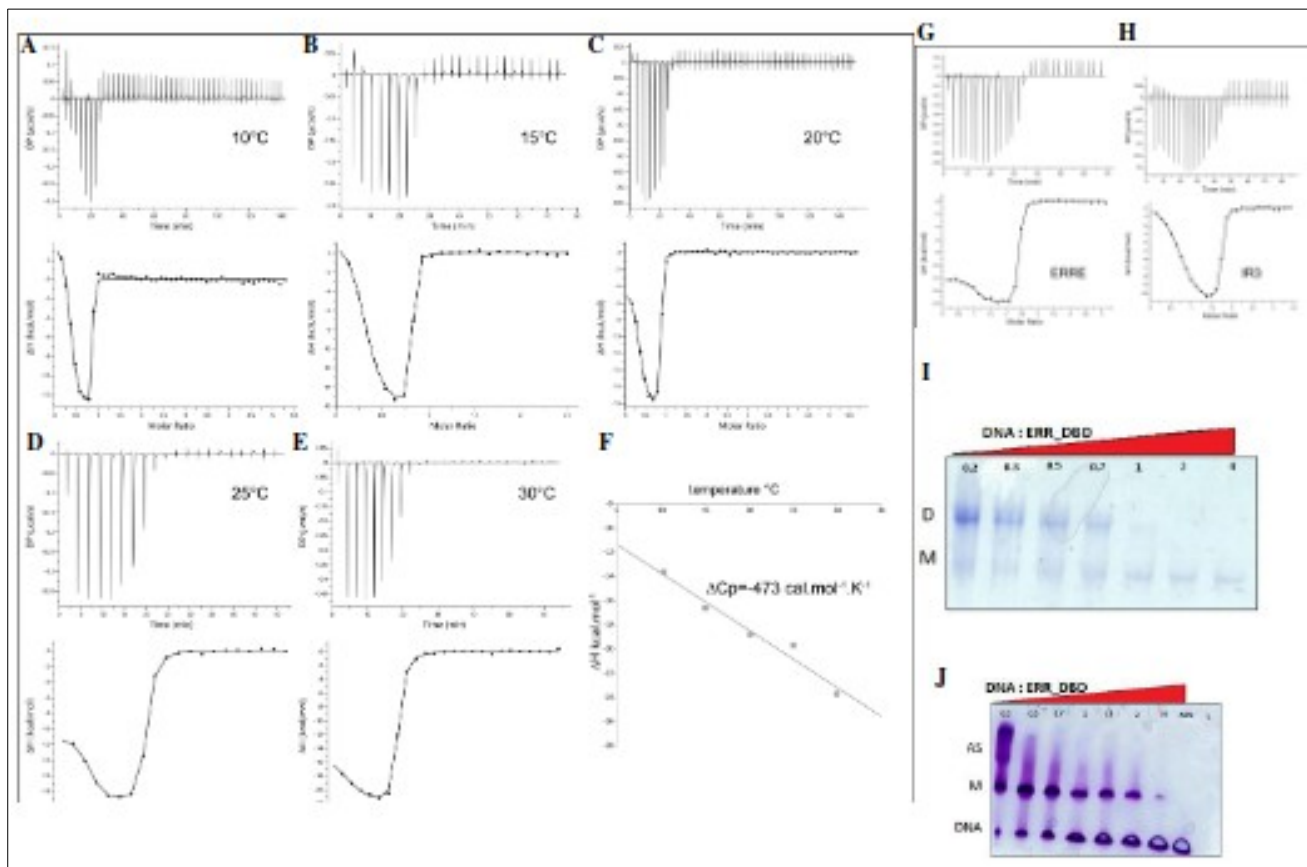
Suppl. Figure 3: Major groove interactions have not the same strength in DBD1 and DBD2.

A. Electron density in the region of the major groove of DBD1 contoured at 0.9 sigma. **B.** Electron density in the region of the major groove of DBD1 contoured at 0.9 sigma. **C.** DBD3 binds in an unspecific manner to a degenerated core half-site sequence CGGTGA (shown in **Suppl Fig. 1A**) and not to the consensus AGGTCA sequence. It offers a unique opportunity to get insights into base-pair specificity at the atomic level. No base-specific interactions with protein residues were observed for the first base pair (A/T) of the half-site. In contrast, the exchange of the penultimate G/C base pair by a C/G base pair leads to loss of binding specificity in the binding to has an important consequence for base-pair specificity loss in DBD3 (**Sup Fig. 3B**). Rather than binding to the N7 nitrogen atom of G17, it makes non-specific interactions with the DBA backbone.



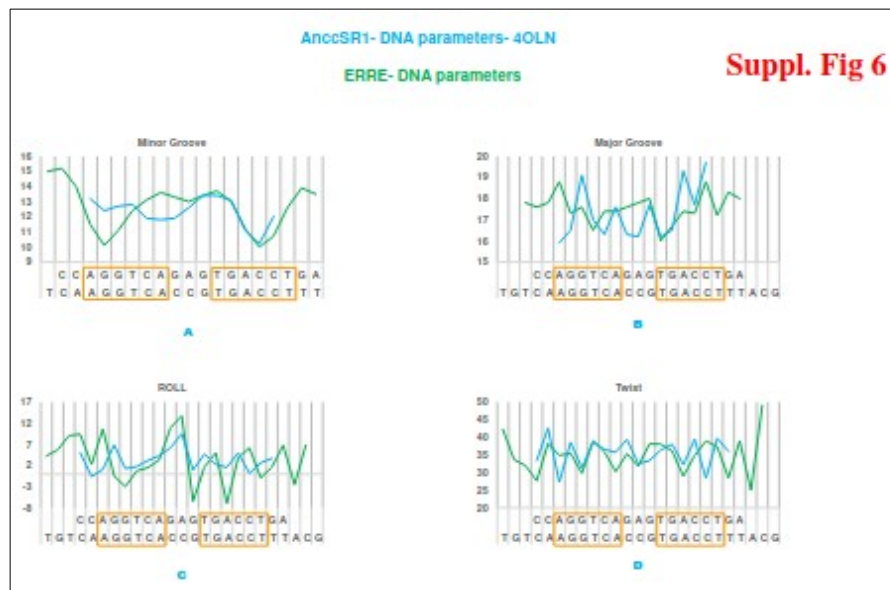
Suppl. Figure 4:

A-B. Polder map analysis of residue Arg158 in DBD1 (**A**) and (**B**) DBD2. **C.** Enlarged view of the interactions seen for Tyr161 with the DNA phosphate backbone for DBD1.



Suppl. Figure 5:

A-F. ITC data for ERR α DBD binding to embERRE DNA at 10 $^{\circ}\text{C}$, 15 $^{\circ}\text{C}$, 20 $^{\circ}\text{C}$, 25 $^{\circ}\text{C}$ and 30 $^{\circ}\text{C}$ (A-E). The protein is placed into the cell and the DNA is injected by the syringe. The baseline-subtracted thermogram is shown on the top and integrated heats on the bottom for each temperature. The stoichiometry observed for the specific binding event ($n=0.5$) is compatible with an ERR α DBD dimer binding to the DNA. The temperature-dependency of the binding enthalpy and the resulting heat capacity change is shown in F. **G-H.** ITC Data for ERR α DBD binding to traERRE half-site at 20 $^{\circ}\text{C}$ (G) and to IR3/ERE at 15 $^{\circ}\text{C}$ (H) DNA sequences. The protein is placed into the cell and the DNA is injected by the syringe. The baseline-subtracted thermogram is shown on the top and integrated heats are shown on the bottom. **I-J.** Native gels showing DBD dimerization according to the molar ratio of embERRE (I) or IR3/ERE (J).



Suppl. Figure 6: Superimposition of the DNA geometrical parameters of $ERR\alpha$ with those of AncSR1 (A-minor groove, B-major groove, C-Roll, and D-Twist).

ERRα-DBD/DNA	
Data collection	
Space group	P 2 ₁ 2 ₁ 2
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	168.12 70.82 94.85
β (°)	90.00 90.53 90.00
Resolution (Å)	94.85-2.48 (2.82-2.48)
<i>I</i> / <i>I</i>	8.9 (1.5)
Completeness (%) [†]	
Spherical	56.8 (9.1)
Ellipsoidal	93.4 (66.4)
Multiplicity	13.5 (14.0)
No. of unique reflections	23098 (1156)
CC _{1/2}	0.996 (0.648)
Wilson B-factor (Å ²)	78.8
Refinement	
Resolution (Å)	94.85 - 2.48
No. of reflections	23098
<i>R</i> _{work} / <i>R</i> _{free}	0.20/0.243
No. atoms	
Protein	8016
DNA	3284
Water molecules	55
<i>B</i> -factors (Å ²)	
Protein	98.05
DNA	115.41
Water molecules	23.20
R.m.s. deviations	
Bond lengths (Å)	0.0082
Bond angles (°)	0.89
Ramachandran favoured (%)	94.43
Ramachandran allowed (%)	5.57
Ramachandran outliers (%)	0.0
PDB ID	Under deposition

Suppl. Table 1. Data collection and refinement statistics

[†]Data were elliptically truncated by Staraniso software from Autoproc suite.

*Values in parentheses are for highest-resolution shell.

DNA	Temperature (°C)	KD (nM)	ΔH (kcal/mol)	ΔG (kcal/mol)	-T ΔS (kcal/mol)
embERRE	10.4	3.63	-13.6	-10.9	2.61
embERRE	15.1	8.92	-16.6	-10.6	6.00
embERRE	20.2	12.9	-18.8	-10.6	7.88
embERRE	25.1	22.2	-19.7	-10.5	9.27
embERRE	30.0	25.1	-23.7	-10.4	13.2
traERRE	20.2	26.8	-17.5	-10.2	7.30
IR3	15.2	28.2	-11.7	-9.96	1.71

Suppl. Table 2. ITC Data for ERR DBD specific binding to various DNA sequences.

References

- Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D* 66, 213-221. doi:10.1107/S0907444909052925.
- Afonine, P.V., Moriarty, N.W., Mustyakimov, M., Sobolev, O.V., Terwilliger, T.C., Turk, D., Urzhumtsev, A., and Adams, P.D. (2015). FEM: feature-enhanced map. *Acta Crystallogr D Biol Crystallogr* 71, 646-666. 10.1107/S1399004714028132.
- Anisimova, M., and Gascuel, O. (2006). Approximate Likelihood-Ratio Test for Branches: A Fast, Accurate, and Powerful Alternative. *Systematic Biology* 55, 539-552. 10.1080/10635150600755453.
- Audet-Walsh, E., and Giguere, V. (2015). The multiple universes of estrogen-related receptor [alpha] and [gamma] in metabolic control and related diseases. *Acta Pharmacol Sin* 36, 51-61. Review.
- Baker, M.E. (2008). Trichoplax, the simplest known animal, contains an estrogen-related receptor but no estrogen receptor: Implications for estrogen receptor evolution. *Biochemical and Biophysical Research Communications* 375, 623-627. <https://doi.org/10.1016/j.bbrc.2008.08.047>.
- Bannister, R., Beresford, N., Granger, D.W., Pounds, N.A., Rand-Weaver, M., White, R., Jobling, S., and Routledge, E.J. (2013). No substantial changes in estrogen receptor and estrogen-related receptor orthologue gene transcription in *Marisa cornuarietis* exposed to estrogenic chemicals. *Aquatic Toxicology* 140-141, 19-26. <https://doi.org/10.1016/j.aquatox.2013.05.002>.
- Bardet, P.L., Laudet, V., and Vanacker, J.M. (2006). Studying non-mammalian models? Not a fool's ERRand! *Trends Endocrinol Metab* 17, 166-171.
- Barry, J.B., and Giguere, V. (2005). Epidermal Growth Factor-Induced Signaling in Breast Cancer Cells Results in Selective Target Gene Activation by Orphan Nuclear Receptor Estrogen-Related Receptor {alpha}. *Cancer Research* 65, 6120-6129.
- Barry, J.B., Laganiere, J., and Giguere, V. (2006). A Single Nucleotide in an Estrogen-Related Receptor {alpha} Site Can Dictate Mode of Binding and Peroxisome Proliferator-Activated

Receptor $\{\gamma\}$ Coactivator 1 $\{\alpha\}$ Activation of Target Promoters. *Molecular Endocrinology* 20, 302-310.

- Bertrand, S., Belgacem, M.R., and Escriva, H. (2010). Nuclear hormone receptors in chordates. *Molecular and Cellular Endocrinology* In Press, Corrected Proof.
- Billas, I.M.L., and Moras, D. (2013). Allosteric Controls of Nuclear Receptor Function in the Regulation of Transcription. *Journal of Molecular Biology* 425, 2317-2329.
- Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O.S., Vornrhein, C., and Womack, T.O. (2017). BUSTER version 2.10.3. Cambridge, United Kingdom: Global Phasing Ltd.
- Deblois, G., Hall, J.A., Perry, M.C., Laganier, J., Ghahremani, M., Park, M., Hallett, M., and Giguère, V. (2009). Genome-Wide Identification of Direct Target Genes Implicates Estrogen-Related Receptor α as a Determinant of Breast Cancer Heterogeneity. *Cancer Research* 69, 6149-6157.
- Delano, W.L. (2002). Pymol: An open-source molecular graphics tool. *CCP4 Newsletter On Protein Crystallography* 40, 82-92.
- Denayer, S., Helsen, C., Thorrez, L., Haelens, A., and Claessens, F. (2010). The Rules of DNA Recognition by the Androgen Receptor. *Molecular Endocrinology* 24, 898-913.
- Dufour, C.R., Wilson, B.J., Huss, J.M., Kelly, D.P., Alaynick, W.A., Downes, M., Evans, R.M., Blanchette, M., and Giguère, V. (2007). Genome-wide Orchestration of Cardiac Functions by the Orphan Nuclear Receptors ERR $\{\alpha\}$ and $\{\gamma\}$. *Cell Metabolism* 5, 345-356.
- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallographica Section D* 66, 486-501. doi:10.1107/S0907444910007493.
- Gearhart, M.D., Holmbeck, S.M., Evans, R.M., Dyson, H.J., and Wright, P.E. (2003). Monomeric complex of human orphan estrogen related receptor-2 with DNA: a pseudo-dimer interface mediates extended half-site recognition. *J Mol Biol* 327, 819-832.
- Giguère, V., Yang, N., Segui, P., and Evans, R.M. (1988). Identification of a new class of steroid hormone receptors. *Nature* 331, 91-94.
- Gomes, I.D.L., Gazo, I., Nabi, D., Besnardeau, L., Hebras, C., McDougall, A., and Dumollard, R. (2019). Bisphenols disrupt differentiation of the pigmented cells during larval brain formation in the ascidian. *Aquatic Toxicology* 216, 105314. <https://doi.org/10.1016/j.aquatox.2019.105314>.
- Gouy, M., Guindon, S., and Gascuel, O. (2009). SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. *Molecular Biology and Evolution* 27, 221-224. 10.1093/molbev/msp259.
- Greschik, H., Althage, M., Flaig, R., Sato, Y., Chavant, V., Peluso-Iltis, C., Choulier, L., Cronet, P., Rochel, N., Schule, R., et al. (2008). Communication between the ERR α homodimer interface and the PGC-1 α binding surface via the helix 8-9 loop. *J Biol Chem* 283, 20220-20230.
- Gronemeyer, H., Gustafsson, J.A., and Laudet, V. (2004). Principles for modulation of the nuclear receptor superfamily. *Nat Rev Drug Discov* 3, 950-964.
- Guindon, S., and Gascuel, O. (2003). A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. *Systematic Biology* 52, 696-704. 10.1080/10635150390235520.
- Helsen, C., and Claessens, F. (2014). Looking at nuclear receptors from a new angle. *Molecular and Cellular Endocrinology* 382, 97-106.

- Helsen, C., Kerkhofs, S., Clinckemalie, L., Spans, L., Laurent, M., Boonen, S., Vanderschueren, D., and Claessens, F. (2011). Structural basis for nuclear hormone receptor DNA binding. *Molecular and Cellular Endocrinology* 348, 411-417. doi: 10.1016/j.mce.2011.07.025.
- Hochberg, G.K.A., Liu, Y., Marklund, E.G., Metzger, B.P.H., Laganowsky, A., and Thornton, J.W. (2020). A hydrophobic ratchet entrenches molecular complexes. *Nature*. 10.1038/s41586-020-3021-2.
- Horard, B., Castet, A., Bardet, P.L., Laudet, V., Cavailles, V., and Vanacker, J.M. (2004). Dimerization is required for transactivation by estrogen-receptor-related (ERR) orphan receptors: evidence from amphioxus ERR. *Journal of Molecular Endocrinology* 33, 493-509.
- Hudson, W.H., Youn, C., and Ortlund, E.A. (2014). Crystal Structure of the Mineralocorticoid Receptor DNA Binding Domain in Complex with DNA. *PLoS ONE* 9, e107000. doi:10.1371/journal.pone.0107000.
- Huelsenbeck, J.P., Nielsen, R., and Bollback, J.P. (2003). Stochastic Mapping of Morphological Characters. *Systematic Biology* 52, 131-158. 10.1080/10635150390192780.
- Huppunen, J., and Aarnisalo, P. (2004). Dimerization modulates the activity of the orphan nuclear receptor ERR γ . *Biochemical and Biophysical Research Communications* 314, 964-970. <http://dx.doi.org/10.1016/j.bbrc.2003.12.194>.
- Kallen, J., Lattmann, R., Beerli, R., Blechschmidt, A., Blommers, M.J.J., Geiser, M., Ottl, J., Schlaeppli, J.M., Strauss, A., and Fournier, B. (2007). Crystal Structure of Human Estrogen-related Receptor $\{\alpha\}$ in Complex with a Synthetic Inverse Agonist Reveals Its Novel Molecular Mechanism. *Journal of Biological Chemistry* 282, 23231-23239.
- Khalturin, K., Billas, I.M.L., Chebaro, Y., Reitzel, A.M., Tarrant, A.M., Laudet, V., and Markov, G.V. (2018). NR3E receptors in cnidarians: A new family of steroid receptor relatives extends the possible mechanisms for ligand binding. *The Journal of Steroid Biochemistry and Molecular Biology* 184, 11-19. <https://doi.org/10.1016/j.jsbmb.2018.06.014>.
- Lazar, M.A. (2017). Maturing of the nuclear receptor family. *The Journal of Clinical Investigation* 127, 1123-1125. 10.1172/JCI92949.
- Le, S.Q., and Gascuel, O. (2010). Accounting for Solvent Accessibility and Secondary Structure in Protein Phylogenetics Is Clearly Beneficial. *Systematic Biology* 59, 277-287. 10.1093/sysbio/syq002.
- Liebschner, D., Afonine, P.V., Moriarty, N.W., Poon, B.K., Sobolev, O.V., Terwilliger, T.C., and Adams, P.D. (2017). Polder maps: improving OMIT maps by excluding bulk solvent. *Acta Crystallographica Section D* 73, 148-157. doi:10.1107/S2059798316018210.
- Little, T.H., Zhang, Y., Matulis, C.K., Weck, J., Zhang, Z., Ramachandran, A., Mayo, K.E., and Radhakrishnan, I. (2006). Sequence-Specific Deoxyribonucleic Acid (DNA) Recognition by Steroidogenic Factor 1: A Helix at the Carboxy Terminus of the DNA Binding Domain Is Necessary for Complex Stability. *Molecular Endocrinology* 20, 831-843. 10.1210/me.2005-0384.
- Lu, X.J., and Olson, W.K. (2003). 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Research* 31, 5108-5121. 10.1093/nar/gkg680.
- Luisi, B.F., Xu, W.X., Otwinowski, Z., Freedman, L.P., Yamamoto, K.R., and Sigler, P.B. (1991). Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. *Nature* 352, 497-505.
- McCoy, A. (2007). Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallographica Section D* 63, 32-41. doi:10.1107/S0907444906045975.

- McKeown, Alesia N., Bridgham, Jamie T., Anderson, Dave W., Murphy, Michael N., Ortlund, Eric A., and Thornton, Joseph W. (2014). Evolution of DNA Specificity in a Transcription Factor Family Produced a New Gene Regulatory Module. *Cell* 159, 58-68. <https://doi.org/10.1016/j.cell.2014.09.003>.
- Meijnsing, S.H., Pufall, M.A., So, A.Y., Bates, D.L., Chen, L., and Yamamoto, K.R. (2009). DNA Binding Site Sequence Directs Glucocorticoid Receptor Structure and Activity. *Science* 324, 407-410.
- Mohideen-Abdul, K., Tazibt, K., Bourguet, M., Hazemann, I., Lebars, I., Takacs, M., Cianféroni, S., Klaholz, B.P., Moras, D., and Billas, I.M.L. (2017). Importance of the Sequence-Directed DNA Shape for Specific Binding Site Recognition by the Estrogen-Related Receptor. *Frontiers in Endocrinology* 8. 10.3389/fendo.2017.00140.
- Oda, M., Furukawa, K., Ogata, K., Sarai, A., and Nakamura, H. (1998). Thermodynamics of specific and non-specific DNA binding by the c-myc DNA-binding domain. Edited by P. E. Wright. *Journal of Molecular Biology* 276, 571-590. <https://doi.org/10.1006/jmbi.1997.1564>.
- Ortlund, E.A., Bridgham, J.T., Redinbo, M.R., and Thornton, J.W. (2007). Crystal Structure of an Ancient Protein: Evolution by Conformational Epistasis. *Science* 317, 1544-1548.
- Peters, W.B., Edmondson, S.P., and Shriver, J.W. (2004). Thermodynamics of DNA Binding and Distortion by the Hyperthermophile Chromatin Protein Sac7d. *Journal of Molecular Biology* 343, 339-360. <https://doi.org/10.1016/j.jmb.2004.08.042>.
- Piñeiro, Á., Muñoz, E., Sabín, J., Costas, M., Bastos, M., Velázquez-Campoy, A., Garrido, P.F., Dumas, P., Ennifar, E., García-Río, L., et al. (2019). AFFINImeter: A software to analyze molecular recognition processes from experimental data. *Analytical Biochemistry* 577, 117-134. <https://doi.org/10.1016/j.ab.2019.02.031>.
- Privalov, P.L., Dragan, A.I., and Crane-Robinson, C. (2010). Interpreting protein/DNA interactions: distinguishing specific from non-specific and electrostatic from non-electrostatic components. *Nucleic Acids Research* 39, 2483-2491. 10.1093/nar/gkq984.
- R-Core-Team (2021). R: A language and environment for statistical computing; <https://www.R-project.org/>. R Foundation for Statistical Computing.
- Revell, L.J. (2012). phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* 3, 217-223. <https://doi.org/10.1111/j.2041-210X.2011.00169.x>.
- Roemer, S.C., Donham, D.C., Sherman, L., Pon, V.H., Edwards, D.P., and Churchill, M.E.A. (2006). Structure of the Progesterone Receptor-Deoxyribonucleic Acid Complex: Novel Interactions Required for Binding to Half-Site Response Elements. *Molecular Endocrinology* 20, 3042-3052.
- Rohs, R., West, S.M., Sosinsky, A., Liu, P., Mann, R.S., and Honig, B. (2009). The role of DNA shape in protein-DNA recognition. *Nature* 461, 1248-1253. http://www.nature.com/nature/journal/v461/n7268/supinfo/nature08473_S1.html.
- Schwabe, J.W., Chapman, L., Finch, J.T., Rhodes, D., and Neuhaus, D. (1993a). DNA recognition by the oestrogen receptor: from solution to the crystal. *Structure* 1, 187-204.
- Schwabe, J.W.R., Chapman, L., Finch, J.T., and Rhodes, D. (1993b). The crystal structure of the estrogen receptor DNA-binding domain bound to DNA: How receptors discriminate between their response elements. *Cell* 75, 567-578.
- Shaffer, P.L., Jivan, A., Dollins, D.E., Claessens, F., and Gewirth, D.T. (2004). Structural basis of androgen receptor binding to selective androgen response elements. *Proc. Natl Acad. Sci. U. S. A* 101, 4758-4763.

- Sievers, F., and Higgins, D.G. (2014). Clustal Omega, Accurate Alignment of Very Large Numbers of Sequences. In *Multiple Sequence Alignment Methods*, D.J. Russell, ed. (Humana Press), pp. 105-116. 10.1007/978-1-62703-646-7_6.
- Sladek, R., Bader, J.A., and Giguère, V. (1997). The orphan nuclear receptor estrogen-related receptor alpha is a transcriptional regulator of the human medium-chain acyl coenzyme A dehydrogenase gene. *Mol Cell Biol* 17, 5400-5409.
- Solomon, I.H., Hager, J.M., Safi, R., McDonnell, D.P., Redinbo, M.R., and Ortlund, E.A. (2005). Crystal Structure of the Human LRH-1 DBD-DNA Complex Reveals Ftz-F1 Domain Positioning is Required for Receptor Activity. *Journal of Molecular Biology* 354, 1091-1102.
- Tennesen, J.M., Baker, K.D., Lam, G., Evans, J., and Thummel, C.S. (2011). The Drosophila Estrogen-Related Receptor Directs a Metabolic Switch that Supports Developmental Growth. *Cell Metabolism* 13, 139-148. <https://doi.org/10.1016/j.cmet.2011.01.005>.
- Tennesen, Jason M., and Thummel, Carl S. (2011). Coordinating Growth and Maturation — Insights from Drosophila. *Current Biology* 21, R750-R757. <https://doi.org/10.1016/j.cub.2011.06.033>.
- Thornton, J.W., Need, E., and Crews, D. (2003). Resurrecting the ancestral steroid receptor: ancient origin of estrogen signaling. *Science* 301, 1714-1717.
- Tremblay, A.M., Dufour, C.R., Ghahremani, M., Reudelhuber, T.L., and Giguere, V. (2010). Physiological genomics identifies estrogen-related receptor alpha as a regulator of renal sodium and potassium homeostasis and the renin-angiotensin pathway. *Mol. Endocrinol* 24, 22-32.
- Tremblay, A.M., and Giguère, V. (2007). The NR3B subgroup: an ovERRview. *Nucl. Recept. Signal* 5, e009.
- Vanacker, J.-M., Bonnelye, E., Chopin-Delannoy, S., Delmarre, C., Cavallès, V., and Laudet, V. (1999). Transcriptional Activities of the Orphan Nuclear Receptor ERR α (Estrogen Receptor-Related Receptor- α). *Molecular Endocrinology* 13, 764-773. doi:10.1210/mend.13.5.0281.
- Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T., and Bricogne, G. (2011). Data processing and analysis with the autoPROC toolbox. *Acta Crystallographica Section D* 67, 293-302. doi:10.1107/S0907444911007773.
- Xia, H., Dufour, C.R., and Giguère, V. (2019). ERR α as a Bridge Between Transcription and Function: Role in Liver Metabolism and Disease. *Frontiers in endocrinology* 10, 206. 10.3389/fendo.2019.00206.
- Zhang, Z., and Teng, C.T. (2000). Estrogen Receptor-related Receptor +1 Interacts with Coactivator and Constitutively Activates the Estrogen Response Elements of the Human Lactoferrin Gene. *Journal of Biological Chemistry* 275, 20837-20846.
- Zhao, L., Zhou, S., and Gustafsson, J.-Å. (2019). Nuclear Receptors: Recent Drug Discovery for Cancer Therapies. *Endocrine Reviews* 40, 1207-1249. 10.1210/er.2018-00222.
- Zhou, T., Yang, L., Lu, Y., Dror, I., Dantas Machado, A.C., Ghane, T., Di Felice, R., and Rohs, R. (2013). DNASHape: a method for the high-throughput prediction of DNA structural features on a genomic scale. *Nucleic Acids Research* 41, W56-W62. 10.1093/nar/gkt437.

Partie II – PGC-1 α régule l'assise de ERR α sur l'ADN

Introduction

Nous venons de voir le rôle des paramètres de forme de l'ADN sur la dimérisation des DBDs de ERR α . Il convient à présent de déterminer ce qu'il en est du récepteur nucléaire entier, ainsi que des mécanismes régulant son activité. Nous avons donc décidé de nous focaliser sur le récepteur entier en présence de son coactivateur PGC-1 α , tout en conservant l'élément de réponse composite. Le but de cette étude sera de déterminer la topologie du complexe et de caractériser l'impact de PGC-1 α sur la liaison entre le récepteur et son ADN.

De précédents jeux de données acquis au sein de notre laboratoire par cryo-EM ont permis d'obtenir une première reconstruction *ab initio* du complexe ERR-PGC sur un élément de réponse composite (**Fig. 34**). Cette carte permet de replacer le récepteur nucléaire et l'ADN par comparaison avec la structure cristallographique d'un autre récepteur nucléaire homodimérique sur son élément de réponse DR1 : HNF-4 (Chandra *et al.*, 2013). La région qui peut correspondre à PGC-1 α est attribuée à l'extension sortant de la région assignée aux LBDs. Sachant que le fragment du coactivateur utilisé faisait 55 kDa et 472 résidus d'acides aminés, il est surprenant que cette région soit plus petite que les LBDs. Cela implique qu'une partie de la construction de PGC-1 α utilisée est flexible et/ou adopte plusieurs conformations distinctes. La stœchiométrie du complexe est ici de un PGC-1 α pour deux ERR α , traduisant la présence d'un effet allostérique lors de l'interaction empêchant la liaison sur le second site potentiel de fixation.

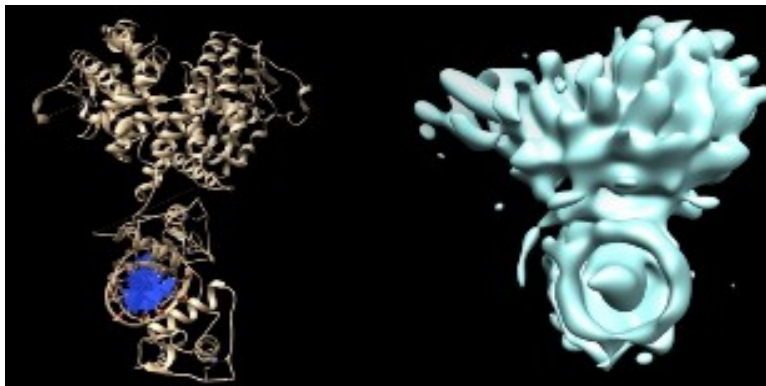


Figure 34: Analyse d'une première carte du complexe ERR α /PGC-1 α /ADN par comparaison avec les structures existantes

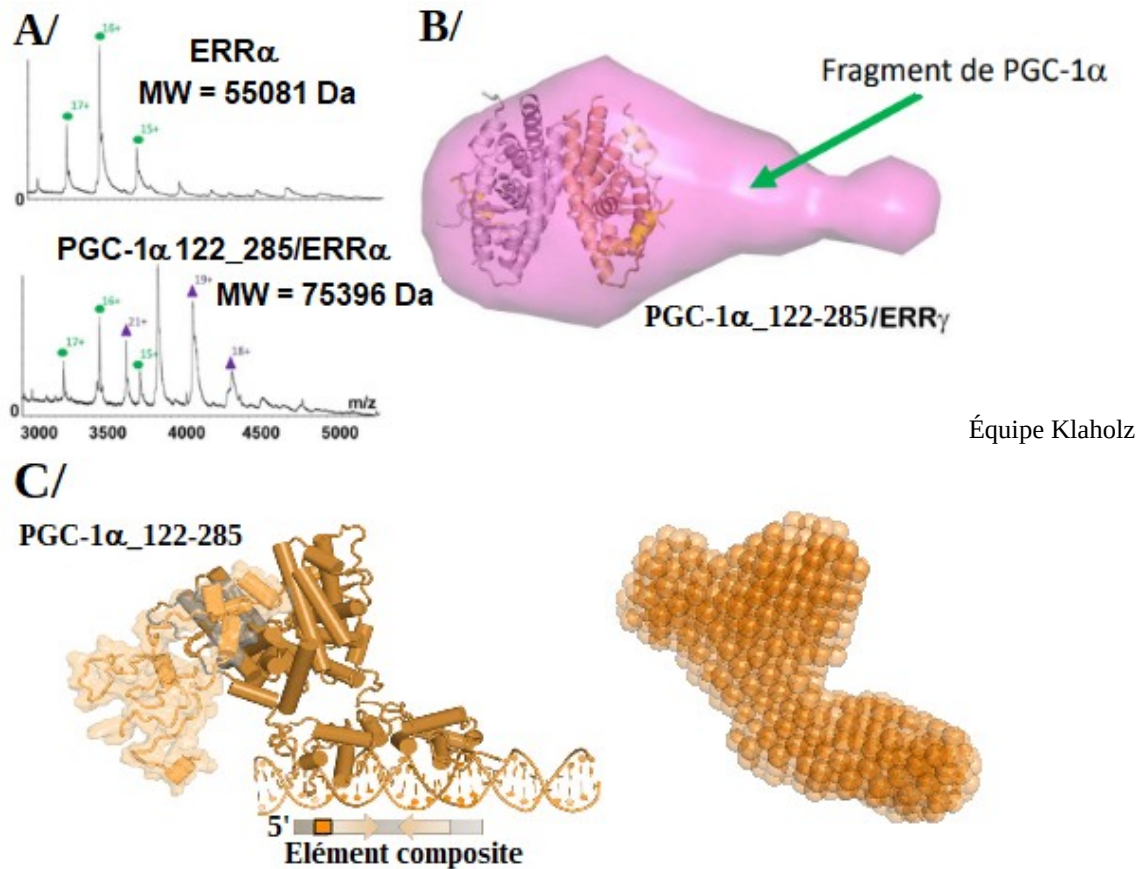
(selon la thèse « Origine et évolution des récepteurs nucléaires et étude structurale du premier stéroïdien, ERR » de Brice Beinsteiner, (2018))

Légende :

Structure du récepteur HNF-4 sur son élément de réponse (à gauche) et reconstruction du complexe ERR α -PGC-1 α -wc29embedded (séquence dans le **tableau 4**) réalisée par B. Beinsteiner (2018).

Ce rapport stœchiométrique et l'hypothèse d'un effet allostérique sont cohérents avec les résultats acquis au sein de notre équipe par plusieurs expériences de spectrométries de masses natives (**Fig. 35-A**), de RMN, de chromatographie d'exclusion stérique (SEC) couplée à une diffusion de lumière multi-

angles (SEC-MALS) et de SAXS en utilisant les LBDs de $ERR\alpha$ et $ERR\gamma$ en présence de plusieurs constructions de $PGC-1\alpha$ (**Fig. 35-B et C**).



Équipe Klaholz

Figure 35: Mise en évidence de l'effet allostérique induit par la fixation de $PGC-1\alpha$

Légende :

A/ Résultats de spectrométrie de masse native attestant de la présence d'un seul $PGC-1\alpha$ en présence du dimère de LBDs de $ERR\alpha$ d'après la masse molaire des LBDs de $ERR\alpha$.

B/ Enveloppe SAXS des LBD de $ERR\gamma$ avec un fragment de $PGC-1\alpha$ dans laquelle la structure cristallographique du LBD sous forme dimérique (3D24) a été positionnée.

C/ Modèle produit à partir de l'enveloppe SAXS de $ERR\gamma \Delta A/B$ en présence de $PGC-1\alpha_{122-285}$ sur un élément de réponse composite (à gauche) et une des enveloppes SAXS du complexe à une résolution estimée de 25-30 Å (à droite).

Nous compléterons donc ici l'analyse biophysique de l'interaction entre $ERR\alpha$ et $PGC-1\alpha$ commencée par mes prédécesseurs et mesurerons l'impact de cette interaction sur la fixation à l'ADN tout en décrivant les étapes de l'étude bio-structurale autour de ce complexe. Cette étude aurait aussi pu se focaliser sur $ERR\gamma$, notamment pour l'analyse structurale par cryo-EM, mais l'intérêt biologique de $ERR\alpha$ étant déterminant et la bibliographie le concernant plus riche, nous avons choisi de nous focaliser sur celui-ci.

Matériels & Méthodes

1/ Choix des constructions

Les constructions choisies pour obtenir la structure du récepteur nucléaire $ERR\alpha$ en complexe avec son coactivateur PGC-1 α et son élément de réponse sont essentiellement motivées par les contraintes des techniques employées pour la détermination structurale, les systèmes d'expression de chaque protéine et l'expérience du laboratoire dans l'étude de ce type de complexe.

Dans le cadre de l'étude par cristallographie, les constructions PGC-1 α _74-219 et PGC-1 α _138-219 sont conçues. PGC-1 α étant une protéine flexible dont la taille pourrait nuire à l'empilement cristallin, mais étant aussi difficile à produire lorsque trop longue au sein d'*Escherichia coli* (*E. coli*), ces constructions ne se composent que d'une partie de la région N-terminale de PGC-1 α comportant le motif d'interaction atypique avec les LBDs de $ERR\alpha$, ainsi que d'une protéine de fusion de 8 kDa liant le calcium (Fh8) en position N-terminale pour faciliter leurs solubilités lors de l'expression.

Dans le cadre de l'étude par cryo-EM, c'est la détection du complexe sur les grilles qui risque de poser problème, ce dernier demeurant un petit objet à faible contraste. Ainsi, afin d'augmenter sa taille et donc sa détectabilité, la construction PGC-1 α _1-285 a été choisie. Les tentatives d'exprimer une construction plus longue et soluble de PGC-1 α ont pour l'instant échoué en *E. coli*.

Pour l'ensemble de ces constructions de PGC-1 α produites en bactéries, une construction du $ERR\alpha$ humain entier, human_ $ERR\alpha$ _1-423, a également été exprimée en bactérie indépendamment. De plus, le laboratoire ayant déjà procédé à des tentatives avec un complexe coproduit en cellules d'insectes infectées par baculovirus, les constructions mouse_ $ERR\alpha$ _1-422 et PGC-1 α 1-472 ont également été utilisées. Ces dernières présentent l'avantage potentiel de disposer de modifications post-traductionnelles, mais aussi de permettre la formation du complexe *in cellulo* ce qui pourrait le rendre plus stable. Ce système d'expression nous permet aussi d'utiliser des constructions plus grandes de PGC-1 α lesquelles sont PGC-1 α _1-635 et PGC-1 α _1-798. La coexpression des fragments de PGC-1 α avec $ERR\alpha$ augmente considérablement leur expression par comparaison avec celle des fragments seuls.

Enfin, dans l'optique de produire un complexe plus important par la taille, mais aussi plus stable, nous avons procédé à la production d'une construction de la protéine p300 et de la protéine PCAF, lesquelles sont connues pour leur implication dans la formation de complexes d'activation de la transcription et pour leur rôle dans la régulation des complexes impliquant ERR et PGC. Cette approche a été confirmée par les articles de l'équipe de Bert O'Malley qui dévoile l'existence de complexes impliquant ER et p300 (Yi *et al.*, 2015 et 2017).

L'ensemble des constructions utilisées, leurs systèmes d'expressions, leurs vecteurs et leurs étiquettes (tag) sont récapitulés dans les Tableaux ci-dessous (**Tableaux 2 et 3**).

Noms	Acides Aminés	Étiquette/protéine de fusion	Position de l'étiquette	Site de clivage	Souches	[IPTG]	Milieu de Culture	Température D'induction
PGC-1 α _74-219	74-219	His10-Strep-Strep-Fh8	N-terminal	P3C	pRARE2	100 mM	LB	18 °C
PGC-1 α _138-219	138-219	His10-Strep-Strep-Fh8	N-terminal	P3C	pRIL	100 mM	LB	18 °C
PGC-1 α _1-285	1-285	His6	N-terminal	Thrombine	BL21(DE3)	500 mM	TB	18 °C
human_ERR α _1-423	1-423	His6-Trx	N-terminal	TEV	pRARE2	100 mM	LB	18 °C

Tableau 2: Constructions destinées à la production en E. coli

Noms	Acides Aminés	Étiquette	Position de l'étiquette	Souches
PGC-1 α _1-472	1-472	Strep	N&C-terminal	Sf9
PGC-1 α _1-635	1-635	Strep	C-terminal	Sf9
PGC-1 α _1-798	1-798	Strep	C-terminal	Sf9
mouse_ERR α _1-422	2-422	His6	C-terminal	Sf9
p300	2-2414	His6	C-terminal	Sf21
PCAF	66-832	FLAG	N-terminal	Sf21

Tableau 3: Constructions destinées à la production en cellules d'insecte

2/ Hybridation des ADN

Les ADN utilisés pour la formation des complexes sont commandés chez Eurogentec sous la forme de simple brin lyophilisés. Une fois centrifugés à 10 000 rpm durant 10 minutes et resuspendus dans une solution composée de 10 mM tris(hydroxyméthyl)aminométhane (Tris) HCl pH 7.5, 100 mM NaCl, 1 % DMSO et 0.2 mM EDTA, les brins sens et antisens de chaque ADN sont placés dans les mêmes proportions molaires jusqu'à un volume final de 100 μ L, puis l'hybridation des ADN est réalisée en chauffant les tubes 2 minutes à 95°C, puis en les refroidissant à 25°C graduellement durant 45 minutes en utilisant une machine Thermal Cycler (Biorad). Les ADN sont ensuite conservés à 4 °C ou -20 °C après vérification de leur état double brin par migration sur gel de polyacrylamide 25 % pendant 30 minutes à 2W constant et coloration au bleu de toluidine.

L'ensemble des ADN hybridés par cette méthode se trouvent dans le tableau récapitulatif suivant :

Noms	Sequence
TrαERRE	ATG TCAAGGTCA CACAGTGAGTTAC ACAG TCCAGT GTGTCACCTCAATGT
wc26embedded	ATG TCAAGGTCA CCGT GACCT TTACG ACAG TCCAGT GGC ACTGGAA ATGCT
wc29embedded	AATTG TCAAGGTCA CCGT GACCT TTACCG AAC AGTTCAGT GGC ACTGGAA ATGGCTT
BE42embedded	AGTCCCATG TCAAGGTCA CCGT GACCT TTACGAATTTAATAA TCAGGGTAC AGTTCAGT GGC ACTGGAA ATGCTTAAATTATT
BE42embedded-bio	Biotine – AGTCCCATG TCAAGGTCA CCGT GACCT TTACGAATTTAATAA TCAGGGTAC AGTTCAGT GGC ACTGGAA ATGCTTAAATTATT

Tableaux 4: ADN utilisés

3/ Expression des protéines

3.1/ Infection et expression des protéines produites en baculovirus

Les baculovirus porteurs des gènes codant pour les constructions mouse_ERR α _1-422 et PGC1- α _1-472 sont incubés en présence de cellules sf9 ou sf21 Oxford durant 72h. Les cultures en plaques sont centrifugées puis resuspendues en tampon de lyse (20 mM Tris HCl pH 8.0, 200 mM NaCl, 4 mM CHAPS, 1 mM TCEP, 0.2 mM MgCl₂). Les cellules sont lysées par sonication, puis les lysats sont centrifugés 1 h à 11 000 *g* et à 4 °C. La purification est réalisée à l'aide d'une résine Strep tactin Sepharose préalablement rincées et équilibrées en tampon de lyse. Après 45 minutes d'incubation, la résine est lavée dans le tampon à trois reprises, puis elle est mise en présence de Laemmli et chauffée à 95°C pendant 2 minutes. Elle est centrifugée et le surnageant est déposé sur gel (**Fig. S1**).

Les constructions PGC-1 α _1-635 et PGC-1 α _1-798 sont transformées au sein de bacmides, extraites par Miniprep (kit NucleoSpin Plasmid de Machery-Nagel), amplifiées par PCR, puis insérées au sein de baculovirus. Ceux-ci sont employés pour l'infection de cellules sf21 sur 48 à 72h (**Fig. S2**).

Le vecteur contenant le gène codant pour la protéine PCAF est inséré au sein des Bacmid YFP Multibac et CFP Multibac à l'aide d'un choc thermique de 1 minute à 42°C puis les bactéries sont étalées sur boîte et incubées sur la nuit à 37°C. Les colonies sont ensuite repiquées afin d'isoler les bactéries ayant été transformées efficacement. L'ADN codant pour PCAF est ensuite extrait par miniprep (kit NucleoSpin Plasmid de Machery-Nagel), amplifié par PCR et analysé sur gel d'agarose. L'ADN provenant du bacmid YFP Multibac est donné à la plate-forme de baculovirus du CBI qui teste les différents temps d'incubation et différentes quantités de virus. La lyse est réalisée par sonication et la purification de la construction s'effectue par l'utilisation d'une incubation des lysats sur des billes couvertes d'anticorps anti-FLAG. L'incubation en présence du virus est arrêtée à 48h pour une quantité de 30 virions par cellule (**Fig. S3**).

Différentes quantités de virus sont testées dans le cadre de la production de la protéine p300 dont le temps d'infection est de deux jours. Les cultures sont déposées en tubes Eppendorf, centrifugés 20 minutes à 1000 rpm, puis les culots sont resuspendus en tampon de lyse (10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.5, 1 mM MgCl₂, 400 mM NaCl, 2 mM TCEP, 2 mM CHAPS) avant de subir deux minutes de sonication à l'aide d'un sonicateur Labsonic P (Sartorius) à 30 % d'amplitude pour une valeur de cycle de 0.5. Les lysats sont ensuite centrifugés 10 minutes à 11 000 *g* et le surnageant est déposé sur des billes de nickel Ni-NTA agarose destinées à la réalisation de chromatographie d'affinité aux ions métalliques (IMAC). Ces billes sont préalablement lavées dans l'eau à trois reprises et dans le tampon de lyse à trois reprises. L'incubation dure un total de 45 minutes, puis les billes sont lavées à trois reprises dans le tampon de lyse. Elles sont ensuite mises en présence du tampon de lyse supplémenté de 500 mM d'imidazole. Les éluats sont ensuite déposés sur gel après ajout de bleu de Laemmli et cinq minutes de choc thermique à 95°C (**Fig. S4**).

Toutes les constructions produites ont été contrôlées par spectrométrie de masse par la plate-forme de spectrométrie de masse de l'IBMC en découpant directement les bandes concernées sur gel afin d'identifier les protéines.

3.2/ Transformation et expression des protéines produites en bactéries

Les bactéries *E. coli* sont transformées en déposant 50 ng du vecteur contenant la construction choisie sur le culot bactérien, sur glace pendant 10 minutes. S'en suit un choc thermique de 45 secondes à 42°C, puis une incubation de 2 minutes sur glace. 1 mL de LB est ensuite ajouté aux bactéries et ces dernières sont incubées 1 h à 37°C avant d'être déposées sur boîte de Pétri contenant un milieu LB gélifié à l'aide d'agarose et supplémenté de l'antibiotique correspondant à la résistance encodé dans le vecteur transfecté. Les colonies présentes après une incubation d'une nuit à 37°C sont placées en milieu LB supplémenté par les antibiotiques correspondant aux résistances encodés dans le vecteur porteur de la construction et laissées pousser à 37°C sur la nuit.

Enfin, les bactéries ainsi cultivées sont mises dans 1 L de milieu de culture (voir tableau 2). L'expression est induite par ajout d'isopropyl β -D-1-thiogalactopyranoside (IPTG) une fois que la densité optique à 600 nm atteint 0.6 et que les bactéries se trouvent donc en phase exponentielle. Les cellules sont ensuite centrifugées, congelées dans l'azote liquide et conservées à -80°C.

Toutes ces constructions ont été contrôlées par spectrométrie de masse par la plateforme de spectrométrie de masse de l'IBMC en découpant directement les bandes concernées sur gel afin d'identifier les protéines.

4/ Purification des protéines

4.1/ Purification des protéines produites en cellules d'insectes

Un même protocole a été utilisé pour purifier le complexe mouse_ERR α _1-422-PGC-1 α _1-472 (**Fig. S5**), la protéine PGC-1 α _1-635 (**Fig. S6**) et la protéine PCAF (**Fig. S7**) produits en cellules d'insectes après leur infection par baculovirus. Les variations entre les protocoles sont indiquées dans le texte ci-dessous ou dans la partie ANNEXES correspondante.

Les culots de cellules d'insectes ayant exprimé la protéine ou le complexe d'intérêt sont resuspendus dans le tampon de lyse (20 mM Tris HCl pH 8.0, 200 mM NaCl, 4 mM CHAPS, 5 mM MgCl₂, 2 mM TCEP). Ce tampon est supplémenté d'une pastille d'inhibiteurs de protéases par litre de culture, de 5 μ L benzonase nucléase par litre de culture et de 1 mM de fluorure de phénylméthylsulfonyl (PMSF). Après une incubation de 30 minutes suivant l'ajout de benzonase, suivi d'un choc thermique de 5 minutes à 37°C en bain-marie, les cellules sont lysées par deux cycles de sonication d'une minute à l'aide d'un sonicateur Labsonic P (Sartorius) fonctionnant à 30 % d'amplitude selon un cycle de 0.5. Les lysats sont ensuite centrifugés 1 h à 40 000 rpm et 4 °C, puis le surnageant est filtré au filtre Milipore 0.44 μ m. Les filtrats sont ensuite passés sur une colonne d'affinité sensible à la streptavidine (StrepTrap™ HP 5 mL), sauf dans le cas du filtrat de PCAF pour lequel 40 μ L d'une résine de billes d'agarose couvertes d'anticorps anti-FLAG est utilisée par litre de culture. Les protéines/complexes sont élués par le tampon de lyse supplémenté par 5 mM α -desthiobiotine, sauf pour PCAF où celui-ci est supplémenté par 5 mg/mL de peptide FLAG. Les éluats sont concentrés et passés sur colonne de SEC de modèle Superdex 200 increase 10/300 GL (Cytiva) (10 mM HEPES pH 7.9, 120 mM KCl, 0.2 mM MgCl₂, 1 mM TCEP). Les fractions centrales du pic obtenu sont récupérées en vue de former les complexes après contrôle sur gel polyacrylamide en conditions dénaturantes par utilisation de SDS. Les protéines sont contrôlées par diffusion dynamique de la lumière (DLS) pour s'assurer de leur monodispersité et visualiser la présence d'agrégats.

4.2/ Purification des protéines produites en bactéries

Les culots bactériens sont resuspendus et incubés 30 minutes dans un tampon de lyse (40 mM imidazole, 20 mM Tris HCl pH 8.0, 400 mM NaCl, 5 mM MgCl₂, 4 mM CHAPS, 4 mM TCEP, 10 % Glycérol) supplémenté par 1 tablette d'inhibiteurs de protéases par litre de culture (C0mplete EDTA-free), 25 µg de lysozyme par litre de culture, 1 mM de PMSF et 5 µL de benzonase nucléase (BaseMuncher Endonucléase (25 000 U)). Pour human_ERRα_1-423, la concentration d'imidazole dans le tampon de lyse est diminuée à 10 mM.

Puis, après un choc thermique de 5 minutes à 37°C, les cellules sont remises sur glace puis lysées par sonication en deux cycles de 1 minute à l'aide d'un sonicateur Labsonic P fonctionnant à une amplitude de 60 % et une valeur de cycle de 0.5. Les lysats sont centrifugés 45 minutes à 4°C et 40 000 rpm. Les surnageants sont filtrés par filtre Milipore 0.22 µm puis passés sur colonne de nickel en éluant la protéine selon un gradient d'imidazole (0-500 mM d'imidazole, 400 mM NaCl, 4 mM CHAPS, 4 mM TCEP, 20 mM Tris HCl pH 8.0). PGC-1α 74-219 et PGC-1α 138-219 sont digérés par la protéase HRV P3C et human_ERRα_1-423 l'est par la protéase du virus de la gravure du tabac (TEV) sur la nuit à 4°C après avoir effectué un échange de tampon pour enlever l'excès d'imidazole. Une Reverse-IMAC (RIMAC) est réalisée pour éliminer le fragment porteur de l'étiquette poly-histidines (**Fig. S8**).

Quelques modifications ont été apportées au protocole de purification de human_ERRα_1-423. L'avantage de ces changements est de permettre une coupure globalement plus efficace par la protéase en éliminant les traces d'imidazole, de permettre une purification plus efficace par l'ajout de la reverse-IMAC et de la chromatographie échangeuse d'ions, laquelle permet de retirer davantage de HSP70 en un temps beaucoup plus court que la colonne héparine. Après l'IMAC, la protéine était coupée sur la nuit par l'action de la protéase du TEV en ratio molaire 1 : 30 à 4°C avant d'être passée sur colonne d'Héparine et éluée selon un gradient de sel afin d'éliminer les éventuelles contaminations dues à l'ADN. Ce protocole a été modifié afin de passer human_ERRα_1-423 sur colonne échangeuse d'anions (HiTrap™ Q HP 5 mL) avant la coupure par la TEV. L'éluion s'effectue grâce à un gradient de chlorure de sodium (20 mM Tris HCl pH 8.0, 4 mM CHAPS, 2 mM TCEP, 150-1000 mM NaCl, 10 mM MgCl₂). La protéine est ensuite digérée par l'action de la protéase du TEV sur la nuit à 4°C en ratio 1 : 30, puis la protéine est passée sur colonne de Nickel au cours d'une RIMAC, récupérée dans la fraction non retenue sur colonne, avant d'être finalement passée sur colonne de chromatographie d'exclusion stérique de modèle Sepharose increase S200 10/300 GL (10 mM HEPES pH 7.9, 120 mM KCl, 0.2 mM MgCl₂, 1 mM TCEP). PGC-1α_1-285 est purifié selon les mêmes étapes (**Fig. S9**).

4.3/ Purification de p300

Les culots obtenus par centrifugation des cellules d'insectes sont lysés par sonication durant 2 cycles de 1 minutes à l'aide d'un sonicateur Labsonic P fonctionnant à 30 % d'amplitude pour une valeur de cycle de 0.5 en tampon de lyse (50 mM Tris HCl pH 7.5, 400 mM NaCl, 4 mM CHAPS, 1 mM TCEP, 1 mM MgCl₂, 10% glycérol) supplémenté d'une tablette d'inhibiteur de protéase par litre de culture, de benzonase nucléase et de 1 mM de PMSF. Les lysats sont ensuite centrifugés 45 minutes à 4°C à 40 000 rpm et le surnageant est filtré au filtre Millipore 0.44 µm. P300 est ensuite purifiée par une première étape de IMAC en éluant la protéine selon un gradient d'imidazole (0-500 mM imidazole, 50 mM Tris HCl pH 7.5, 400 mM NaCl, 4 mM CHAPS, 1 mM TCEP, 1 mM MgCl₂, 10% glycérol). La protéine est ensuite soumise à une chromatographie échangeuse d'anions et est éluée selon un gradient de sel (220-1000 mM NaCl, 20 mM Tris HCl pH 7.5, 4 mM CHAPS, 2 mM TCEP, 1 mM MgCl₂). La protéine est ensuite finalement soumise à une SEC (10 mM HEPES pH 7.5, 200 mM KCl, 1 mM

CHAPS, 0.2 mM MgCl₂, 1 mM TCEP) (**Fig. S10**). Son intégrité au sein des fractions est contrôlée par DLS.

5/ Formation des complexes

5.1/Formation des complexes ERR-PGC-ADN

Concernant les complexes reconstitués impliquant human_ERR α _1-423, le fragment de PGC-1 α voulu est placé en présence du récepteur en rapport molaire PGC-1 α : ERR α = 3 : 1. Le complexe protéique incube au moins 20 minutes, puis il est ensuite purifié et séparé de l'excès de PGC-1 α par SEC (10 mM HEPES pH 7.9, 120 mM KCl, 0.2 mM MgCl₂, 1 mM TCEP). Il est concentré entre 1 et 4 mg/mL et mis en présence de l'ADN wc26embedded, wc29embedded ou BE42embedded en rapport molaire ERR-PGC : ADN = 1 : 1,1. L'excès d'ADN est retiré par une ultime étape de concentration avant de contrôler la formation du complexe par migration sur gel natif de TRIS CAPS à pH 9.5 pendant 45 minutes à 2 W constant. Le gel est coloré successivement par l'utilisation de Bromure d'éthidium (BET) ou de bleu de toluidine, puis de bleu de coomassie (**Fig. S11**). Dans le cadre de la réalisation des grilles streptavidine, l'ADN biotinyllé BE42embedded est ajouté à une molarité de 90 % de la concentration de ERR-PGC, le but étant d'avoir la protéine en excès dans ce cas particulier. La durée de vie des complexes reconstitués à partir des protéines produites en bactéries est d'environ 4 à 5 jours dans ces conditions après la lyse. Cette durée de vie augmente pour atteindre une semaine après ajout de l'ADN avec une précipitation croissante au cours du temps, notamment de PGC. Il s'agit probablement des régions de la protéine qui ne sont pas protégées par son interaction avec ERR.

Concernant le complexe exprimé par coinfection virale mouse_ERR α _1-422-PGC-1 α _1-472, ce dernier est purifié par SEC de la même manière que les complexes reconstitués. Il est cependant concentré à des valeurs comprises entre 0,1 et 0,6 mg/mL avant ajout de l'ADN (**Fig. S12**). La durée de vie du complexe avant ajout d'ADN est d'environ 36h après la lyse. Après ajout de l'ADN, cette durée de vie augmente sensiblement, mais dépasse rarement les trois jours à 4°C, suite à quoi le complexe est généralement soit partiellement dégradé, soit agrégé, notamment le coactivateur PGC.

5.2/ Formation des complexes ERR-PGC-p300

Le complexe ERR-PGC se voit ajouter p300 dans différents rapport molaire allant de 1 à 5 excès molaire de p300. Les protéines sont ensuite concentrées à une concentration d'au moins 3 mg/mL. Le tampon est échangé lors de la concentration des protéines et le même tampon est ensuite utilisé pour le passage du complexe sur colonne de SEC de modèle Superdex S200 10/300 GL. Les fractions sont ensuite déposées sur gel SDS afin d'identifier les fractions contenant les trois protéines. Les fractions sont également déposées sur gels de polyacrylamide 4 % en conditions natives avant et après le passage par l'étape de SEC à des pH allant de 6.5 à 9.5. La migration s'effectue durant 1 à 4 heures en fonction du pH à 2 W constant.

5.3/ Formation des complexes ERR-PGC-PCAF et ERR-PCAF-ADN

Les protéines ERR et PCAF sont incubées de sorte à ce que ERR soit en excès molaire dans des rapports molaires allant de PCAF : ERR α = 1 : 2 à PCAF : ERR α = 1 : 5. Les protéines sont ensuite concentrées à 5 mg/mL avant de réaliser une SEC sur (10 mM HEPES pH 7.5, 120 mM KCl, 0.2 mM MgCl₂, 1 mM TCEP) sur une colonne de modèle Superdex S200 10/300. On ajoute ensuite l'ADN choisi à l'une des fractions tel qu'il soit en excès molaire de 1,1 par rapport à la concentration estimée

du complexe. Les fractions sont déposées sur gel SDS afin d'identifier les fractions contenant les deux protéines, mais aussi sur gel natif afin de conclure sur l'apparition ou non d'un complexe. Le protocole demeure inchangé dans le cas de la formation du complexe ERR-PGC-PCAF. Dans ce cas, les protéines ERR et PGC correspondent aux constructions human_ERR α _1-423 et PGC-1 α _1-285 afin de pouvoir les concentrer efficacement et d'augmenter leur durée de vie individuelles.

6/ Cristallogénèse et diffraction aux rayons X

Les essais de cristallogénèse ont été réalisés par l'utilisation du Mosquito (plateforme de Biologie structurale intégrative de Strasbourg Instruct-ERIC de l'Institut de Génétique et de Biologie Moléculaire et Cellulaire - Centre de biologie structurale intégrative, supportée par FRISBI (ANR-10-INBS-0005)) en utilisant les kits commerciaux suivant : TOP96, PEGS, Morpheus, Classics, JCSG+. Les concentrations de complexes protéines/ADN impliquant PGC-1 α _74-219 vont de 2.2 à 9 mg/mL et de 3.5 à 5 mg/mL pour celles impliquant le fragment PGC-1 α _138-219. L'évaporation s'effectue par la technique de la goutte assise. Les cristaux obtenus à 20°C ont ensuite été pêchés à l'aide d'une boucle litholoop (Molecular Dimensions Limited) et, si besoin, plongés dans une solution cryo-compatibles avant d'être congelés rapidement en bain d'azote liquide, puis testés sur la ligne de lumière du centre de biologie structurale intégrative supportée par FRISBI, puis sur la ligne de lumière Proxima 2 du Synchrotron SOLEIL (Gif-sur-Yvette, France). Les rares cristaux obtenus ne diffractant qu'à une résolution maximale de 30 Å, cette partie ne sera pas abordée dans les résultats.

7/ Analyses biophysiques du complexe ERR α -PGC-1 α

Plusieurs analyses biophysiques ont été réalisées pour caractériser le complexe ERR α -PGC-1 α . Les protocoles décrits ici sont ceux utilisés pour réaliser ses caractérisations.

7.1/ Analyses par la méthode FIDABio

La protéine human_ERR α _1-423 ainsi que le complexe human_ERR α _1-423-PGC-1 α _1-285 sont amenés à une concentration de 20 μ M en tampon de SEC supplémenté par 0,03 % d'acide pluronique pour un volume final de 100 μ L. Ils sont ensuite centrifugés 30 minutes à 15 000 *g*. On prépare ensuite 100 μ L de marqueur blue-NHS à 40 μ M, puis on mixe les deux préparations 30 minutes à température ambiante dans l'obscurité.

L'échantillon est déposé sur une colonne « NAP » (Cytiva) équilibrée en tampon de SEC supplémenté de 0,03 % d'acide pluronique. On ajoute 300 μ L de tampon après passage de l'échantillon, puis on élue avec 600 μ L la protéine marquée en récoltant les fractions. Les fractions sont contrôlées par mesure au Fidabio en s'assurant d'avoir un rapport signal sur bruit suffisant (valeur supérieure à 30). Les mesures sont réalisées à 10°C avec human_ERR α _1-423 marqué à 400 nM et le complexe human_ERR α _1-423-PGC-1 α _1-285 à une concentration de 200 nM. Pour cela, des tubes contenant des concentrations croissantes de l'ADN testé en présence d'une concentration fixe de complexe marqué sont préparés. Cela est effectué en réalisant la dilution en série de l'ADN dans un volume fixe à la concentration désirée de protéines. Lors de l'expérience, chaque complexe est analysé dans un tampon contenant l'ADN testé à la même concentration que dans le tube échantillon contenant le complexe.

7.2/ Préparation des échantillons pour les analyses par spectrométrie de masse

Les échantillons human_ERR α _1-423, human_ERR α _1-423-PGC-1 α _1-285, PGC-1 α _1-285, human_ERR α _1-423-BE42embedded et human_ERR α _1-423-PGC-1 α _1-285-BE42embedded sont mis en tampon de SEC puis centrifugés 30 minutes à 15 000 g afin de s'affranchir des agrégats. Les échantillons sont ensuite portés chez nos collaborateurs de l'équipe de Sarah Cianféroni de l'institut pluridisciplinaire Hubert Curien. Puis, deux autres préparations effectuées en remplaçant l'HEPES du tampon de SEC par un tampon KH₂PO₄/KHPO₄ à pH 6.5 et pH 8.5 ont été effectuées et données à la plate-forme de spectrométrie de masse de l'institut pluridisciplinaire Hubert Curien pour les analyses de XL-MS et de HDX.

7.3/ Préparation des échantillons pour analyse par la méthode Protéoplex

Les protéines human_ERR α _1-423, PGC-1 α _1-285, p300 et PCAF sont purifiées selon la méthode précédemment décrite. Le complexe human_ERR α _1-423-PGC-1 α _1-285-p300-ADN est formé à 2,5 μ M (20 mM Tris HCl pH 8.0, 1 mM CHAPS, 1 mM TCEP, 200 mM KCl, 0.2 mM MgCl₂). Le complexe human_ERR α _1-423-PCAF-ADN est également formé à une concentration de 2.26 μ M dans le même tampon. Ces complexes sont ensuite fournis à la plate-forme de Novalix. Ils procèdent à la dilution en trois étapes des échantillons. Puis, deux molécules de marquages sont testées sur nos complexes : le Dye Sypro Orange et le Bodipy FL L-Cystéine. En fonction de la molécule de marquage retenue (Dye Sypro Orange pour le complexe ERR-PGC-p300-ADN et Bodipy FL L-Cystéine dans le cas de ERR-PCAF-ADN), l'échantillon est dilué et marqué à une certaine concentration avant d'être placé dans plusieurs conditions de tampons différentes. Les mesures sont ensuite réalisées selon un gradient de température allant de 20 à 90°C.

8/Vitrification du complexe ERR-PGC

8.1/ Vitrification du complexe mouse ERR α 1-422-PGC1-472

Le complexe est amené à une concentration entre 0.3 et 0.6 mg/mL pour vitrification au vitrobot. Les grilles Quantifoil R2/2 et R1.2/1.3 non supportées ou avec une couche de graphène oxydé sont soumises à une décharge lumineuse de 1 minute sous vide à l'aide d'un gaz contenant 80 % d'argon et 20 % d'oxygène. On place ensuite les grilles une par une dans la chambre du vitrobot à 100 % d'humidité et à 4°C. On dépose l'échantillon à hauteur de 3 μ L juste après y avoir ajouté 0,0025 % de n-dodecyl- β -D-maltoside (DDM), puis le vitrobot absorbe l'excès de l'échantillon en appliquant une pression à l'aide du papier filtre durant 1.5 s et avec un paramètre de force de 7.

Des grilles de Streptavidine constituées d'une couche continue de carbone, d'une monocouche de lipides biotinylés et d'une monocouche de cristaux 2D de streptavidines sont utilisées pour capturer le complexe mouse_ERR α _1-422-PGC1-472-BE42embeddedbio dans le but de l'observer après sa coloration négative. Il s'agit d'une approche que novatrice développée par Nils Maréchal sur la plate-forme de biologie structurale intégrative de l'institut de génétique et de biologie moléculaire et cellulaire que j'ai eu l'occasion de tester vers la fin de ma thèse. Les grilles ont été réhydratées 20 minutes dans l'eau, puis lavées 3 fois dans le tampon avant qu'on y dépose 3 μ L d'échantillons à une concentration de 200 nM durant 20 minutes à 4°C. La grille est ensuite lavée trois fois dans l'eau, puis le surplus d'échantillon est absorbé au papier filtre. La grille est trempée dans une solution contenant 2% d'acétate d'uranyl. Après une incubation de 20 minutes, les grilles sont lavées dans l'eau puis observées au microscope Tecnai F20 (TFS) à 200 kV. Voir ce type d'objet sur ce type de grilles à une

concentration de 200 nM malgré le contraste induit par le support est déjà un premier obstacle de taille de franchi. Des grilles destinées à la cryo-EM sont ensuite réalisées à une concentration de 50 nM de complexe en suivant le même protocole sans l'application de l'acétate d'uranyl. La vitrification a lieu au vitrobot.

8.2/ Vitrification du complexe human $ERR\alpha$ 1-423-PGC-1 α 1-285

Le complexe est amené à une concentration de 0.5 mg/mL pour vitrification au vitrobot et 1,1 mg/mL pour vitrification au Chaméléon.

Pour la vitrification au vitrobot, les grilles d'or UltraGold sont soumises à une décharge lumineuse de 30 s dans une chambre contenant un mélange de 80 % d'argon et de 20 % d'oxygène. Elles sont ensuite placées dans la chambre du vitrobot à 100 % d'humidité et à 4°C. Puis, on dépose l'échantillon à hauteur de 3 μ L sur la grille après y avoir ajouté 0.0025 % de DDM. Le vitrobot absorbe le surplus de l'échantillon au papier filtre durant 1.5 s pour un paramètre de force de 7 et plonge rapidement la grille en éthane liquide. Dans le cas des grilles Quantifoil R2/2 non supportées, ainsi que des grilles de graphène oxydé, la décharge lumineuse dure 1 minute, mais les autres conditions de vitrification demeurent inchangées.

Pour la vitrification au Chaméléon, la décharge lumineuse est appliquée sur les grilles auto-absorbantes (Wei *et al.*, 2018) sous vide avec les traces de l'air ambiant. Plusieurs conditions d'application de la décharge sont testées, ainsi que plusieurs vitesses de congélation allant de 109 ms à 2 s. Deux applications sont effectuées à chaque vitrification : la première sert à quantifier la vitesse nécessaire pour la vitrification et à estimer la qualité de la grille qui sera obtenue, la seconde conduit à la vitrification de la grille d'après les paramètres décidés lors du premier essai. Pour chaque grille réalisée, l'échantillon s'est vu supplémenté de 0.0025 % de DDM lors de sa préparation.

9/ Acquisitions des jeux de données et pré-traitement

Un jeu de données du complexe human_ $ERR\alpha$ _1-423-PGC-1 α _1-285-BE42embbeded vitrifié au vitrobot sur grille or et constitué de 2583 films a été collecté au cryo-microscope électronique Glacios à une tension d'accélération de 200 kV à l'aide d'une caméra à détection directe K2 Gatan en mode Comptage et à une taille de pixel de 0,901 $\text{\AA}^2/\text{pixel}$. La stratégie de collecte a consisté en l'acquisition de trois films par trous en inclinant le faisceau afin d'atteindre les différentes positions ciblées, inclinaison qui fut corrigée en amont lors de l'alignement du microscope. Cette stratégie de collecte permet de limiter les mouvements du stage, aussi parfois appelé « platine », tenant la grille. Il n'y a, avec cette méthode, qu'un mouvement du stage tous les 25 trous, soit toutes les 75 images (**Tableau 5**). Cela permet notamment l'augmentation de la vitesse de collecte et donc l'acquisition de plus de données lors de la durée de l'acquisition, puisque cela permet de passer outre le temps de relaxation qui suit le mouvement du stage et qui est nécessaire afin de s'assurer qu'aucune vibration ne perturbe l'acquisition en cours. L'alignement des 50 images pour produire un micrographe a été réalisé dans RELION et CryoSPARC en utilisant leur implémentation de MotionCorr2. L'estimation globale des paramètres de la CTF a été réalisée grâce à CTFFIND4.1. Les particules ont ensuite été sélectionnées par CrYOLO ou par l'outil « blob picker » de CryoSPARC.

Mag	45000
Physical pixel size (Å/pix)	0.901
C2 Aperture (µm)	30
C2 Current (%)	39.6
Spot size	6
Objective Aperture (µm)	100
Beam size (µm)	~1.1
Target defocus (- µm)	-1.2 → -2.5
Camera	K2 counting mode
Dose rate (e/pix/s)	5.632
Dose rate (e/Å ² /s)	6.938
Time for each exposure (s)	8
Total dose (e/Å ²)	55.501
Number of frames	50
Dose per frame (e/Å ²)	1.110
For SPA	
Number of images per hole	3
Distance (µm)	0.46
Total images	2583
Speed (images/hour)	~172

Tableau 5 : Conditions d'acquisition du complexe human ERR α 1-423-PGC-1 α 1-285-BE42embedded sur grille d'or

Un autre jeu de données du même complexe a été acquis à une taille de pixel de 0,901 Å²/pixel sur une grille réalisée au Chaméleon. Cette grille avait été produite avec l'application d'une décharge lumineuse de 30 secondes et une application de l'échantillon en deux étapes pour un temps de vitrification de 108 ms. 3748 films ont été acquis avec la même stratégie de collecte que décrit précédemment à ceci près que les mouvements du stage ont été effectués tous les 9 trous, puis une nouvelle acquisition sur la même grille a permis de collecter 15 524 films supplémentaires (**Tableau 6**). L'alignement des images a été réalisé « on the fly », c'est-à-dire en même temps que l'acquisition, dans WARP. L'estimation globale de la CTF sur chaque micrographie a été réalisée à l'aide du logiciel CTFFIND4.1. La sélection des particules s'est vu effectuée à l'aide du logiciel CrYOLO ou par l'outil « blob picker » de CryoSPARC.

Mag	45kx TEM nP
Physical pixel size (Å/pix)	0.901
C2 Aperture (µm)	30
C2 lens	40.8
Spot size	5
Objective Aperture (µm)	100
Screen current (nA)	
Beam size (nm)	
Target defocus (- µm)	-1.0 to -2.5
Camera	K2
Dose rate (e/pix/s)	9.4
Dose rate (e/Å ² /s)	11.6
Time for each exposure (s)	6
Total dose (e/Å ²)	69.5
Number of frames	50
Dose per frame (e/Å ²)	1.4
For SPA	
Number of images per hole	3
Distance (µm)	0.46
Number of holes per stage movement	9
Total images	15524
Speed (images/hour)	232

Tableau 6: Conditions de la seconde acquisition du complexe human ERR α 1-423-PGC-1 α 1-285-BE42embedded vitrifié au Chameleon

Un jeu de donnée du complexe mouse_ERR α _1-422-PGC-1 α _1-472-BE42embeddedbio constitué de 2945 films a été collecté au cryo-microscope électronique Glacios à 200 kV à l'aide d'une caméra à détection directe K2 Gatan en mode Comptage. Ce jeu a été collecté sur une grille supportée par un film de carbone et recouverte d'une monocouche 2D de cristaux de streptavidine à une taille de pixel de 0,901 Å²/pixel. Il s'agit d'un jeu de donnée effectué à partir d'une grille vitrifiée au vitrobot. L'alignement des images pour produire les micrographes a été réalisé avec les logiciels RELION et CryoSPARC en utilisant leur implémentation de MotionCorr2. L'estimation globale de la CTF a été effectuée à l'aide du logiciel CTFFIND4.1. Les particules ont ensuite été sélectionnées par crYOLO ou par l'outil « blob picker » de cryoSPARC.

Un jeu de donnée du complexe mouse_ERR α _1-422-PGC-1 α _1-472-BE42embedded constitué de 4770 films a été collecté au Titan Krios à 300 kV à l'aide d'une caméra à détection directe K3 Gatan en mode Comptage. Ce jeu a été collecté sur une grille Quantifoil R2/2 à une taille de pixel de 0,862 Å²/pixel. Il s'agit d'un jeu de donnée effectué à partir d'une grille vitrifiée au vitrobot (**Tableau 7**). L'alignement des images pour produire les micrographes a été réalisé avec les logiciels RELION et cryoSPARC en utilisant leur implémentation de MotionCorr2. L'estimation globale de la CTF a été effectuée à l'aide du logiciel CTFFIND4.1. Les particules ont ensuite été sélectionnées par crYOLO ou par l'outil « blob picker » de cryoSPARC.

Microscope	Titan Krios G1
Pixel size	0.862
Target defocus	-1 to - 3,5 μm
Caméra	Gatan K3
Time for each exposure	8
Total Dose	43
Number of frames	40
Dose per frame	0.93
Number of image per holes	1
Total images	4770

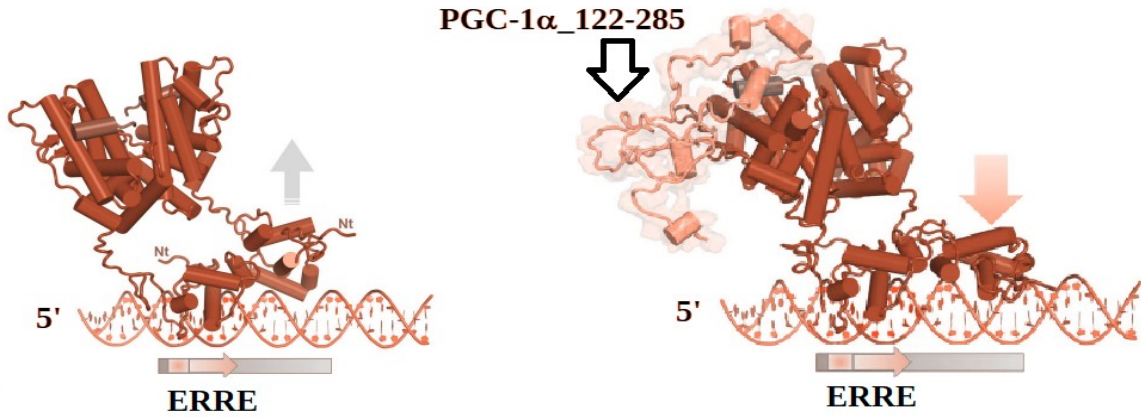
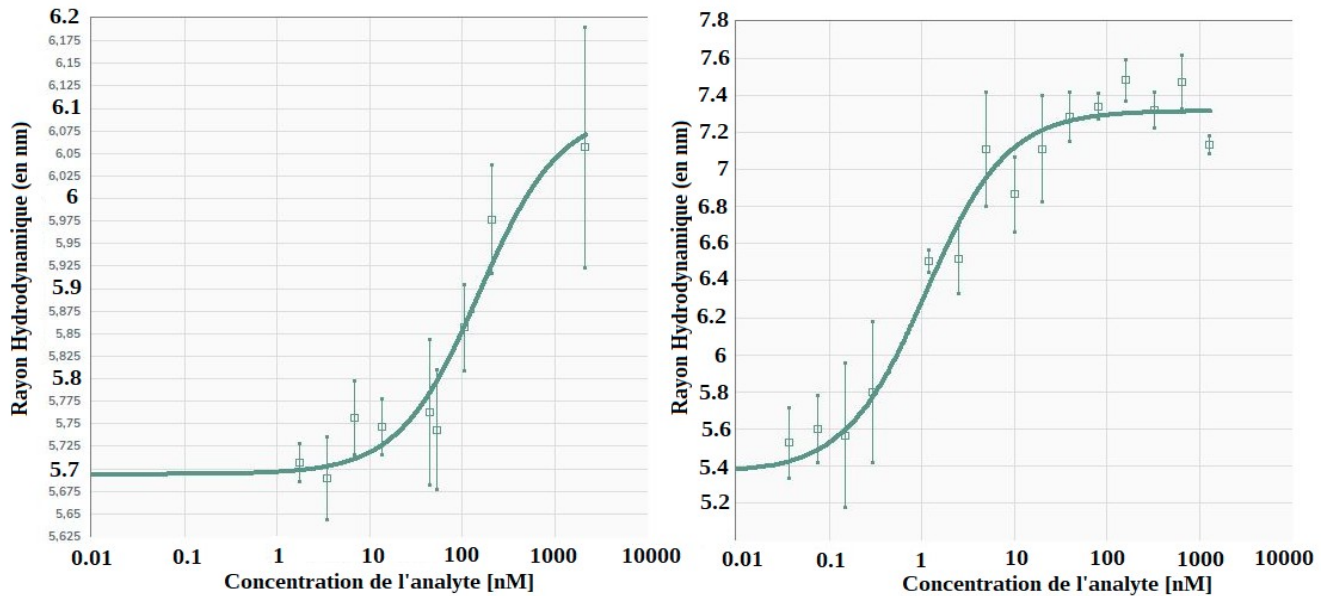
Tableau 7: Conditions de la seconde acquisition du complexe human $ERR\alpha$ 1-422-PGC-1 α 1-472-BE42embedded vitrifié au Vitrobot

Résultats

1/ PGC-1 α permet l'assise de ERR α sur ses ERRE

L'analyse des données SAXS du récepteur $\Delta A/B_ERR\gamma$ sur un ERRE effectuée au sein de notre équipe a permis de produire des enveloppes et un modèle tout atome décrivant l'effet stabilisateur du coactivateur PGC-1 α sur la liaison à l'ADN (**Fig. 36-A**). Le modèle obtenu indique que le second DBD de ERR γ ne serait pas lié sur l'élément de réponse, mais que l'interaction avec PGC-1 α sur le LBD situé en position 5' induirait la fixation du DBD en 3' sur le second demi-site.

Cette hypothèse est renforcée par les mesures effectuées au FIDA bio impliquant ERR α sur le même élément de réponse de type ERRE (**Fig. 36-B**). Les mesures réalisées à 10°C dévoilent un changement du K_d de l'interaction avec l'ERRE dans le cadre du complexe ERR α -PGC-1 α par comparaison avec le récepteur seul, passant de 144 nM environ à moins de 1 nM lors de l'ajout de PGC-1 α (**Fig. 36-C**). Ces mesures, réalisées à l'aide de protéines marquées et de la dilution en série de l'ADN, peuvent contenir des incertitudes quant à la concentration exacte de l'ERRE, celle-ci étant trop basse pour être contrôlée par l'analyse de l'absorbance à 260 nm. Néanmoins, le changement de K_d lors de l'ajout de PGC-1 α demeure significatif. Le rayon hydrodynamique entre ERR α sans et avec ADN varie assez peu, passant de 5.69 à 6 nm. Une telle différence demeure dans la gamme de mesure de l'appareil, bien qu'il s'agisse de sa limite basse. Les mesures ont été reproduites à plusieurs reprises afin de s'assurer de la validité des résultats, mais davantage de mesures seront nécessaires afin d'établir l'incertitude sur le K_d de l'interaction. La faible variation observée peut s'apparenter à un faible changement de rayon hydrodynamique global. La variation de rayon hydrodynamique entre le récepteur seul et le récepteur lié à PGC-1 α est très faible, passant de 5.69 nm à 5.36 nm. Cela pourrait indiquer que le complexe est plus compact que le récepteur seul et/ou que PGC-1 α n'est pas détecté et ne contribue pas à la mesure du rayon hydrodynamique. La variation de rayon hydrodynamique entre le complexe ERR α -PGC-1 α avant et après ajout de l'élément de réponse permet d'observer un changement plus important du rayon hydrodynamique, lequel passe de 5.36 à 7.31 nm. Cette différence est significative et traduit un changement de la conformation du complexe lors de l'ajout de l'ADN par comparaison avec le récepteur seul.

A/**B/****C/**

Affinity constant	144 nM
Indicator size	5,69 nm ~ 0,04 nm
Complex size	6,097 nm ~ 0,13 nm
R ²	0,93
Model	1:1

Affinity constant	0,82 nM
Indicator size	5,36 nm ~ 0,02 nm
Complex size	7,31 nm ~ 0,18 nm
R ²	0,96
Model	1:1

Figure 36: PGC-1 α permet l'ancrage des ERR sur les ERRE

Légende :

A/ Modélisations (rigid-body modeling) avec tout atome de $\Delta A/B_ERR\gamma$ et de $\Delta A/B_ERR\gamma$ -PGC-1 α _122-285 sur l'élément ERRE générées par Dadimodo (Evrard et al., 2011) à partir des données SAXS. **B/** Courbes décrivant l'évolution du rayon hydrodynamique de human_ERR α _1-423 (à gauche) et du complexe formé par human_ERR α _1-423 et PGC-1 α _1-285 (à droite) en fonction de la concentration théorique de l'ERRE Tr α ERRE obtenues par FIDABio. **C/** Résultats obtenus par l'analyse de la courbe d'interaction entre human_ERR α _1-423 (à gauche) et le complexe human_ERR α _1-423-PGC-1 α _1-285 sur l'élément Tr α ERRE (à droite).

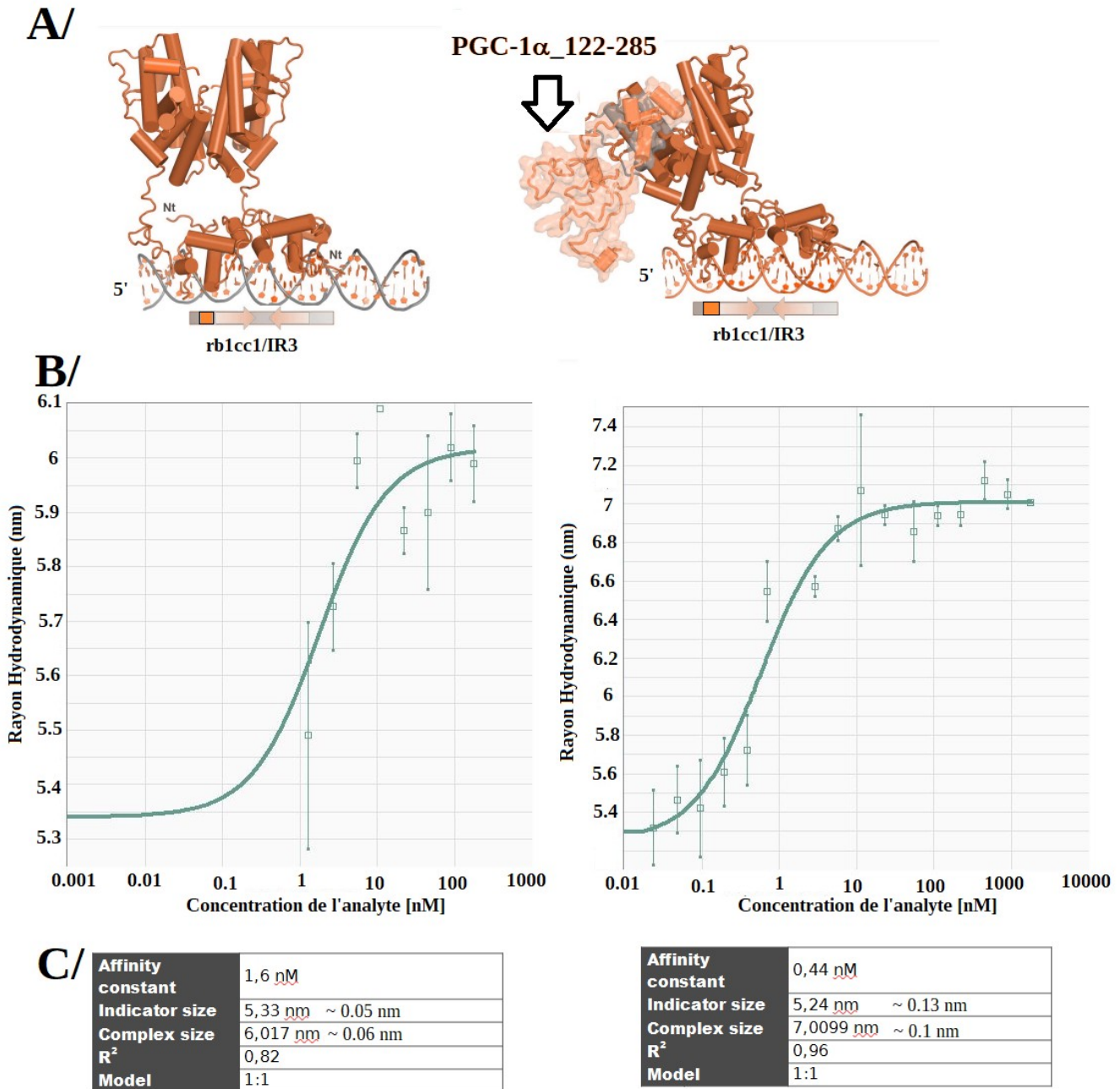


Figure 37: Effet de PGC-1 α sur l'interaction entre ERR α et un élément composite

Légende :

A/ Modélisations (rigid-body modeling) avec tout atome de $\Delta A/B_ERR\gamma$ et de $\Delta A/B_ERR\gamma$ -PGC-1 α _122-285 sur un élément rb1cc1/IR3 générées par Dadimodo (Evrard et al., 2011) à partir des données SAXS. **B/** Courbes décrivant l'évolution du rayon hydrodynamique de human_ERR α _1-423 (à gauche) et du complexe formé par human_ERR α _1-423 et PGC-1 α _1-285 (à droite) en fonction de la concentration théorique de l'élément composite obtenues par FIDAbio. **C/** Résultats obtenus par l'analyse de la courbe d'interaction entre human_ERR α _1-423 (à gauche) et le complexe human_ERR α _1-423-PGC-1 α _1-285 sur l'élément composite (à droite).

L'analyse des données SAXS du récepteur $\Delta A/B_ERR\gamma$ sur un élément composite rb1cc1/IR3 effectuée par l'équipe a permis de produire des enveloppes et un modèle tout atome décrivant le récepteur doublement lié à cet ADN (**Fig. 37-A**). Le modèle obtenu indique que le second DBD de $ERR\gamma$ serait lié sur l'élément de réponse et que l'interaction avec $PGC-1\alpha$ sur le LBD situé en position 5' induirait une modulation du positionnement des LBDs.

Les mesures effectuées au FIDAbio impliquant $ERR\alpha$ sur l'élément composite BE26embedded (**Fig. 37-B**) et réalisées à 10°C dévoile un K_d plus favorable que lors de l'interaction avec un ERRE. Le K_d mesuré est de l'ordre de 1.6 nM (**Fig. 37-C**). Dans le cadre du complexe $ERR\alpha$ - $PGC-1\alpha$ par comparaison avec le récepteur seul, le K_d passe à une valeur plus favorable : 0.44 nM (**Fig. 36-C**). Ces mesures souffrent des mêmes problèmes que ceux décrits précédemment et sont très dépendantes de la dilution en série effectuée de l'ADN. Elles ont été effectuées le même jour et avec le même stock de protéines que les mesures avec l'élément $Tr\alpha$ ERRE (**Fig. 36**). Le K_d mesuré est ici plus discutable et il est difficile d'affirmer que la différence de K_d est réellement représentative tant que les résultats n'ont pas été reproduit et que l'incertitude sur le K_d n'a pas été définie. Le rayon hydrodynamique entre $ERR\alpha$ sans et avec ADN varie un peu plus que les mesures effectuées avec l'élément $Tr\alpha$, passant de 5.33 à 6 nm. Le récepteur complexé avec l'ADN dispose du même rayon hydrodynamique avec les deux ADN testés. La variation de rayon hydrodynamique entre le récepteur seul et le récepteur lié à $PGC-1\alpha$ est très faible, passant de 5.33 nm à 5.24 nm, ce qui correspond au même phénomène de diminution du rayon hydrodynamique que précédemment observé. La variation de rayon hydrodynamique entre le complexe ERR - PGC avant et après ajout de l'élément de réponse permet d'observer un changement plus important du rayon hydrodynamique, lequel passe de 5.24 à 7 nm. Cette différence est significative et traduit un changement de conformation global du complexe lors de l'ajout de l'ADN plus important que celui du le récepteur seul. Le rayon hydrodynamique de ce complexe est, semble-t-il, plus petit que celui impliquant l'élément $Tr\alpha$. La variation observée entre les rayons hydrodynamique est cependant du même ordre de grandeur.

$PGC-1\alpha$ augmente donc l'affinité du récepteur pour les éléments de réponse ERRE. Le K_d de l'interaction avec les ERRE pour le complexe $ERR\alpha$ - $PGC-1\alpha$ devient alors du même ordre de grandeur que pour l'interaction de $ERR\alpha$ avec un élément composite. Ces résultats vont donc dans le sens de l'hypothèse suggérée par les résultats des données SAXS.

2/ PGC-1 α interagit avec ERR α au niveau d'une région très spécifique

L'analyse des données récoltées par XL-MS de ERR α avec ou sans le coactivateur, ainsi qu'avec ou sans l'élément de réponse composite permet d'identifier les régions de l'espace physiquement proches en solution de chacun des partenaires. L'expérience suivante a été réalisée trois fois pour ERR α et le complexe ERR α -PGC-1 α en triplicat, et une fois pour les conditions avec l'élément de réponse composite. Les expériences ont été réalisées deux fois à un pH de 8.5 et une fois à un pH de 6.5. Seuls les résultats de l'expérience incluant les quatre conditions avec l'ADN, à pH 8,5, et apparaissant dans au moins deux des trois répliquas seront discutés ici (**Tables S1, S4, S5 et S7**).

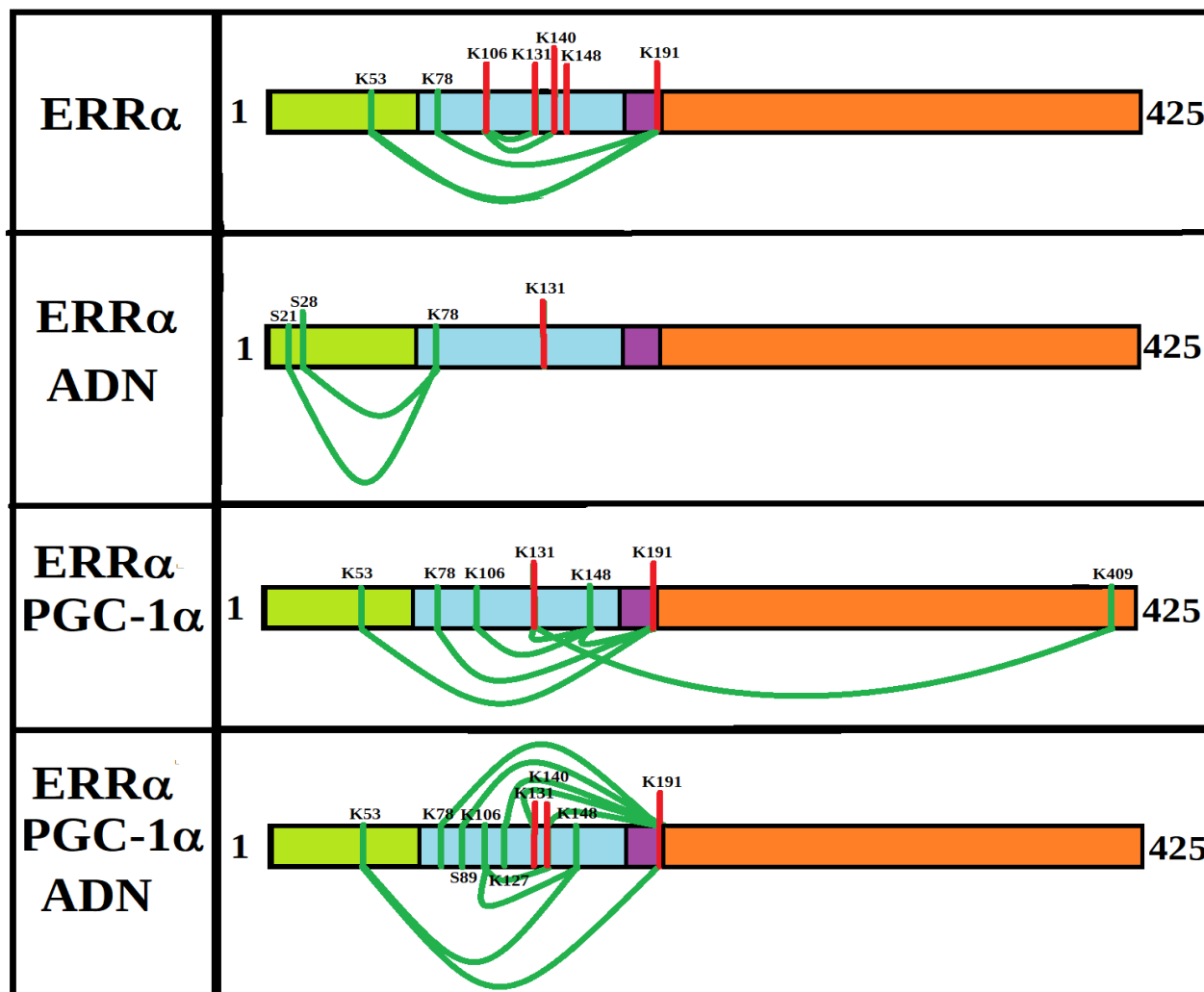


Figure 38: Représentation des réticulations intra-moléculaires de ERR α en fonction de PGC-1 α et l'ADN

Les réticulations intramoléculeaires sont indiquées en vert dans les différentes conditions entre ERR α . Les domaines de la construction human_ERR α _1-423 sont indiqués comme suit : le domaine N-terminal en vert, le DBD en bleu clair, la région coude en violet, LBDs en orange. Les barres rouges indiquent une réticulation entre le même résidu sur chacun des monomères de ERR α .

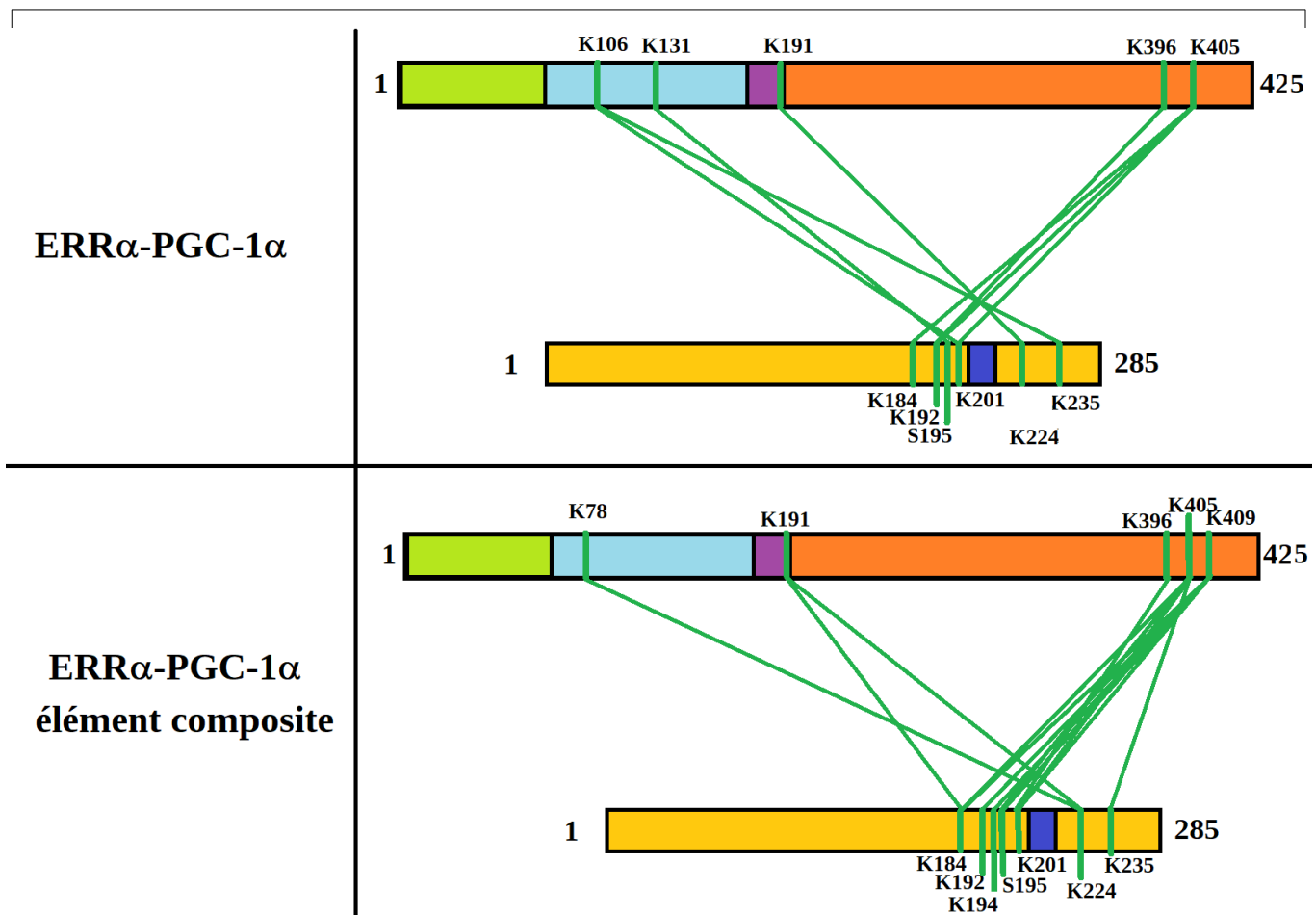


Figure 39: Représentation des réticulations observées entre $ERR\alpha$ et $PGC-1\alpha$

Les réticulations sont indiquées en vert dans les différentes conditions entre $ERR\alpha$ et $PGC-1\alpha$. Les domaines de la construction human_ $ERR\alpha$ _1-423 sont indiqués comme suit : le domaine N-terminal en vert, le DBD en bleu clair, la région coudeuse en violet, LBDs en orange. La construction human_ $PGC-1\alpha$ _1-285 est indiquée en jaune avec le motif « PCSELLKYL » représenté en bleu foncé.

Légende :

Réticulations inter-moléculaires entre $ERR\alpha$ et $PGC-1\alpha$ dans les deux conditions impliquant $PGC-1\alpha$.

Le suivi des réticulations de $ERR\alpha$ dans les différentes conditions testées permet d'apprécier la formation et la disparition de certaines réticulations (**Fig. 38**). La comparaison entre la condition avec et sans élément de réponse montre une nette disparition de la plupart des réticulations. Ce phénomène peut-être lié à l'utilisation d'ADN qui, d'une part, interagit avec un grand nombre de lysines des DBDs et peut ainsi les camoufler de l'agent réticulant et, d'autre part, interagit potentiellement avec ses groupements chimiques avec les agents réticulants. Il peut néanmoins être observé qu'un rapprochement spatiale a lieu entre la région N-terminale, et plus particulièrement entre la région impliquée dans les modifications post-traductionnelles des Ser19, Ser21 et Lys14 et les DBDs. Ce résultat donne des indications sur la façon dont les modifications post-traductionnelles pourraient inhiber l'activité transcriptionnelle du récepteur. De plus, les réticulations observées au sein du récepteur en l'absence d'ADN suggère que les DBDs seraient proches l'un de l'autre et probablement

en interaction l'un avec l'autre en solution en l'absence d'ADN. Cette interaction serait donc déstabilisée si la séquence de l'élément de réponse ne permet pas la fixation du second DBD.

La comparaison entre la condition avec et sans PGC permet d'apprécier la disparition et l'apparition plus importante de certaines réticulations intra-moléculaires entre les DBDs. Il y a notamment un rapprochement des résidus de la région DBD avec la région LBD et donc un changement dans la topologie du récepteur. Il apparaît aussi qu'au moins l'un des DBDs se rapproche spatialement d'au moins l'un des LBDs et plus particulièrement du site d'interaction connu entre ERR et PGC. Ce résultat sera confirmé par les deux autres expériences réalisées en XL-MS laissant voir qu'au moins l'un des DBDs se rapproche spatialement du site d'interaction avec PGC-1 α (**Table S.6**).

La comparaison de la condition ERR-PGC et ERR-PGC-élément-composite révèle un plus grand nombre de réticulations dans la condition avec l'élément composite. Ces dernières concernent plus particulièrement la Lys191 de ERR α avec des résidus de la région DBD, indiquant un rapprochement global des domaines DBD et des domaines LBDs. Ces rapprochements indiquent également que la structure globale du récepteur est probablement plus compact. Ces résultats sont néanmoins à pondérer avec ceux obtenus pour le complexe sans ADN et au récepteur seul lors des deux autres expériences (**Tables S.2 et S.6**).

Le suivi des réticulations intermoléculaire entre ERR α et PGC-1 α permet de délimiter les régions spatialement proches dans l'espace entre les protéines, de poser certaines bases sur la topologie du complexe et de considérer l'impact de l'élément de réponse sur l'interaction entre les deux protéines (**Figure 39**). Les conditions ERR α -PGC-1 α et ERR α -PGC-1 α -élément-composite permettent d'apprécier une région de 51 résidus (184-235) du coactivateur spatialement proches du récepteur et délimitant la région d'interaction réelle entre les deux protéines. Cette région inclue le motif « PCSELLKYL » connue pour interagir avec ERR α d'après la structure cristallographique de ses LBDs (PDB : 3D24), ce qui est un gage de qualité de l'expérience. Plus surprenamment, des régions de PGC-1 α se trouve spatialement proches des DBDs, ce qui constitue un résultat original qui peut traduire, d'une part, le rapprochement d'au moins l'un des LBDs des DBDs, mais aussi le rapprochement physique et l'hypothèse d'une interaction entre le coactivateur et les DBDs. Les résidus concernés étant de part et d'autres du motif « LLKYL » de PGC-1 α dans le cas du complexe sans ADN et en position C-ter par rapport au motif dans le cas du complexe avec ADN, il est possible que cette interaction soit influencée par la présence de l'élément de réponse. Les autres expériences de XL-MS confirment ces résultats pour le complexe ERR α -PGC-1 α (**Table S.6**).

3/ PGC-1 α induit le changement conformationnel de ERR α

Les expériences de HDX réalisées sur le complexe ERR α -PGC-1 α révèlent des changements des régions protégées et déprotégées par comparaison avec ERR α (**Fig. 40**).

Au niveau du domaine LBD, ERR α présente plusieurs régions protégées. Les hélices H11 et H12, ainsi que la boucle H11-H12, sont protégées de l'échange hydrogènes-deuterium lors de l'interaction avec PGC-1 α . Considérant la structure cristallographique des LBDs (PDB : 3D24), la protection de l'hélice H12 est un résultat attendu et un gage de qualité de l'expérience réalisée. Il est raisonnable de penser, au vu de ces résultats, et de ceux de XL-MS précédent que PGC-1 α interagit directement avec la boucle H11-H12. La région correspondant aux résidus 213 à 230 est également protégée lors de l'interaction avec le coactivateur, ainsi que la région correspondant aux résidus 205 à 211, bien que cette dernière le soit plus faiblement. Ces régions correspondent à la boucle entre H1 et H3 du récepteur ERR α . Ce résultat va dans le sens de l'hypothèse selon laquelle, dans les récepteurs de classe I, l'hélice H1 des LBDs et la boucle H1-H3 seraient impliquées au sein d'un canal de communication entre le site de liaison des coactivateurs et l'interface de dimérisation en passant par les hélices H9 et H10 pour lesquelles une très faible protection est observable (Billas et al., 2013). Enfin, les résidus 296 à 304 sont déprotégés en présence de PGC-1 α . Ces résidus correspondent à l'hélice H6 des LBDs. La déprotection de ces résidus peut s'associer à une déstabilisation dynamique de l'hélice H6.

Au niveau du domaine DBD, deux régions sont protégées sous l'action de PGC-1 α . Il s'agit des régions correspondant aux boîtes fonctionnelles D (résidus 116 à 122) et T (résidus 147 à 155), lesquelles sont impliquées dans la dimérisation des DBDs. Leur protection peut donc être associées à un changement d'état conformationnel et/ou à la mise en place de l'interface de dimérisation. Deux régions du DBD sont déprotégées lors de l'interaction de PGC-1 α . Il s'agit de la région correspondant à l'hélice H1 des DBDs et comportant la boîte P (résidus 100 à 114), ainsi qu'une partie de l'hélice H2 (résidus 143 à 146). La déprotection de ces régions s'associe vraisemblablement à leur exposition au solvant et à un possible changement conformationnel ou, au moins, à une déstabilisation dynamique.

Enfin, une déprotection forte des résidus 16 à 38 du domaine N-terminale se produit en présence de PGC-1 α . Cette région est prédite comme n'étant pas structurée en solution. Elle contient trois sites de modifications post-traductionnelles dont le rôle est l'inhibition de l'activité transcriptionnelle associée au récepteur : la Lys16, la Ser21 et la Ser24. La déprotection de cette région pourrait signifier que cette dernière est enfouie ou structurée en l'absence du coactivateur et qu'elle se retrouve exposée au solvant lors de l'interaction. Il s'agit probablement d'un mécanisme de régulation de l'activité du complexe. La région correspondant aux résidus 7 à 14 est légèrement protégée lors de la liaison de PGC-1 α .

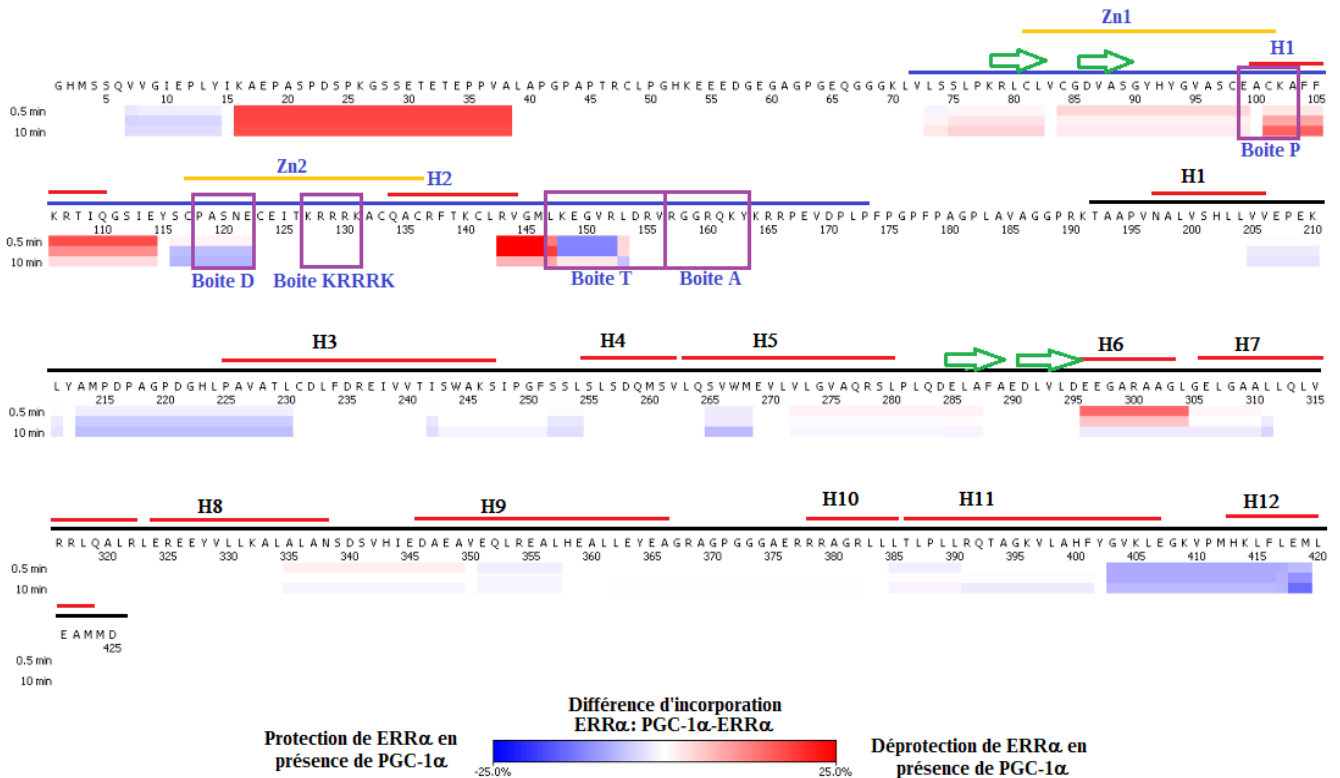


Figure 40: Schéma des régions protégées et déprotégées sous l'influence de PGC

Légende :

La séquence de human_ERRα_1-423 est présentée ici avec les régions déprotégées en rouge et protégées en bleu pour différents temps de marquage du complexe ERRα-PGC-1α par comparaison avec l'absence du coactivateur. La barre noire délimite le domaine LBD en accord avec la structure cristallographique (PDB : 3D24) et la barre bleue délimite le domaine DBD en accord avec la structure décrite dans la partie I de cette thèse. Les barres rouges indiquent les hélices alpha du récepteur et les flèches vertes indiquent les feuillets bêta. Les boîtes fonctionnelles du DBD sont annotées par des rectangles violets et les doigts de zinc sont indiqués en jaune.

En conclusion, PGC-1α change de façon importante la stabilité et la conformation de ERRα. Il est confirmé, au vu de ces résultats et de ceux de XL-MS, que la liaison du coactivateur affecte la conformation des DBDs et favorise vraisemblablement leur dimérisation, ce qui va dans le sens des résultats de SAXS et des mesures réalisées au FIDAbio. La région N-terminale est aussi affectée par la liaison de PGC-1α par l'exposition de résidus pouvant être modifiés post-traductionnellement pour inhiber l'activité transcriptionnelle associée au récepteur. Enfin, la liaison du coactivateur entraîne la protection et donc la stabilisation des boucles H1-H3 et H11-H12 des LBDs, ainsi que la déstabilisation de l'hélice H6. Ces résultats peuvent être statistiquement appréciés (**Fig. S.13**).

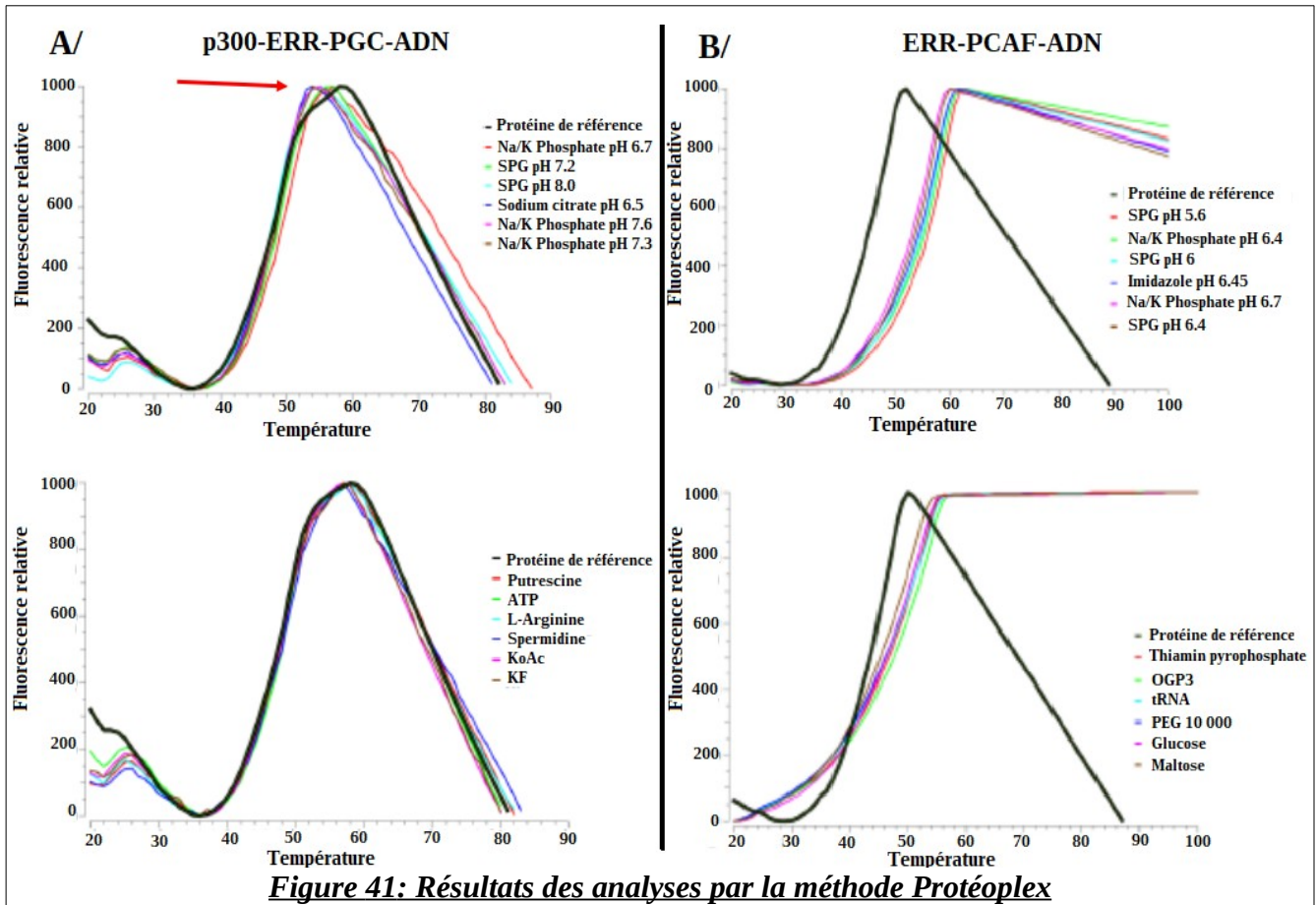
Ce résultat dispose de limites dans son interprétation. La première d'entre elles est qu'il n'est pas possible de suivre un changement d'incorporation du deutérium de PGC-1α. Ce coactivateur étant une protéine intrinsèquement désordonnée, l'incorporation du deutérium dans la condition de base et la condition complexée ne permet pas d'obtenir de différences significatives. De plus, les changements de protection et de déprotection observés ne rendent pas compte de l'existence de non-équivalences entre

les monomère de ERR α formant le dimère actif. En effet, PGC-1 α se liant une seule fois sur ERR α , il est possible que sa liaison affecte certaines régions de l'un des ERR uniquement.

4/ P300 et PCAF forment un complexe avec ERR α

Les résultats du criblage par la méthode Protéoplex effectuée par Novalix indiquent que les conditions propices à l'interaction entre p300 et ERR-PGC-ADN, mais aussi pour PCAF avec ERR-ADN, seraient en tampon NaKPO₄/NaKHPO₄ à un pH compris entre 5.6 et 7.6 (**Fig. 41**). Ces conditions ont été testées avec différents rapports molaires de chacune des protéines sans mener à la formation du complexe, du moins après analyse des chromatogrammes suivant les SEC, des expériences de SEC-MALLS et des différentes conditions de migration en gel natif testées.

La co-lyse de culots de cellules d'insectes ayant coproduit mouse_ERR α _1-422-PGC-1 α _1-472 d'une part et p300 d'autre part ont permis de mettre en évidence la présence potentielle d'un complexe après purification sur colonne d'affinité pour l'étiquette streptavidine (StrepTrapTM HP 5 mL, en tampon : 20 % glycérol, 200 mM NaCl, 20 mM Tris HCl pH 7.5, 0.2 mM MgCl₂, 2 mM CHAPS). Le complexe n'est pas observable après passage sur colonne de SEC, ni par SEC-MALLS. Il est donc difficile de conclure quant à son existence. Il est à la fois possible que les bonnes conditions n'aient pas été trouvées, mais aussi que l'interaction soit faible de par l'absence d'un partenaire nécessaire à la formation du complexe.



Légende :

A/ Courbes de dénaturation du complexe p300-h_{ERR}α₁₋₄₂₃-PGC-1α₁₋₂₈₅-wc29embedded. La flèche rouge indique le sommet unique des meilleures conditions.

B/ Courbes de dénaturation du complexe PCAF_{ERR}α_{FL}-wc29embedded. Dans ce cas, du fait de la molécule de marquage utilisée, la courbe ne retourne pas au niveau basal.

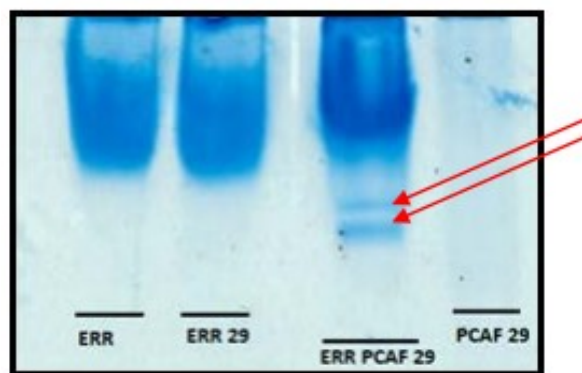


Figure 42: Migration sur gel natif du complexe human ERRα 1-423-PCAF-wc29embedded

Légende :

Gel de polyacrylamides 6 % Tris Borate pH 8.5 après migration en conditions natives du complexe human_{ERR}α₁₋₄₂₃-PCAF-ADN pendant 1 heure à 2 W constant. Les flèches rouges indiquent les bandes apparues suite à la formation du complexe.

L'utilisation de $ERR\alpha$ en présence d'ADN sans PGC-1 α a permis d'identifier un tampon (50 mM $KH_2PO_4/KHPO_4$ pH 7.0, 200 mM NaCl, 1 mM $MgCl_2$, 2 mM TCEP) où PCAF se lie (**Fig. 42**). Néanmoins, l'excès molaire de ERR est conséquent, dix fois plus que PCAF ici, ce qui pourrait traduire une interaction de très faible affinité. Cela peut s'expliquer par l'absence de partenaires permettant la stabilisation du complexe ou par l'action de PCAF qui serait transitoire.

5/ Effet de la nature de l'échantillon et du support sur le complexe

Dans le but de procéder à l'étude structurale du complexe $ERR\alpha$ -PGC-1 α , j'ai comparé les échantillons human_ $ERR\alpha$ _1-423-PGC-1 α _1-285-BE42embedded et mouse_ $ERR\alpha$ _1-422-PGC-1 α _1-472-BE42embedded après leur vitrification sur différents supports.

D'après les micrographes obtenues sur grilles Quantifoil, on constate que le complexe produit en cellules d'insectes tend à s'agréger lors de la vitrification au vitrobot. L'ajout de DDM à hauteur de 0,005 % permet d'éviter partiellement ce problème en permettant une meilleure répartition des particules. De plus, les protéines tendent à être très concentrées à l'extrémité des trous, proche du carbone. Enfin, les classes 2D révèlent un phénomène de dissociation du complexe. Par comparaison, le complexe formé par les protéines produites en bactéries est moins prône à l'agrégation sur grille et semble moins propice à se concentrer aux abords du carbone. Cependant, il semble davantage dissocié de l'ADN.

La nature des grilles impacte la qualité des images. L'utilisation de grilles d'or améliore la répartition des particules sur grilles, ainsi que l'homogénéité de l'épaisseur de glace. Par comparaison, les grilles commerciales Quantifoil R2/2 ne permettent pas d'augmenter visiblement la concentration en particules sur la grille et la glace y apparaît souvent plus épaisse et moins homogène (**Fig. 43-A et B**). A l'inverse, les complexes semblent davantage dissociés sur grille or (**Fig. 44-A et B**). Les grilles comportant une monocouche de cristaux de streptavidine permettent de concentrer l'échantillon plus aisément en capturant l'ADN biotinylé, mais le bruit de fond y est plus important et le signal du réseau cristallin empêche l'alignement correct des particules au vu de leur petite taille, et donc de leur faible signal en comparaison (**Fig. 43-C et Fig. 44-C**).

Enfin, les grilles réalisées au Chaméléon, comparativement à celles réalisées au vitrobot, montre une épaisseur de glace plus fine et plus homogène sur un même carreau (**Fig. 43-D**). De plus, les particules y apparaissent visiblement moins dissociées de l'ADN pour le complexe human_ $ERR\alpha$ _1-423-PGC-1 α _1-285-BE42embedded et aucune agrégation ne semble particulièrement visible. Cependant, le temps de vitrification y est critique et il est difficile de reproduire de façon fidèle ce type de résultat (**Fig. 44-D**).

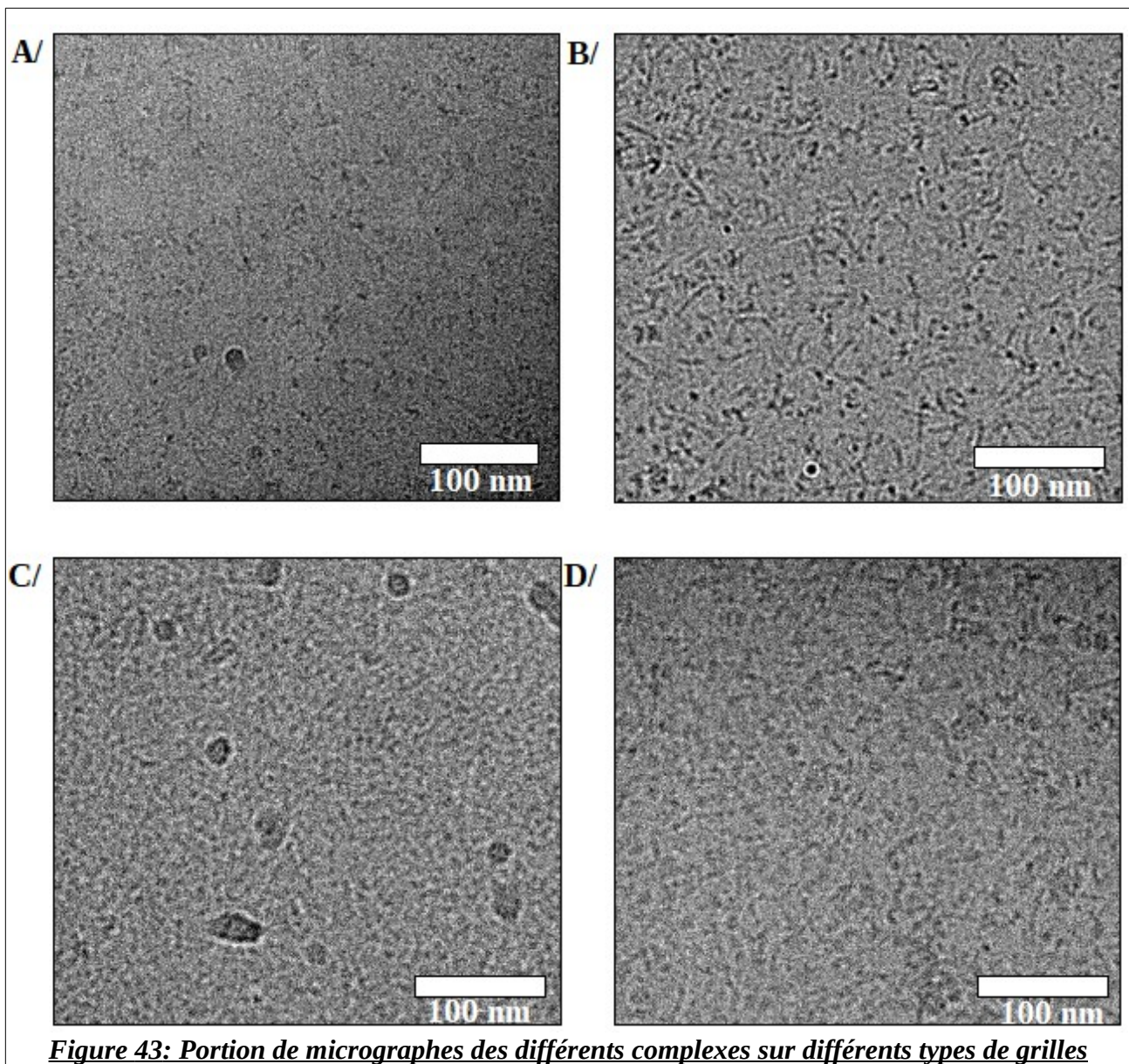


Figure 43: Portion de micrographes des différents complexes sur différents types de grilles

Légende :

A/ Complexes mouse *ERR α* _1-422-PGC1-472-BE42embedded sur grille de carbone sans support.

B/ Complexe human *ERR α* _1-423-PGC-1 α 1-285-BE42embedded sur grille or sans support.

C/ Complexe human *ERR α* _1-423-PGC-1 α 1-285-BE42embedded-bio sur grille disposant d'une monocouche de cristaux de streptavidines.

D/ Complexe human *ERR α* _1-423-PGC-1 α 1-285-BE42embedded sur grille carbone auto-abosrobante réalisée au Chaméléon.

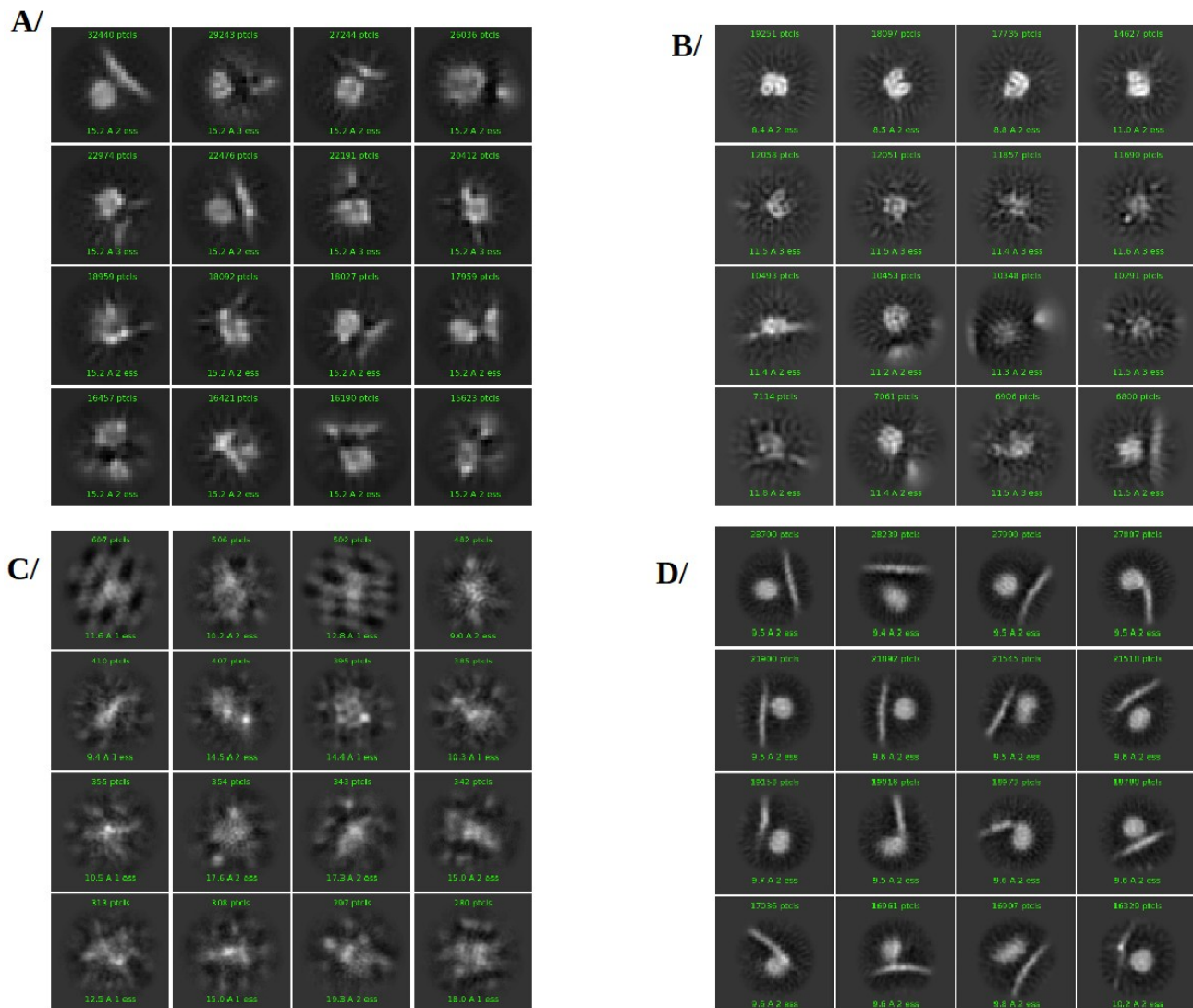


Figure 44: Classification 2D des complexes selon les jeux de données sur différents supports

16 classes 2D obtenue par l'utilisation du logiciel cryoSPARC et représentatives de chaque jeu collecté sont présentées ici.

Légende :

A/ Classes 2D du complexe mouse *ERRα*₁₋₄₂₂-PGC1-472-BE42embedded sur grille de carbone R2/2 en 256*256 pixels.

B/ Classes 2D du complexe human *ERRα*₁₋₄₂₃-PGC-1α₁₋₂₈₅-BE42embedded sur grille d'or en boites de sélection de 256*256 pixels.

C/ Classes 2D du complexe human *ERRα*₁₋₄₂₃-PGC-1α₁₋₂₈₅-BE42embedded-bio sur grille streptavidine après sélection manuelle des particules et extraction en boites de 256*256 pixels.

D/ Classes 2D du complexe human *ERRα*₁₋₄₂₃-PGC-1α₁₋₂₈₅-BE42embedded sur grille de carbone auto-absorbante réalisée au Chaméléon en boites de 300*300 pixels.

6/ Traitement des données du jeu mouse ERR α 1-422-PGC-1 α 1-472-BE42embedded

Le traitement du jeu de donnée a été effectué en parallèle dans les logiciels cryoSPARC et RELION. Le pré-processing a été effectué dans le logiciel Warp durant la collecte des données afin d'aligner les images composant les films acquis. L'estimation des paramètres de la CTF s'est effectuée avec CTFFIND4.1 et sur les 4776 micrographes, 4281 ont été conservés après inspection manuelle de chacun d'entre eux. La sélection des particules s'est faite selon trois méthodes différentes. La première consiste en une sélection manuelle des particules dans RELION sur quelques micrographes à différentes valeurs de défocalisation, suivie d'une classification 2D et d'une reconstruction *ab initio*. La reconstruction *ab initio* sert ensuite à générer une référence afin que le logiciel sélectionne l'ensemble des particules sur les micrographes. La seconde méthode consiste en une sélection entièrement automatisée et non biaisée par l'utilisateur dans cryoSPARC par l'intermédiaire du « blob picker », une tâche qui sélectionne les particules en fonction du contraste et de la taille des particules, ainsi que de l'espacement entre chaque particule défini par l'utilisateur. La dernière méthode consiste en une sélection par l'intermédiaire du logiciel crYOLO en entraînant une intelligence artificielle à la sélection des particules à partir de quelques particules sélectionnées sur des micrographes à des defocus différents.

L'ensemble des grandes étapes du traitement d'images est résumé dans la figure ci-dessous (**Fig. 45 et 46**). Chaque tâche a été répétée avec des paramètres différents afin de trouver les bonnes conditions pour obtenir un résultat de qualité. Pour donner un exemple, il faut en général optimiser entre 3 et 6 paramètres pour obtenir de bonnes classes 2D : tester s'il faut ou non prendre en compte le premier pic de la CTF (basses résolutions), s'il faut augmenter le facteur T ce qui change le poids mis sur les données expérimentales par rapport à la courbe de tendance des données, s'il faut changer la taille du masque pour prendre en compte plus de solvant et donc moyenniser plus efficacement le bruit de fond ou au contraire contraindre l'alignement de l'objet en réduisant la taille du masque considérant la difficulté d'analyse d'un objet de ce type (petite taille, dissociation partielle de l'ADN ou du coactivateur, faible contraste...), le nombre d'itérations en fonction de la capacité du logiciel à aligner facilement ou non les particules, ou encore le nombre de particules par itération, notamment lorsque ces dernières présentent un ratio signal sur bruit faible. Les tâches intermédiaires ou celles n'ayant pas conduit à de bonnes cartes 3D (trop d'artefacts, pas la bonne longueur de l'ADN, pas de connectivité entre les domaines...) ne sont donc pas représentées dans les descriptions des résultats.

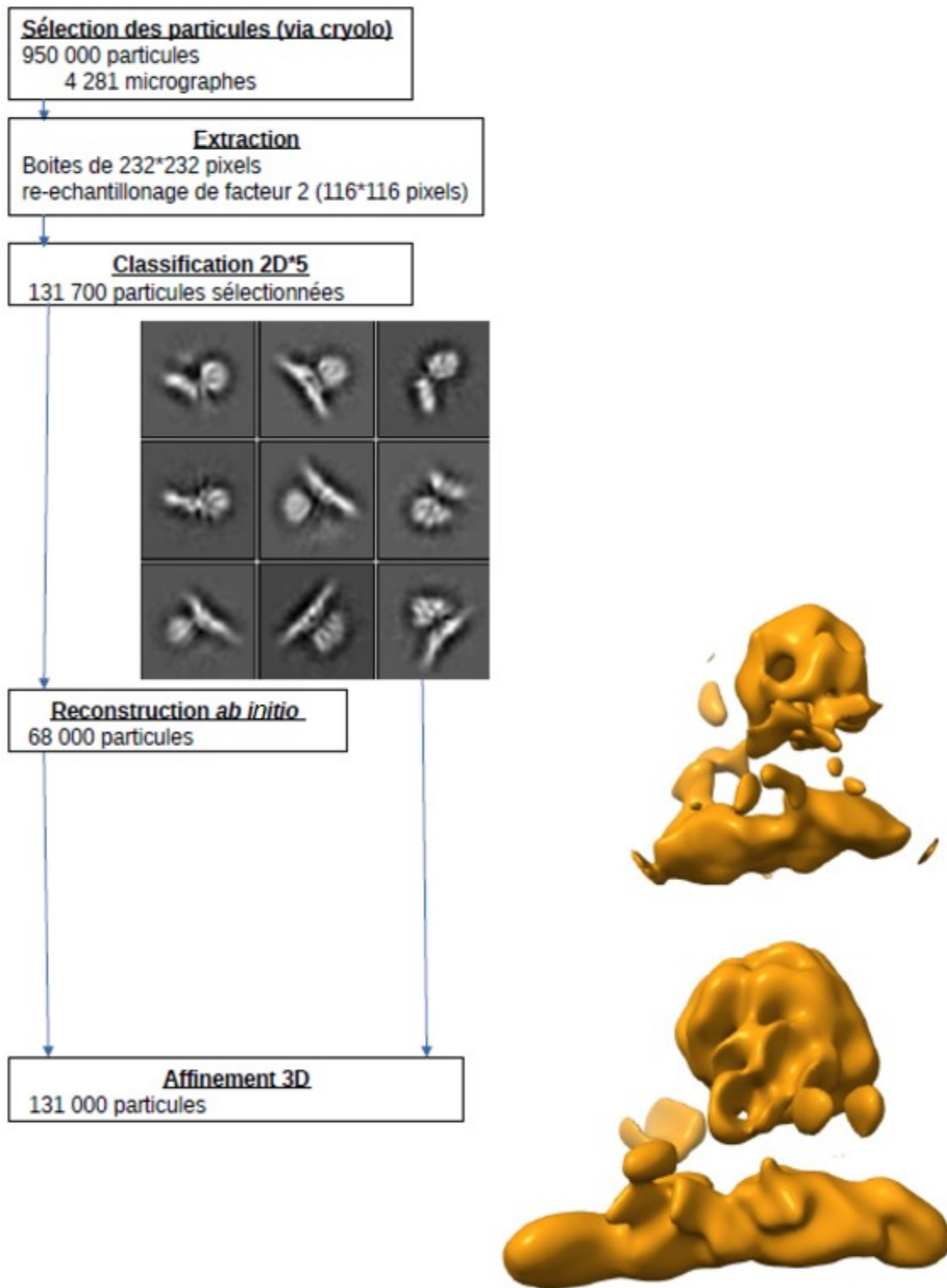


Figure 45: Figure du traitement de données dans RELION

Légende :

Figure récapitulative des grandes étapes du traitement d'images effectué dans RELION.

Les classes 2D obtenues par l'intermédiaire du logiciel RELION, dont quelques unes sont présentées sur la Figure 45, présentent une forme globale et une taille qui correspondent à celle de l'objet, malgré des traces de bruit de fond, d'artefacts en bordure de masques et d'un manque de connectivité continue entre les grandes masses de densités qui correspondent aux domaines du récepteur nucléaire (**Fig. 45**). La reconstruction *ab initio* à partir des particules sélectionnées contient beaucoup d'artefacts indiquant des difficultés à aligner les vues de l'objet. Des artefacts de bords sont visibles et causés par la proximité du masque choisi. Le choix d'un masque à 140 Å appliqué à l'objet lors de la reconstruction fut néanmoins nécessaire afin de permettre un alignement correct des particules et l'obtention d'une reconstruction *ab initio* correcte. Lors de l'affinement 3D, le masque a été étendu à 180 Å afin de faciliter l'alignement correct de l'objet et la prise en compte de la taille complète de l'ADN (140 Å). La carte présente un manque de densité continue entre les deux masses principales correspondant aux LBDs du récepteur nucléaire et à ses DBDs sur son élément de réponse. L'ensemble des étapes découlant de la sélection manuelle des particules et de la sélection à partir d'une référence produite dans RELION à partir d'une sélection manuelle ne sont pas montrée ici, ces sélections étant globalement moins bonnes que celles effectuées dans crYOLO.

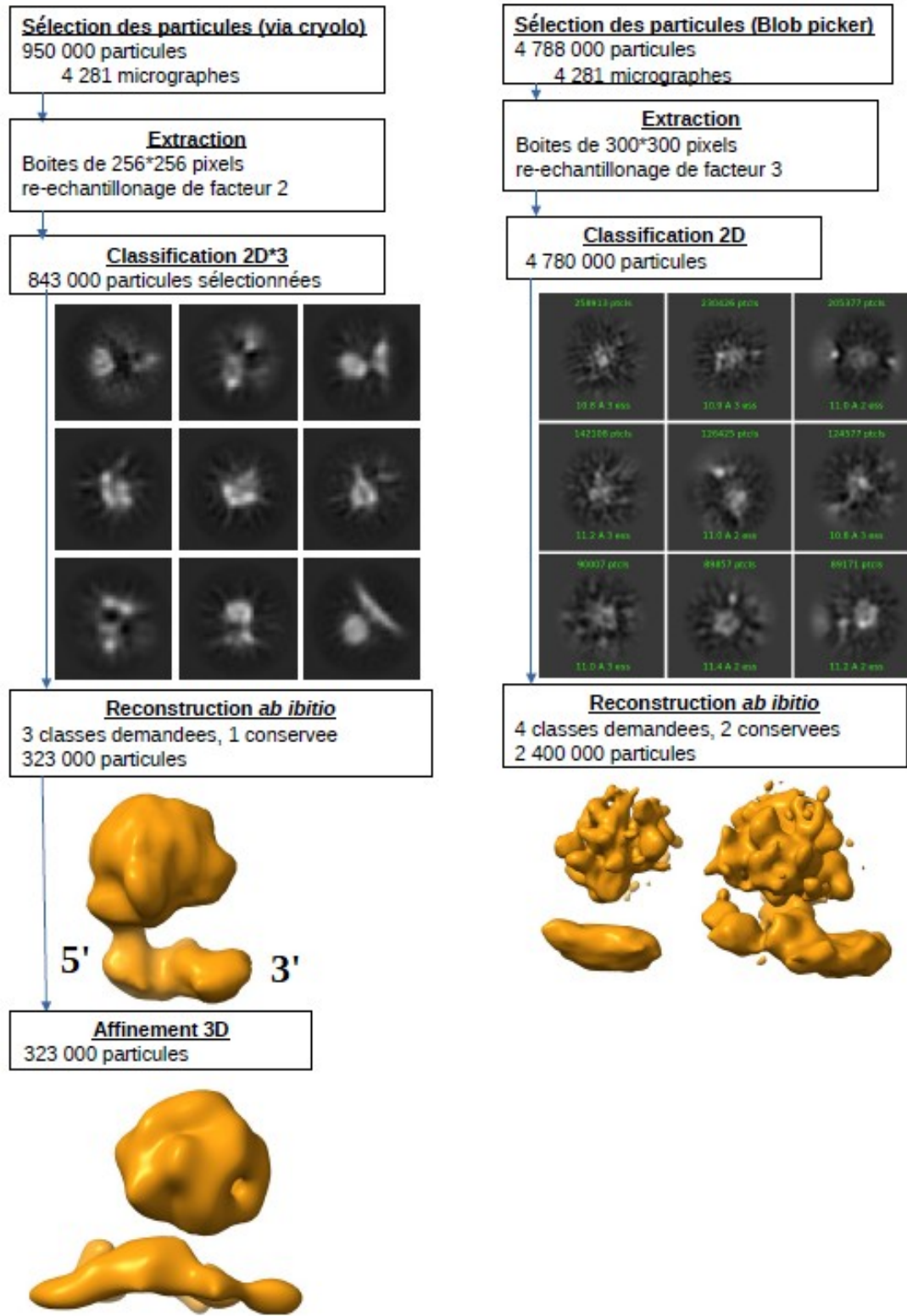


Figure 46: Figure Traitement des données dans le logiciel cryoSPARC

Figure récapitulative des grandes étapes du traitement d'images effectué dans le logiciel cryoSPARC.

Les classes 2D obtenues à l'aide du logiciel cryoSPARC à partir de la même sélection de particules effectuée à l'aide du logiciel crYOLO présentent, comparativement à celles obtenues dans RELION, moins de bruit de fond et d'artefacts (**Fig. 46**). Cela peut notamment s'expliquer par une meilleure qualité de l'alignement des particules, ainsi que par une meilleure moyennage du bruit de fond à taille de masque choisie identique. De plus, cryoSPARC dispose d'une option par défaut qui ne force pas la réduction du nombre de classes 2D obtenues en combinant « trop » de particules par classes, mais sépare strictement les particules en un nombre défini par l'utilisateur sans qu'aucune classe ne finisse sans particules. La reconstruction *ab initio* semble également plus qualitative puisqu'elle n'a pas autant d'artefacts que celle obtenue avec RELION. Il est également notable que cette dernière ait des contours plus uniforme et mieux définis, résultat de la filtration qu'opère cryoSPARC sur la carte finale par défaut, là où RELION fournit une carte brute non filtrée. Le tri des particules a notamment été réalisé par la reconstruction *ab initio* de plusieurs objets plutôt que sur une sélection de classes 2D, laquelle peut, dépendamment des cas, engendrer un biais lié à l'utilisateur. Cette façon de procéder fonctionne mieux sur le logiciel cryoSPARC puisque ce dernier est capable de ré-échantillonner à chaque tâche les images, lui permettant de travailler plus rapidement que RELION par lequel ce type de procédé demanderait plus d'étapes. Enfin, la sélection manuelle des particules demande une grande quantité de mémoire vive, chose que ne nécessite pas, ou du moins pas autant, la reconstruction *ab initio* pour laquelle les cartes sont téléchargeables séparément. En d'autres termes, il est impossible de sélectionner manuellement des classes 2D à partir d'un ensemble composé de plus de 3 millions de particules dans cryoSPARC. L'affinement 3D ne semble, pour sa part, pas fondamentalement meilleure que celui dans RELION au vu de la carte finale obtenue, si ce n'est dans la réduction de certains artefacts.

Ce jeu étant le premier à avoir chronologiquement été analysé, il fut aussi celui sur lequel j'ai passé le plus de temps afin de me former au traitement d'images en cryo-microscopie. Il m'a notamment permis de définir les paramètres par défaut de nombreuses tâches pour le traitement des jeux suivants. Il m'a permis de constater l'efficacité de crYOLO dans la sélection des particules lesquelles sont globalement mieux centrées qu'une sélection automatisée avec des outils basés sur le contraste des images. Ce fait s'observe notamment lors de la reconstruction *ab initio* (**Fig. 46**). De plus, j'ai pu constater l'avantage de cryoSPARC dans le tri des particules directement en 3D, ce qui limite l'introduction de biais par l'utilisateur liés à la sélection manuelle de classes 2D. Une sélection biaisée avec un fort a priori sur la forme de l'objet peut conduire à ne pas conserver certaines classes à cause de leur faible rapport signal sur bruit et la conservation par l'utilisateur d'une majorité de classes mieux définies. Ces classes peuvent être le fruit d'une orientation préférentielle de l'objet, engendrant des problèmes lors de la reconstruction en 3D à cause du manque de vues plus rares. Enfin, l'interface graphique de ce logiciel offre un confort d'utilisation et une rapidité dans l'analyse de la répartition des angles de vues des particules et dans l'estimation de la résolution des cartes obtenues. Si RELION me paraissait meilleure dans la production de classes 3D, l'ajout de cette tâche en version Bêta dans les dernières versions de cryoSPARC semble comparable.

7/ Traitement des données du jeu human ERR α 1-423-PGC-1 α 1-285-BE42embedded (Chameleon)

Le traitement des jeux de données du même complexe qu'auparavant, mais avec une construction plus courte de PGC-1 α , est effectué dans cryoSPARC après une sélection manuelle des particules dans crYOLO, l'entraînement d'une intelligence artificielle à partir des particules sélectionnées, puis l'application du modèle aux micrographes pour sélection.

L'ensemble des étapes importantes est résumé dans la figure suivante où les deux collectes de la grille réalisées grâce au système de vitrification Chameleon sont utilisées en exemple (**Fig. 47**).

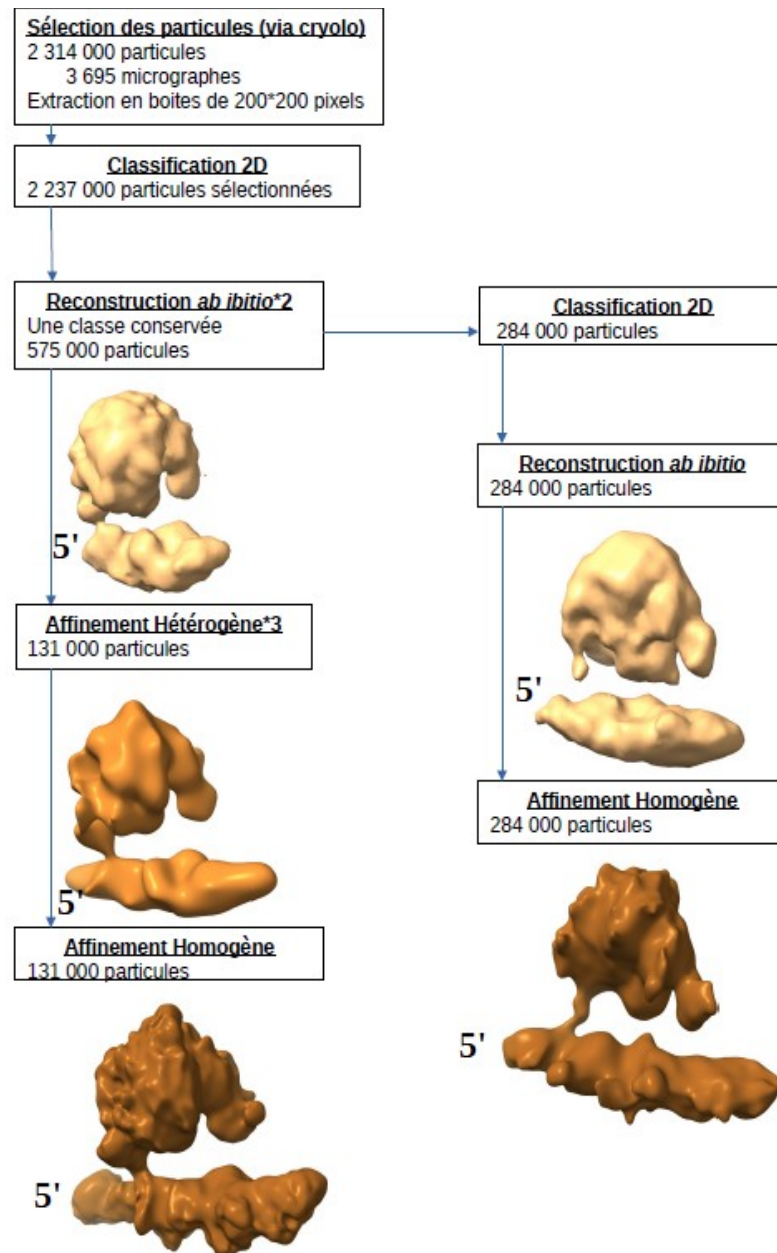
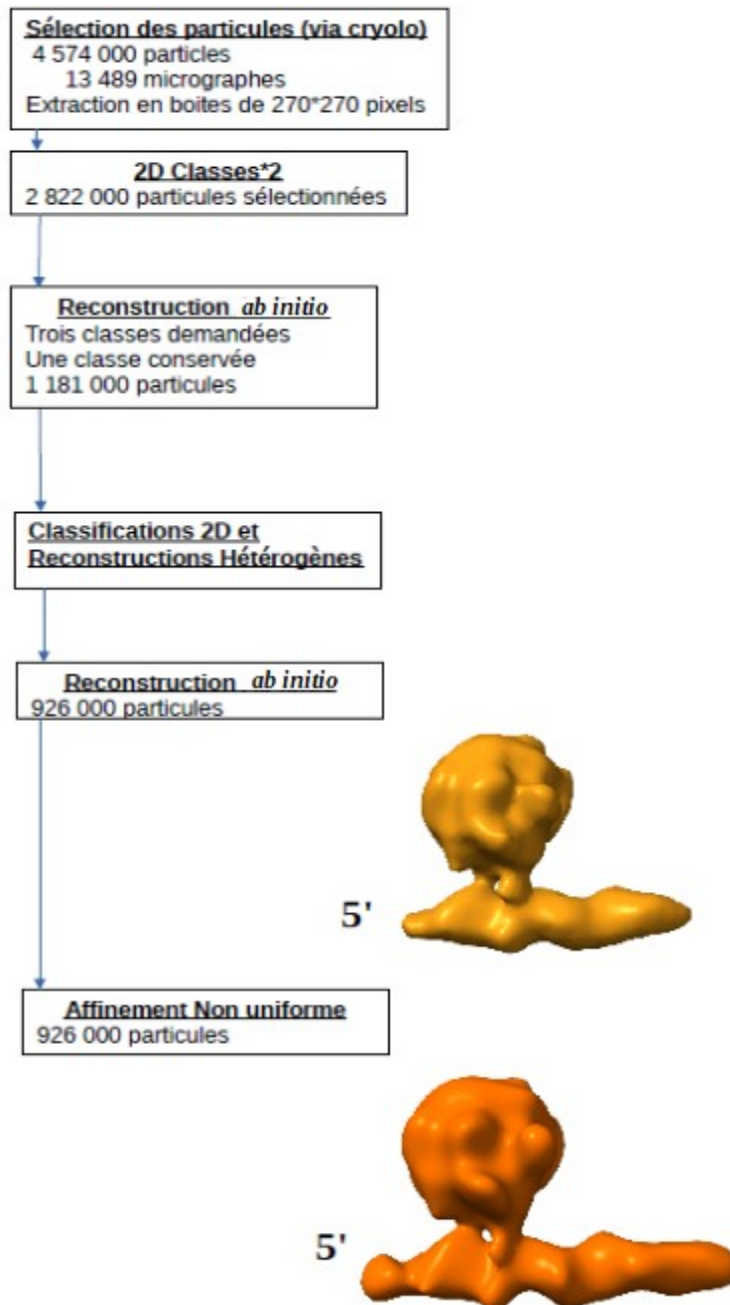


Figure 47: Principales étapes de la reconstruction de l'objet à partir de la première acquisition de la grille Chameleon

Sur plus de 2 millions d'objets sélectionnés au départ, seules 131 000 sont conservées durant le premier traitement. Il s'avère que certaines particules correspondant au complexe sont rejetées lors de l'affinement hétérogène car ces dernières présentent probablement divers types d'hétérogénéités : dissociation partielle de l'ADN, dissociation du coactivateur ou mauvais alignement. Une classification 2D permettra de récupérer un certain nombre de particules parmi celles précédemment rejetées sans permettre l'amélioration significative de la carte obtenue. Cette dernière présente un unique lien entre les domaines LBDs et l'ADN situé probablement en 5'. Il y a aussi la présence d'une extension de

voxels au niveau des LBDs. Chose intéressante : pour le même nombre, approximatif, de particules, la carte affinée de ce complexe produit en *E.coli* est bien plus compréhensible et détaillée (bien qu'il s'agisse encore de basse résolution) que celles obtenues à partir des complexes produits en cellules d'insectes. Cela peut être attribué d'une part à l'échantillon qui se trouve dans un meilleur état, mais aussi à l'épaisseur de la glace plus fine associée au système de vitrification chameleon qui permet un meilleur contraste et la vitesse de la vitrification qui réduit le temps d'interaction de l'échantillon avec l'interface air-eau.



La seconde acquisition contient bien plus de micrographes et, surtout, plus de particules du complexe utilisables. En effet, sur une sélection de 4 millions et demi de particules, 926 000 ont été conservées (**Fig. 48**). La carte obtenue est différente de celle du premier traitement et présente notamment deux liens reliant la région des LBDs et de l'ADN. Il est plus aisé de replacer la structure cristallographique des LBDs (PDB : 3D24) dans la carte, ainsi qu'un modèle de l'élément composite BE42embedded produit dans le logiciel coot. Cette carte a été obtenue en limitant volontairement la résolution de l'alignement à 9 Å de résolution dans cryoSPARC afin d'éviter la génération d'artefacts ou de bruit de haute résolution.

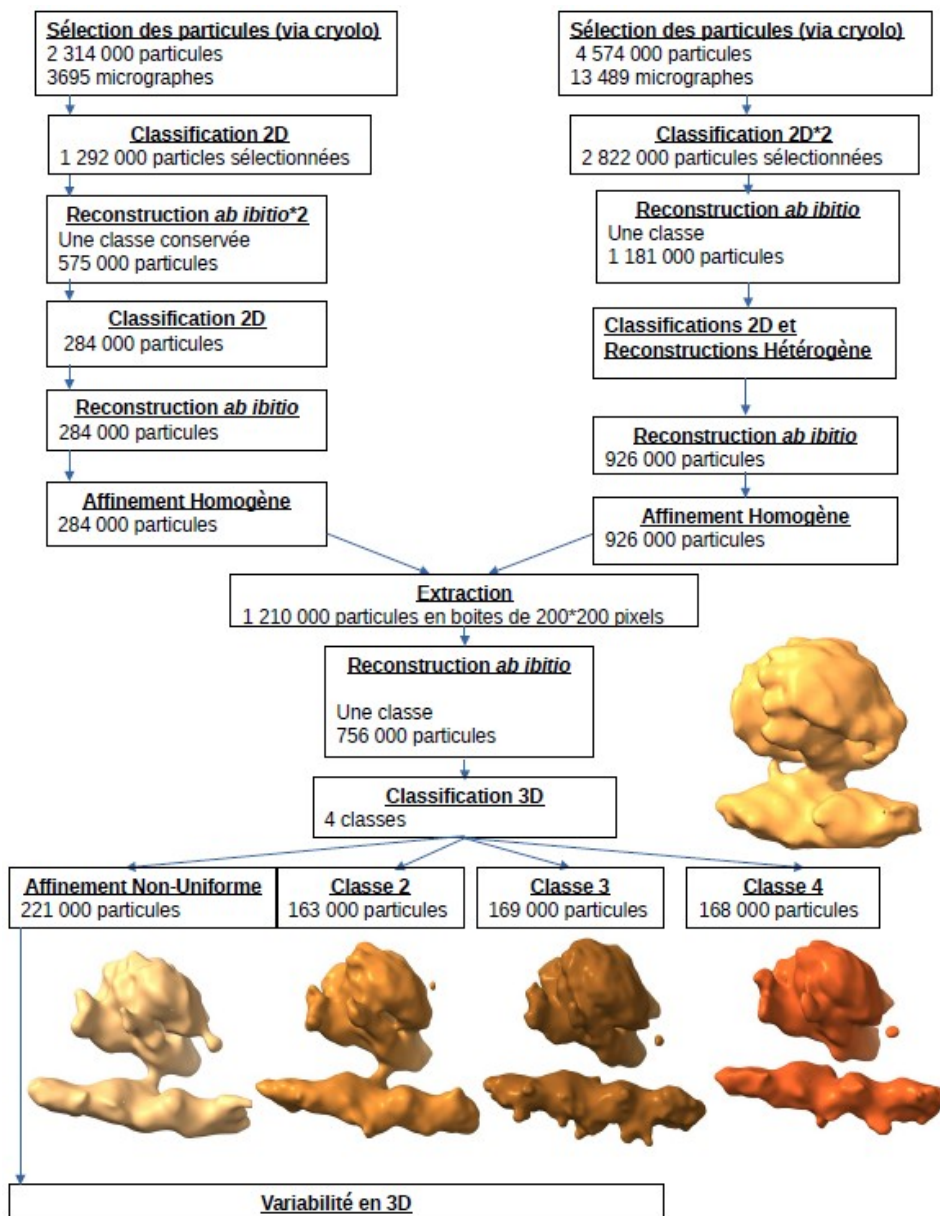


Figure 49: Principales étapes de la reconstruction de l'objet à partir des particules des deux acquisitions extraites en boîtes de 200 pixels

Dans l'optique d'analyser plus en détails l'objet, les deux acquisitions sont combinées et les particules sont extraites de nouveau à partir de leurs coordonnées affinées en boîtes de 200 pixels (**Fig. 49**). La reconstruction *ab initio* est suivie d'une classification 3D qui permet l'obtention de cartes potentiellement différentes. Ces classes se distinguent notamment par la variation de la distance entre la région associées à l'emplacement des LBDs et la région associée aux DBDs sur BE42embedded. Néanmoins, à ce niveau de résolution, les cartes peuvent être considérées comme globalement identiques. La pseudo-symétrie observée nuit probablement à une reconstruction correcte de l'objet en 3D. L'une de ces classes (en orange clair) est soumise à une analyse de sa variabilité 3D selon un algorithme basé sur l'analyse des composantes principales de l'objet.

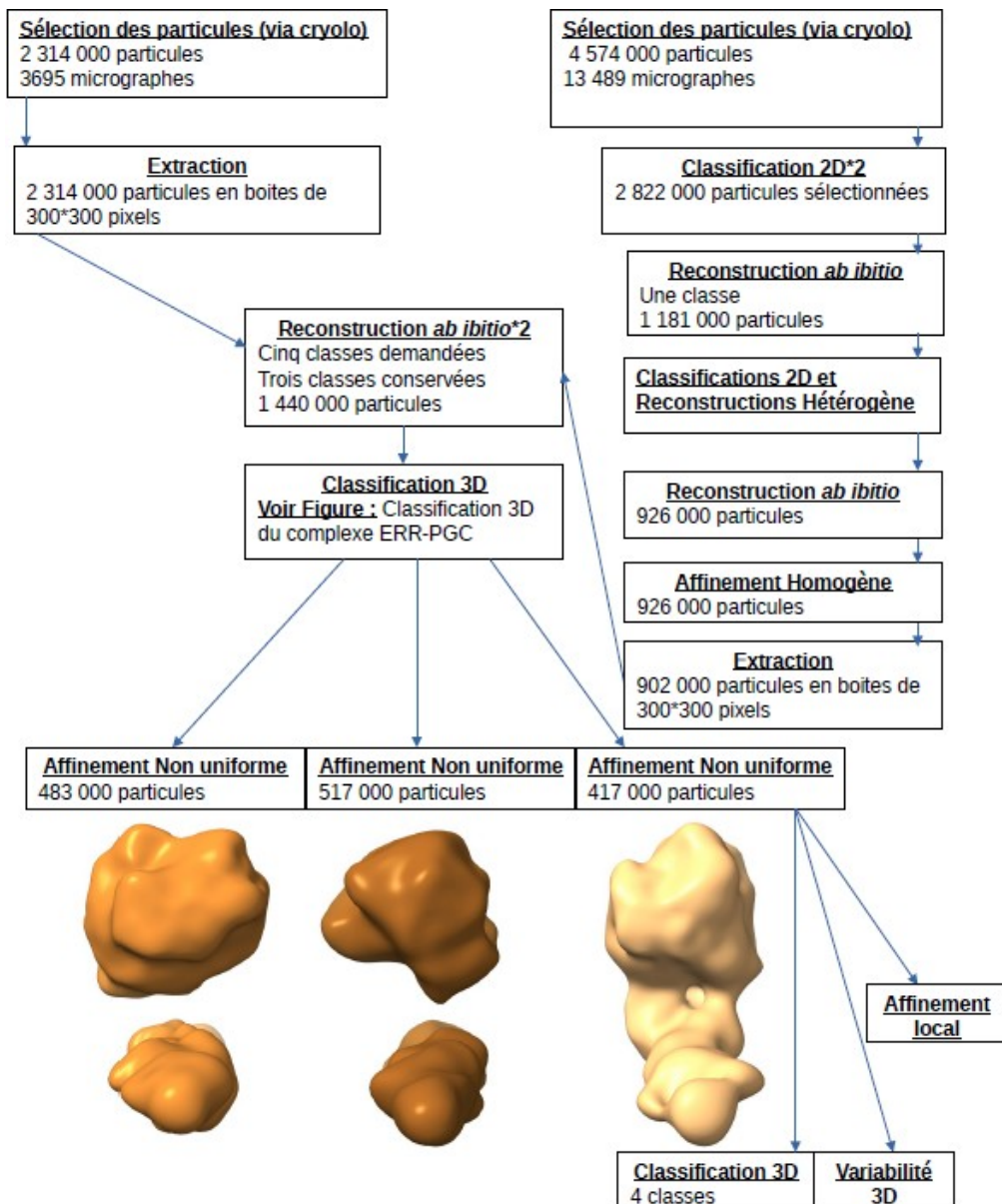


Figure 50: Principales étapes de la reconstruction de l'objet à partir des deux acquisitions

Les particules sont de nouveau extraites à partir d'une boîte plus grande (300*300 pixels) afin de permettre un théorique gain de signal par une meilleure moyennation du solvant, c'est-à-dire du bruit de fond (**Fig. 50**). Les particules sélectionnées au cours de la première acquisition par crYOLO sont cette fois triées directement après la combinaison pour ne pas introduire de biais quant à la surreprésentation d'une conformation. La reconstruction *ab initio* et la classification 3D qui s'en suit nous fournit trois classes distinctes et globalement différentes des classes de la précédente classification avec une taille de boîte plus petites (**Fig. 49**). Cette fois, l'une des classes présente une connectivité mieux définie entre les LBDs et l'ADN (en orange clair). Les deux autres classes obtenues présentent une région associée aux LBDs plus grosse (orange) ou plus petite (marron) et l'absence de densités continues entre les LBDs et l'ADN. La classe présentant la connectivité la plus marquée est ensuite analysée par classification 3D, affinement local de la région associée aux LBDs et analyse de sa variabilité en 3D de façon indépendante. Les résultats sont commentés dans la partie suivante.

8/ Identification des domaines individuels

Les cartes du complexe obtenues entre 10 et 14 Å de résolution montrent la topologie et l'organisation générale du récepteur sur l'élément de réponse composite et permettent de replacer les domaines isolés composant le récepteur nucléaire. Mises en perspective avec les classes 2D correspondantes, ces cartes dévoilent l'orientation des domaines les uns par rapport aux autres. La comparaison des différentes cartes issues du traitement d'images pour le complexe human_ERR α _1-423-PGC-1 α _1-285-BE42embedded montrent aussi les limites intrinsèques de l'alignement sur ce type d'objet.

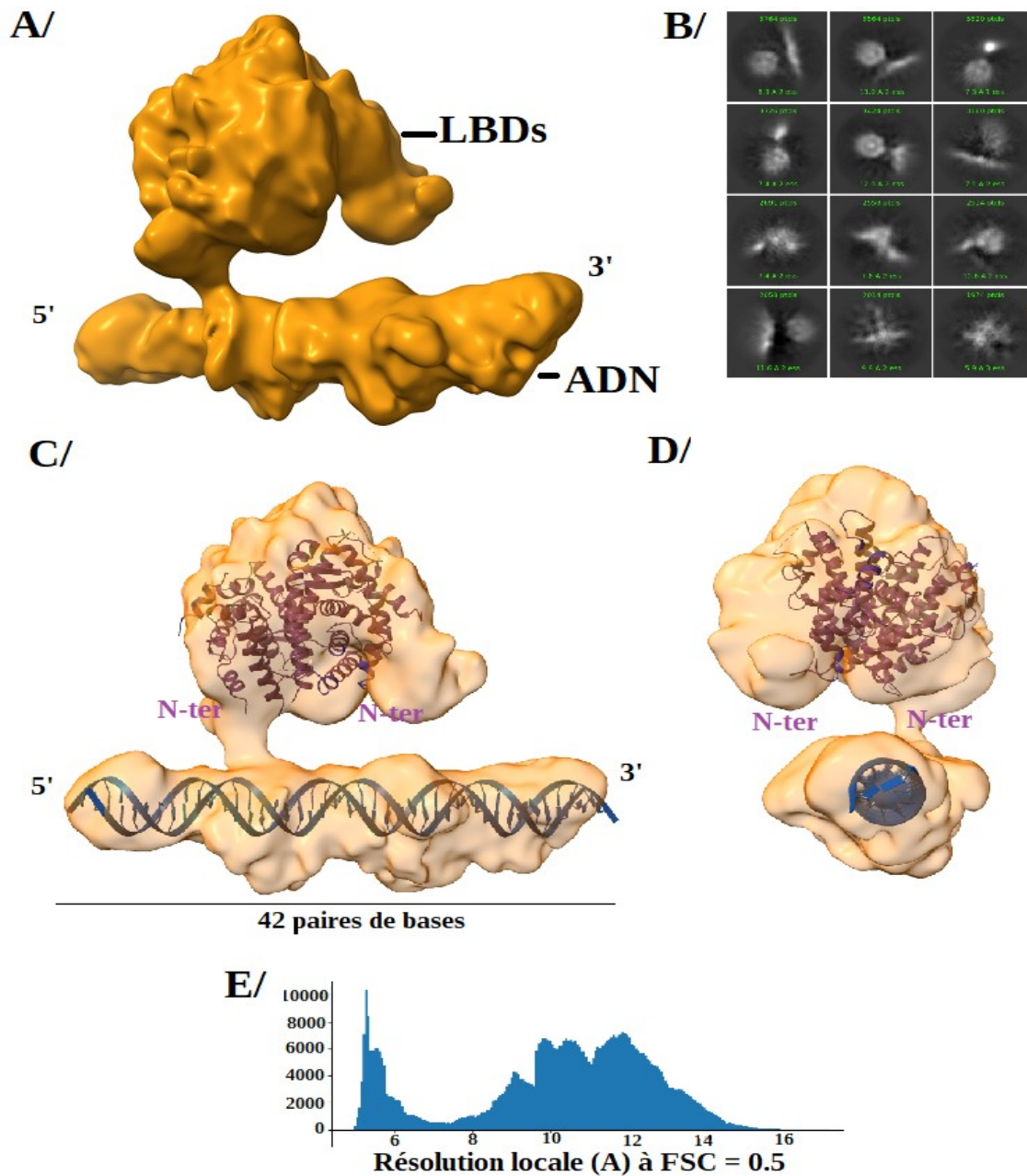


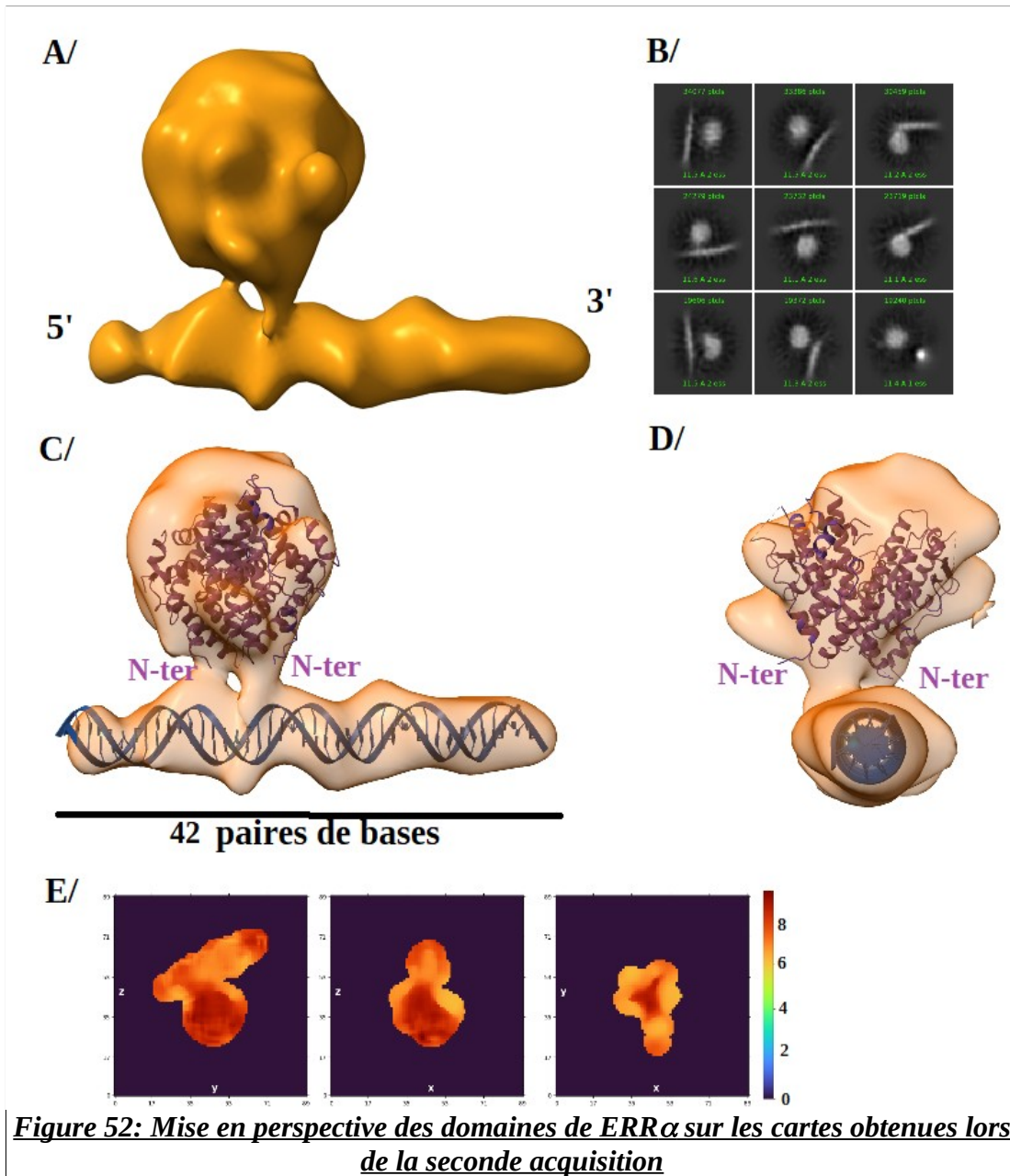
Figure 51: Mise en perspective des domaines de $ERR\alpha$ sur les cartes obtenues lors de la première acquisition

Légende :

A/ Carte affinée du complexe suivant le traitement de la première acquisition à partir de 131 000 particules. **B/** Quelques classes 2D associées à la carte affinée. **C/** Carte affinée du complexe dans laquelle sont positionnés les modèles atomiques des LBDs (PDB : 3D24, en violet) et de l'élément de réponse composite (en bleu). L'alignement des domaines a été réalisé manuellement dans le logiciel ChimeraX. L'extrémité N-ter de chacun des LBDs a été annotée. **D/** Carte affinée du complexe dans laquelle sont positionnées les modèles atomiques en vue depuis la position 3'. **E/** Répartition de la contribution de chaque image à la résolution maximale atteignable dans la carte obtenue. La majorité des images donnent une résolution comprise entre 10 et 14 Å. Le pic à 6 Å correspond au bruit de fond.

Sur la carte qui correspond au traitement de la première acquisition, une région en forme de cylindre correspond en taille à un ADN de 42 paires de bases, soit quatre tours de la double hélice que l'on peut placer (environ 140 Å de longueur) (**Fig. 51-A**). Il est possible de distinguer deux régions attribuées à l'ADN sur lesquelles la densité est plus forte, indiquant vraisemblablement la présence des DBDs. Considérant la séquence de l'ADN BE42embedded, l'élément composite reconnu par les DBDs se trouve entre les bases 10 et 27, soit entre les tours de double hélices numéros 2 et 3. La position des DBDs dans cette carte correspond donc à leur emplacement attendu sur la séquence, bien qu'il ne soit pas possible, à ce stade, de les placer précisément (**Fig. 51-C-D**).

Une région plus globulaire est attribuée aux LBDs. Néanmoins, les LBDs seuls ne suffisent pas à remplir et expliquer cette région de la carte de cryo-EM. En effet, un prolongement de celle-ci est observable. Cette région pourrait correspondre à la région coudée de l'un des LBDs ou à PGC-1 α . Procéder à davantage de tri des particules par classification 2D ou affinement hétérogène n'a pas permis d'améliorer davantage cette carte. Sa résolution est estimée par le logiciel cryoSPARC à 5.4 Å et son B-factor estimé est très élevé, valant -378.6 après affinement. La résolution réelle est entre 10 et 14 Å d'après l'estimation locale de la résolution (**Fig. 51-E**). Les détails de structures secondaires n'étant pas observables, mais les tours d'hélices du grand sillon de l'ADN pouvant être attribués, la résolution est donc plutôt autour de 10-12 Å. Les valeurs inférieures à 6 Å sont liées au bruit de fond, lequel est facilement pris en compte par cryoSPARC, ce logiciel utilisant des masques très proches des voxels en 3D conduisant à une sur-estimation de la résolution.



Légende :

A/ Carte affinée obtenue à partir du traitement de la seconde acquisition sur la base de 925 000 particules du complexe. **B/** Quelques classes 2D associées à la carte. **C et D/** Positionnement des modèles atomiques des LBDs (PDB : 3D24, en violet) et de BE42embedded (en bleu) au sein de la carte obtenue. Leur positionnement a été calculé avec l'option « fit to volume » de ChimeraX et corrigé manuellement à la main. **E/** L'estimation de la résolution locale calculée par cryoSPARC à partir des « demi-cartes » montre des valeurs de résolution autour de 7-10 Å.

Sur la carte de la seconde acquisition, comportant huit fois plus de particules qui ont été ré-échantillonnées d'un facteur 3, il est possible de replacer les structures des domaines individuels cristallographiques (**Fig. 52-A, C et D**). La topologie globale demeure la même que celle de la précédente carte.

Tout d'abord, les DBDs ne sont pas identifiables sur l'ADN, lequel présente une longueur de 42 paires de bases par comparaison au modèle produit dans coot. Cela peut-être le résultat du ré-échantillonnage qui diminue le nombre de pixels pouvant décrire l'emplacement des DBDs. Concernant les LBDs, la densité observée présente cette fois deux connectivités distinctes avec l'élément de réponse composite. Ces dernières sont cependant très proches l'un de l'autre. L'orientation des LBDs n'est pas clairement définie et la région correspondant aux LBDs à une orientation non parallèle à l'axe de l'ADN et présente un axe de pseudo-symétrie qui nuit probablement au bon alignement des particules. PGC-1 α est difficile à positionner dans cette carte, pour des raisons probablement similaires à celles qui conditionnent le positionnement des DBDs. La résolution de cette carte est estimée à 6.71 Å dans le logiciel cryoSPARC avec un B-factor très important de -550 Å². Ce dernier rend probablement compte des difficultés à aligner les particules au sein du jeu de données, de la difficulté à obtenir une bonne carte 3D initiale pour ce type de complexe et de la surestimation par le logiciel de la résolution réelle de la carte. Ce sont ces différents éléments qui pousseront à la réalisation de classification 3D du complexe.

9/ Amélioration des cartes du complexe ERR α -PGC-1 α

La combinaison des deux jeux de données en extrayant les particules dans des boîtes de 200 pixels permet la production de cartes disposant de beaucoup plus de particules, ce qui permet de les classer sans amoindrir de façon conséquente le rapport signal sur le bruit. Il est possible de classifier en 3D les particules afin d'extraire des conformations différentes de l'objet, mais aussi d'extraire l'hétérogénéité sous forme de films montrant la transition conformationnelle entre les états intermédiaires discrets du complexe grâce à l'opération « 3D variability ». L'analyse de la variabilité en 3D, ainsi que la classification 3D par une analyse des composantes principales, réalisées dans le logiciel cryoSPARC permettent d'identifier au sein des particules des changements conformationnels (**Fig. 53**).

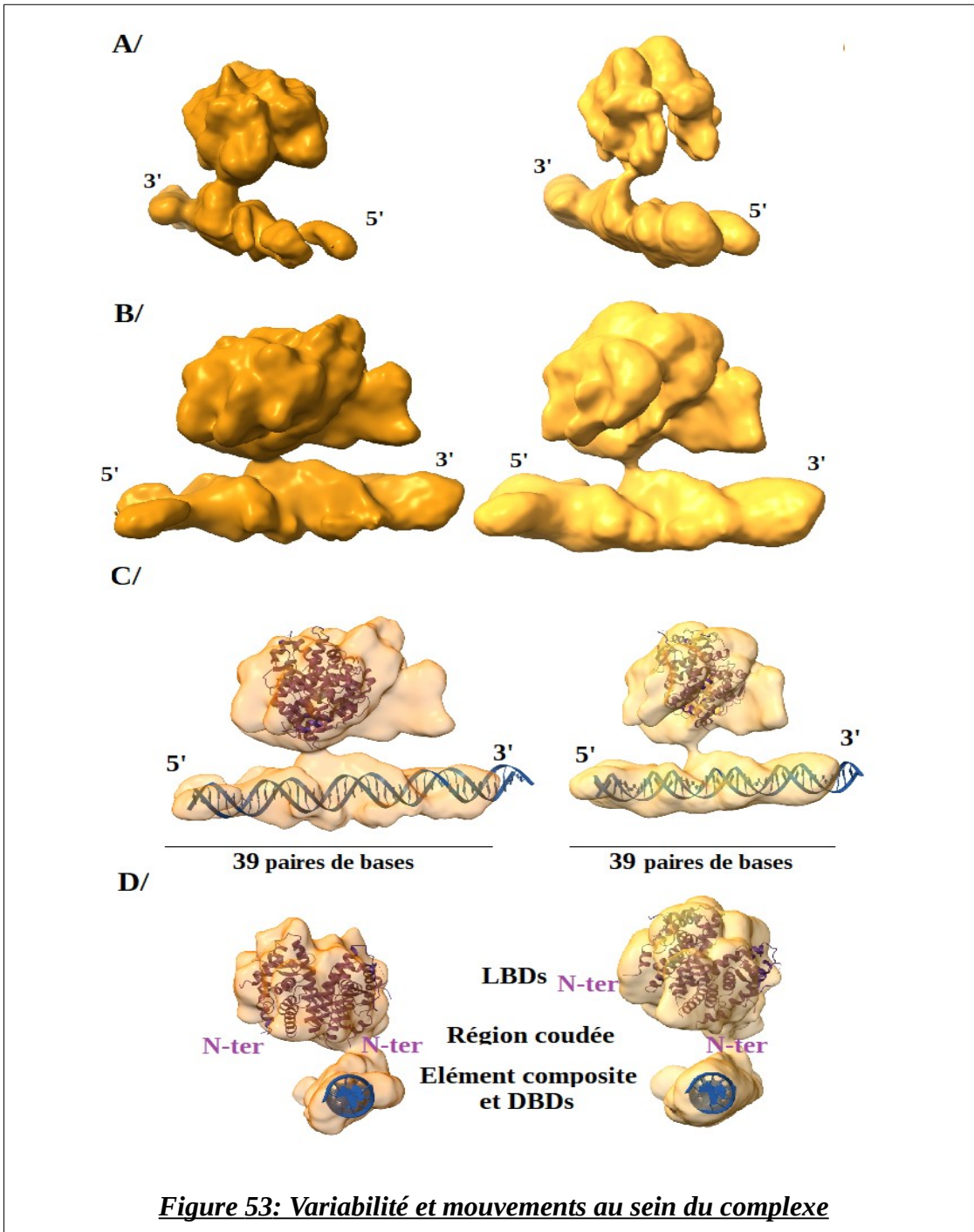


Figure 53: Variabilité et mouvements au sein du complexe

Légende :

A/ Deux cartes du complexe en 200*200*200 voxels issues de l'analyse par la tache « 3D variability » par le logiciel cryoSPARC à partir de 220 000 particules en vue 3'. **B/** Cartes du complexe en vue 5'-3'. **C/** Cartes en vue 5'-3' avec les LBDs (PDB : 3D24, en violet) et l'élément de réponse composite (en bleu) placés dans la carte à la main via le logiciel ChimeraX. Le nombre de bases entrant dans la carte a été annoté. **D/** Cartes en vue 3' avec les LBDs (PDB : 3D24, en violet) et l'élément de réponse composite placés à la main dans le logiciel ChimeraX. Les résidus en position N-terminal des LBDs ont été annotés.

Les cartes obtenues par l'analyse de la variabilité en 3D permettent d'apprécier 3 des 4 tours du grand sillon de l'ADN. La longueur mesurée correspond à 39 des 42 paires de bases de l'ADN BE42embedded, la région correspondant à l'extrémité 3' étant moins bien définie. L'hélicité de l'ADN semble indiquer que ces cartes présentent la bonne chiralité (**Fig. 53-A et B**). Les régions attribuées aux LBDs ne sont pas équivalentes dans les deux cartes qui présentent deux différences principales : une orientation légèrement différente de la région attribuée aux LBDs et une évolution de la distance entre l'ADN et les LBDs. Ces différences peuvent être le fruit d'une mauvaise attribution angulaire des particules à cette résolution ou d'un léger mouvement des LBDs.

Au vue de la protubérance de la région dans laquelle a été placée la structure cristallographique de la région LBDs, il est possible que le coactivateur soit présent. Cette protubérance peut également être le fait d'autres facteurs tels que la difficulté à aligner des particules dans des boîtes de 200 pixels, soit 180.2 Å par 180.2 Å. Cela peut aussi être causé par un mauvais alignement des particules comme semble le suggérer la déformation dans une direction des cartes, laquelle semble comme « étirée » dans l'axe 5'-3'.

Finalement, la classification 3D réalisée dans le logiciel cryoSPARC par une analyse par composante principal sur 1 440 000 particules extraites dans des boîtes de 300 par 300 pixels permet d'identifier trois classes distinctes du complexe sur l'ADN (**Fig. 54-A à C**). Les trois classes présentent un positionnement différent des LBDs par rapport à l'ADN et aux DBDs. La région associée à l'ADN BE42embedded fait entre 39 et 42 paires de bases et les DBDs ne sont pas visibles dans les cartes. Les différences de formes et de tailles de la région attribuée aux LBDs, ainsi que la différence d'orientations des domaines cristallographiques au sein de la carte de cryo-EM suggère que les LBDs disposent d'une mobilité dans l'espace, qui reste néanmoins à confirmer sur des cartes à plus haute résolution. La classe présentant une connectivité avec l'ADN est la seule où la chiralité ne peut pas être attribuée en se basant sur l'hélicité de l'ADN. Les deux autres cartes semblent présenter une double-hélice droite, ce qui signifie que ces cartes auraient la bonne chiralité.

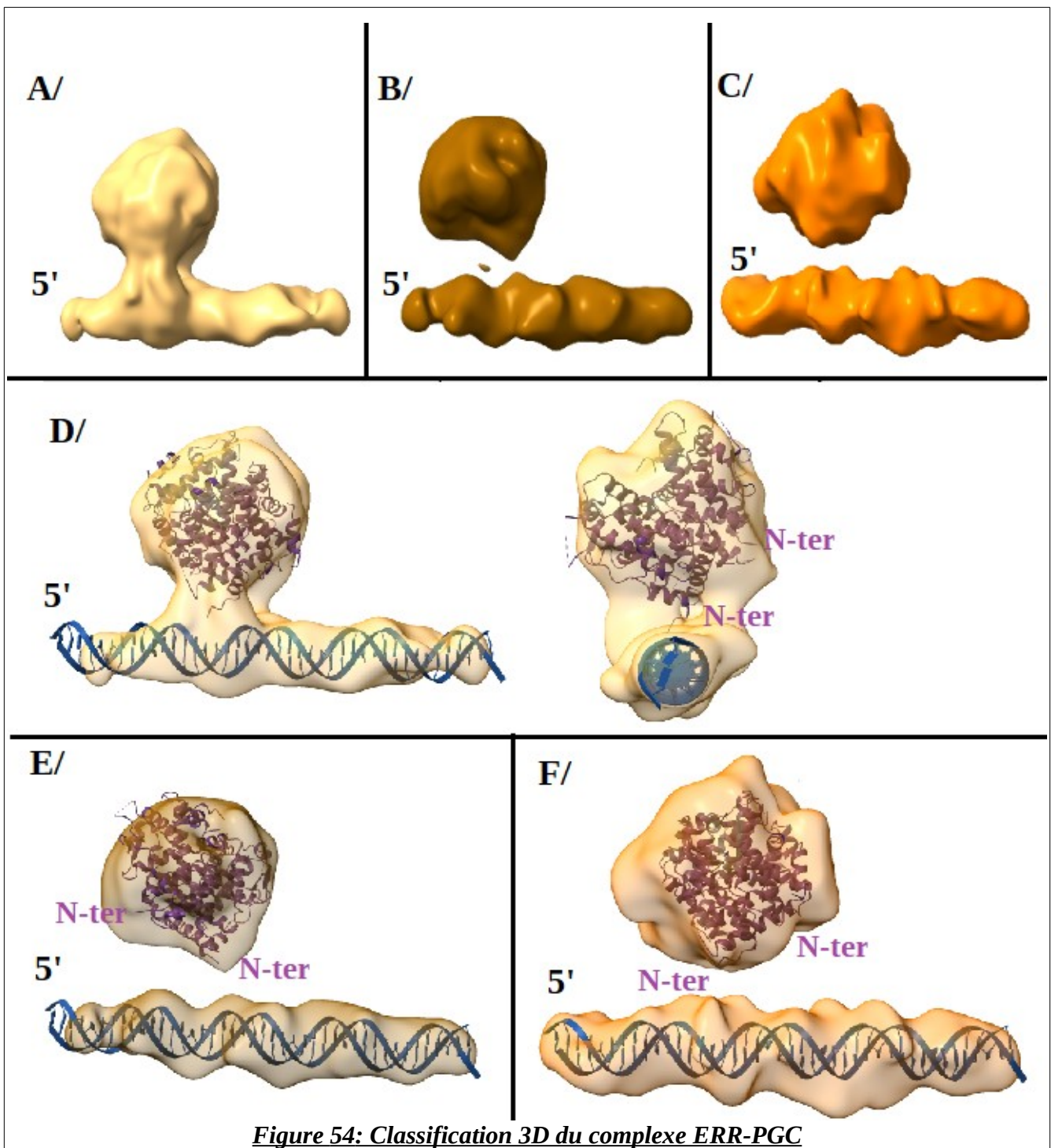


Figure 54: Classification 3D du complexe ERR-PGC

Légende :

A partir d'une sélection de 1 440 000 particules extraites en boîtes de 300*300 pixels, trois classes 3D sont identifiées. La première (A) comporte 417 000 particules. La seconde (B) comporte 539 000 particules. La dernière classe (C) comporte 483 000 particules. Les LBDs cristallographiques (PDB : 3D24) et le modèle atomique de l'élément composite ont été placés à la main dans ChimeraX sur chacune de ces cartes (D à F).

L'étude de la première classe composée de 417 000 particules par classification 3D, affinement local de la région LBD et analyse de la variabilité en 3D a été réalisée en priorité, cette dernière présentant une connectivité entre les LBDs et l'élément de réponse qui peuvent rendre son affinement plus aisé (Fig. 55).

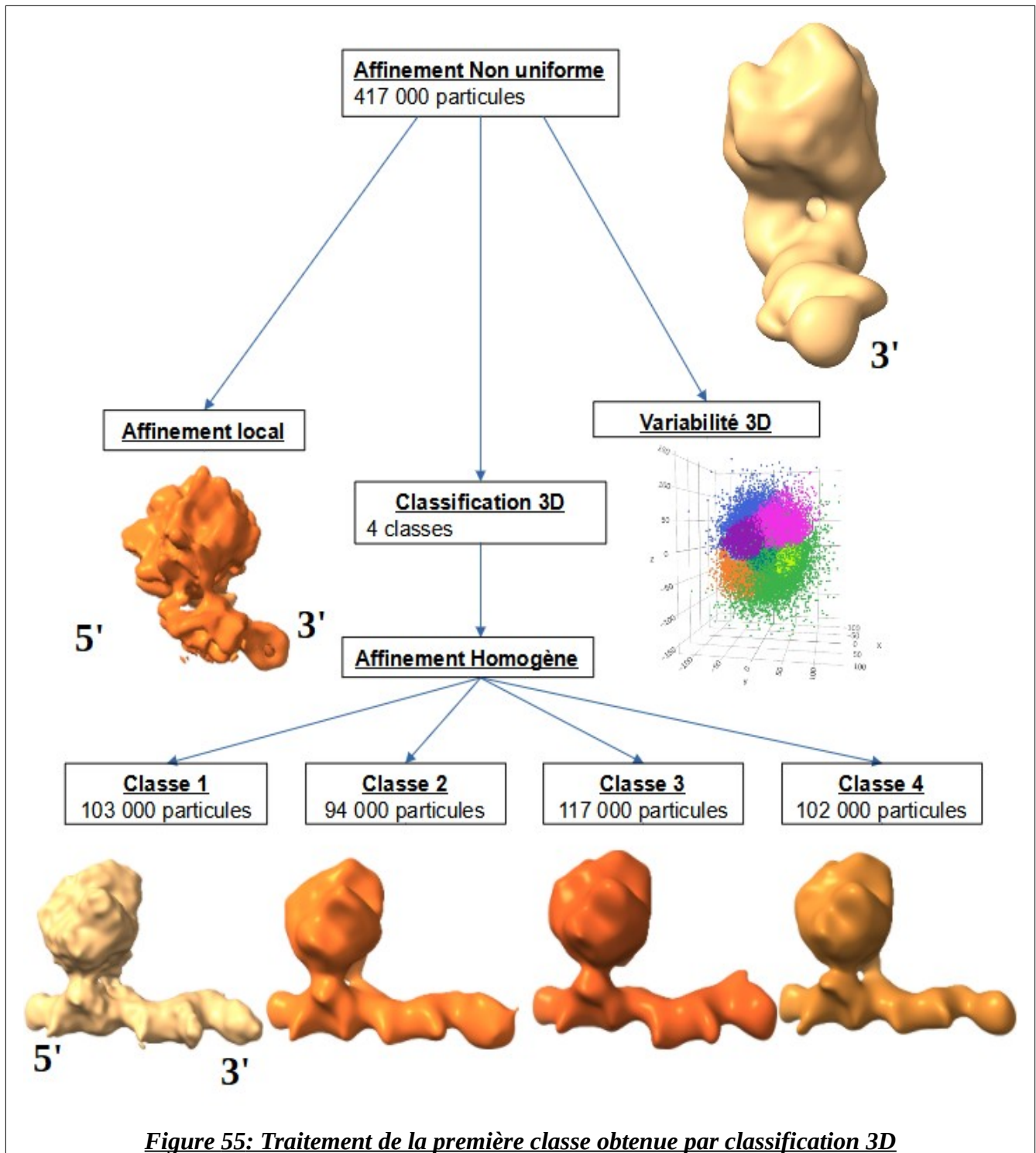


Figure 55: Traitement de la première classe obtenue par classification 3D

La première carte issue de la précédente classification 3D (**Fig. 54-A**) a été analysée de différentes manières pour essayer d'en améliorer la résolution ainsi que de comprendre l'architecture du complexe. La classification 3D révèle que les classes 1, 2 et 4 présentent deux liens avec l'ADN. La classe 3 ne présente qu'une seule connectivité observable qui est celle située en 5'. Il est cependant difficile de conclure sur la disparition de la connectivité en 3' du fait de la résolution et du nombre similaire de particules dans chacune des classes (environ 100 000 chacune).

L'analyse de la variabilité en 3D permet d'apprécier un mode où la distribution des particules s'étale sans pour autant former plusieurs clusters bien définis. Il n'y a donc pas, d'après cette analyse, de conformations différentes et fixes, mais probablement un ensemble de conformations discrètes très proches les unes des autres. Cela peut à la fois rendre compte d'une hétérogénéité réelle liée aux régions flexibles des protéines, mais aussi aux limites de l'alignement des particules qui expliqueraient cette distribution des particules.

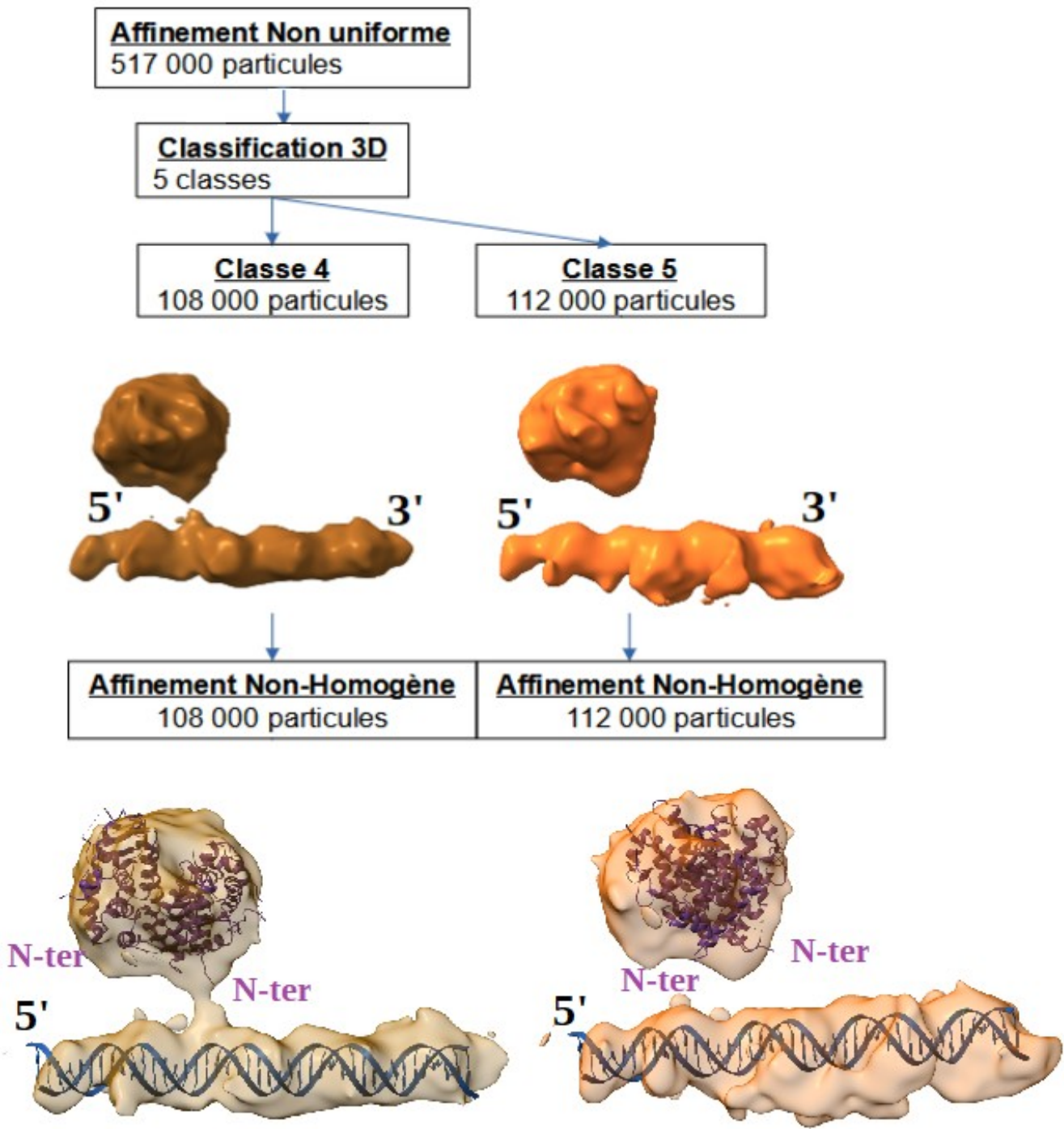


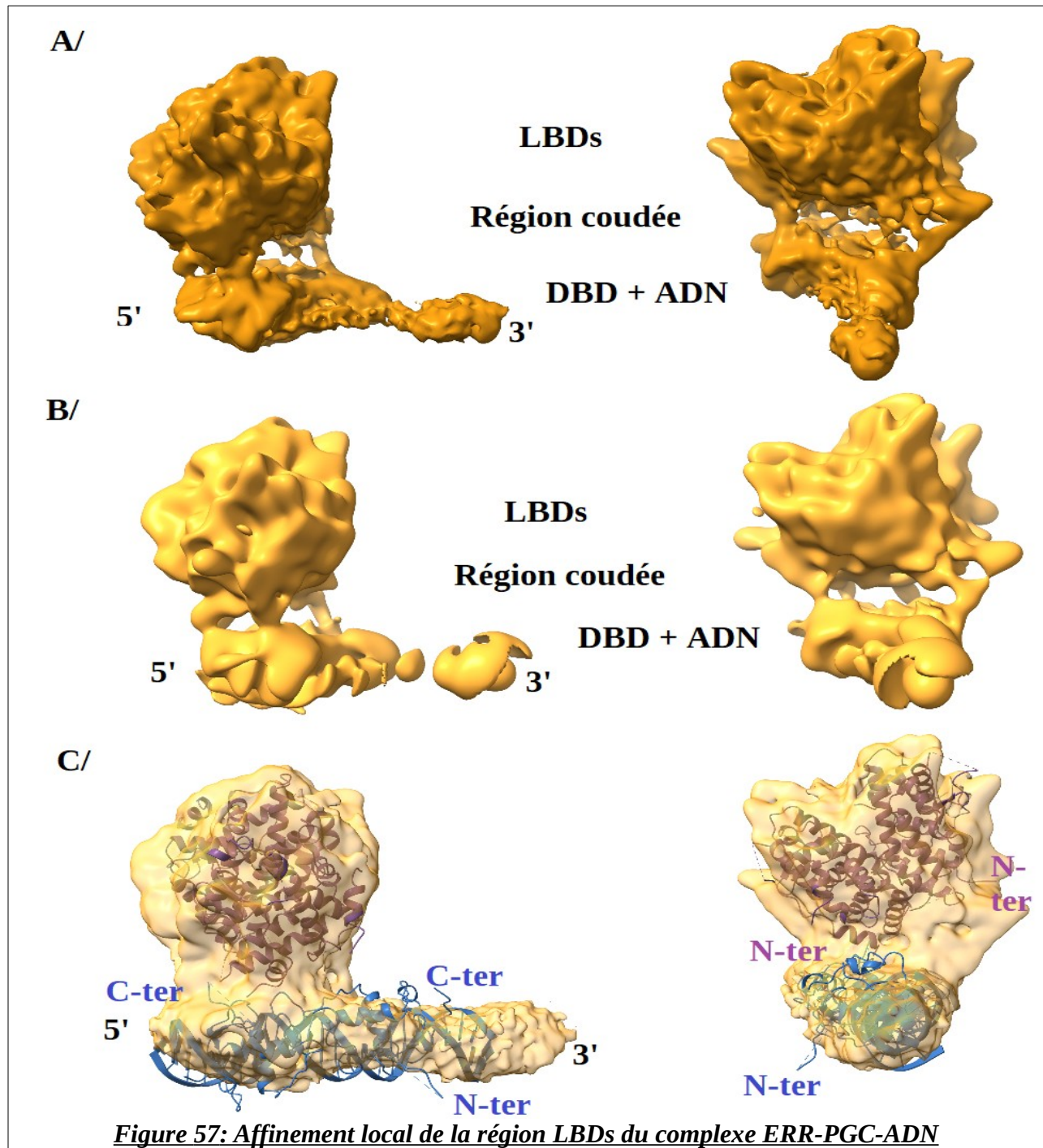
Figure 56: Traitement de la seconde classe obtenue par classification 3D

Après un affinement non-uniforme, les particules constituant la seconde classe 3D (**Fig. 54-B**) sont de nouveau triées en 3D afin de voir si certaines sous-conformations peuvent apparaître (**Fig. 56**). Les 5 classes obtenues contiennent toutes autour de 100 000 particules et sont globalement similaires à ce niveau de résolution. Parmi les cartes reconstruites, de légères différences sont observables : la classe 4 présente un début de connectivité et la classe 5 présente des protubérances sur l'ADN qui peuvent correspondre aux DBDs. L'affinement non-homogène confirmera la connectivité observée pour la classe 4 et la présence d'une région pouvant correspondre aux DBDs sur la classe 5.

10/ Description de la topologie du complexe

L'affinement local permet, à l'aide d'un masque, de délimiter une zone de la carte que l'on désire affiner. Ici, la région LBDs a été affinée localement à l'aide d'un masque comprenant la région attribuée aux LBDs et une partie de la région comportant l'ADN. La carte servant de référence est la carte sortant de l'affinement non-uniforme suivant la première classification 3D (**Fig. 54-A**). La chiralité de la carte demeure cependant à vérifier. Nous détaillerons donc ici l'interprétation des deux cartes possibles.

La carte telle qu'obtenue après l'affinement local (**Fig. 57-A**) dispose d'un B-factor de -150 \AA^2 . La détermination de la résolution en sortie de tâche est estimée par le logiciel à environ 3 \AA , ce qui est ici une surestimation. Cette dernière provient soit de l'affinement du bruit de fond ou d'artefacts de bordure de masques, soit d'une non détérioration de la qualité de la carte lors des étapes d'affinement à haute résolution conduisant le logiciel à estimer que la résolution est meilleure que la résolution réelle de la carte obtenue. La carte filtrée à 8 \AA afin d'éliminer le bruit des hautes fréquences ne permet cependant pas d'observer de différences majeures par rapport à la carte filtrée à 3 \AA , montrant que la carte dispose de peu ou pas d'artefacts pour les niveaux de contours considérés et que la résolution n'est probablement pas meilleure que 8 \AA (**Fig. 57-B**). L'axe de pseudo-symétrie des LBDs peut être apprécié, notamment après avoir placé la structure cristallographique (PDB : 3D24) au sein de la carte par l'utilisation de l'option « fit to map » du logiciel ChimeraX, puis après en avoir corrigé le placement à la main (**Fig. 57-C**). Enfin, bien que cette région ne soit pas comprise initialement dans la région à affiner localement, la reconstruction partielle de l'ADN permet d'apprécier la position relative des DBDs sur l'élément de réponse par comparaison avec la structure cristallographique présente dans la première partie de cette thèse.



Légende :

A/ Carte résultant de l'affinement local à un niveau de contour de 0,15 en vue « de côté » (à gauche) et en vue de 3' (à droite). **B/** Carte résultant de l'affinement local filtrée à 8 Å à un niveau de contour de 0,2. **C/** Carte résultant de l'affinement local à un niveau de contour de 0,065 dans laquelle ont été placées les structures cristallographiques du dimère de LBDs (PDB : 3D24) et des DBDs de ERR α sur l'élément de réponse composite décrite dans la partie I de cette thèse.

La carte permet de replacer la structure cristallographique des LBDs et de constater que le modèle correspond en de nombreux points à la carte obtenue. Tout d'abord, le repliement central comprenant les hélices H3, H4, H5, H7, H8, H9, H10, H11 et H12 des LBDs se trouvent entièrement dans la carte et correspondent aux densités observées. De plus, le LBD situé en 5' est plus proche de l'ADN et des DBDs que celui situé en 3', ce qui rejoint les observations déjà effectuées sur les précédentes cartes (**Fig. 51, 52 et 53**). La région coudée est observée, notamment dans le cas du LBD orienté en 3'. La vue en 3' permet d'apprécier la forme elliptique de l'ADN induite par les DBDs. Ces derniers sont dans une position qui semble compatible avec la position des LBDs au vu de l'orientation de leur région C-terminale.

Concernant le coactivateur PGC-1 α , ce dernier ne peut pas être placé dans la carte, ce qui est probablement dû à la limite de résolution de celle-ci.

Il est possible de comparer cette carte obtenue avec son image miroir, c'est-à-dire d'en changer la chiralité, cette dernière n'étant pas certaine avant affinement local. Les valeurs de résolution et de B-factor demeurent dans ce cas les mêmes, puisque seule la valeur d'un axe a été changée de 180° sans affecter la résolution (**Fig. 58**).

La structure cristallographique des LBDs, avec cette chiralité de la carte, présente une meilleure accommodation globale. L'ensemble des hélices peuvent être placées et correspondent aux contours de la carte. Seule une partie de l'hélice H3 du LBD situé en 3' fait exception et sort légèrement de la carte. Néanmoins, si cette chiralité est plus favorable aux LBDs, elle exclut totalement les deux fragments de PGC-1 α . Cela peut être lié à la résolution de la carte.

Enfin, les DBDs, dans cette orientation, se situent de l'autre « côté » de l'ADN. Cette topologie globale ressemble à celle du complexe VDR/RXR et à celle de l'homodimère HNF-4. Le récepteur nucléaire se lierait sur son premier demi-site en 5', positionnant son LBD à une distance proche du DBD fixé sur le demi-site étendu où l'interaction est plus forte. La région coudée y serait probablement plus compacte. L'interaction de la région CTE du DBD en 5' avec le petit sillon de l'ADN porteur de l'extension de trois paires de bases « TCA » explique probablement cette orientation des LBDs, cette même interaction n'étant pas observée dans la structure cristallographique sur le second demi-site, permettant probablement une plus grande liberté de la région coudée du ERR α en position 3'. Le second DBD positionné dans la carte présente une orientation de sa région C-terminale moins favorable à expliquer la carte que dans le cadre de la figure 58. Le LBD situé en 3' permettrait la liaison sur le second demi-site avec une région coudée plus « dépliée » que celle en 5'. Il est possible qu'il existe des interactions entre la région LBD et la région DBD, mais la résolution de la carte ne permet pas de les identifier.

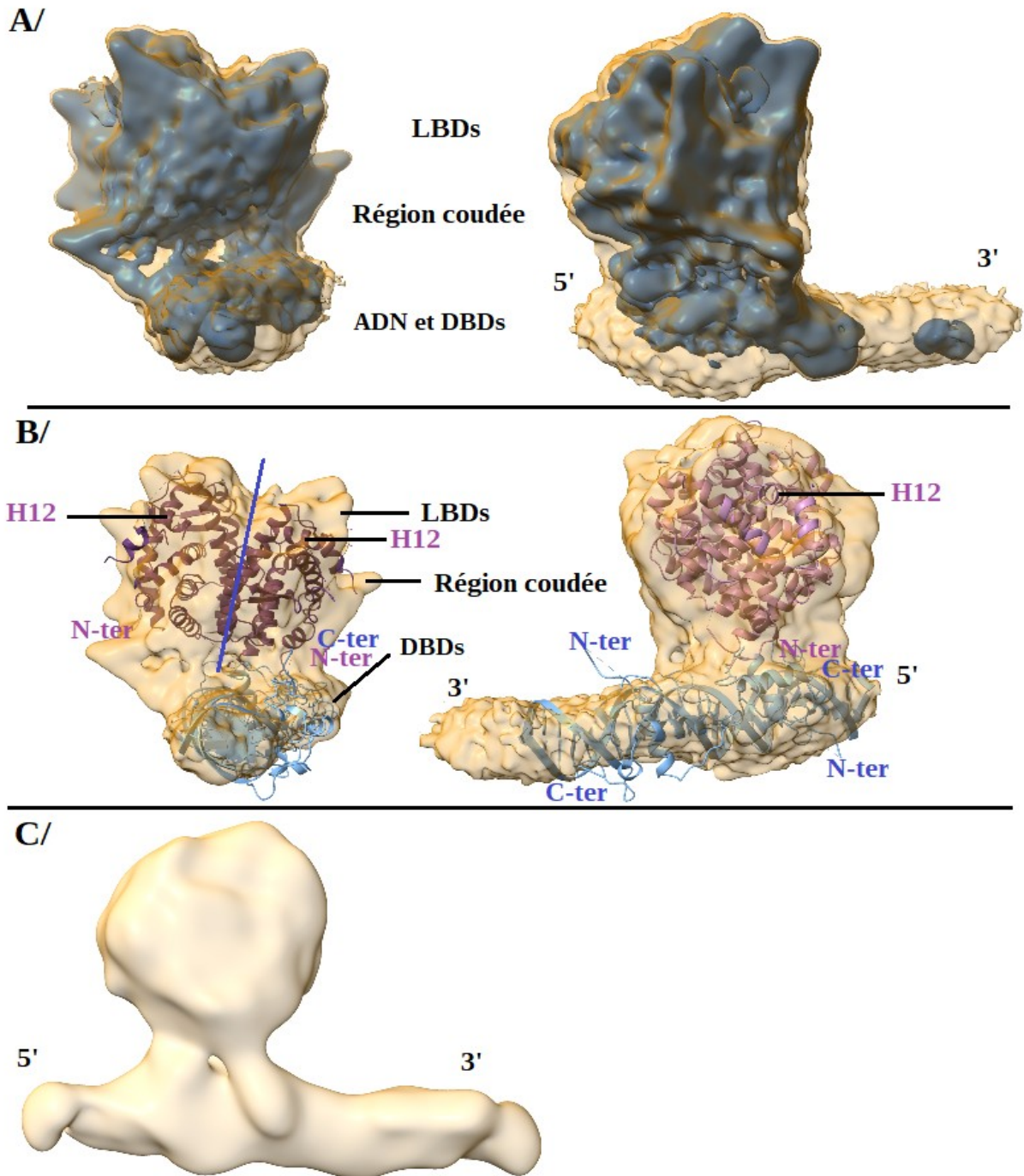


Figure 58: Analyse de la carte chirale

Légende :

A/ Carte issue de l'affinement local après changement de sa chiralité en vue de 3' (à gauche) et en vue de 5'-3' (à droite). La carte bleue correspond à un niveau de contour de 0,15 et la carte en transparence orange correspond à un niveau de contour de 0,065. **B/** Carte en vue de 3' (à gauche) et en vue de 5'-3' (à droite) dans laquelle les domaines cristallographiques des LBDs et des DBDs sur l'élément de réponse composite ont été placés. **C/** Carte précédent l'affinement local dont la chiralité a été changée.

Discussion et Perspectives

L'un des objectifs de cette thèse était de comprendre les mécanismes moléculaires mis en jeu dans la reconnaissance et la fixation de $ERR\alpha$ sur l'élément de réponse composite. Sur ce type de séquence particulière reconnue également par le récepteur aux œstrogènes, les DBDs de $ERR\alpha$ peuvent dimériser, comme cela fut montré au sein de notre équipe (Mohideen-Abdul *et al.*, 2017). Cette dimérisation des DBDs est dépendante des paramètres de formes de l'ADN et n'est pas observée au sein des autres ERRE. L'obtention de la structure cristallographique à 2.48 Å de résolution des DBDs sur l'élément composite a permis de mettre en lumière les interactions mises en jeu entre chacun des DBDs et la mise en place d'une surface d'interaction asymétrique faisant intervenir une boîte fonctionnelle nouvellement identifiée : la boîte « KRRRK ». Celle-ci prend une conformation différente pour chacun des dimères et induit une augmentation de la taille de la surface d'interaction. Cette boîte est spécifique à la famille des ERRs et l'analyse évolutive que nous avons menée montre que l'interface de dimérisation asymétrique qu'elle induit semble avoir été perdue au cours de l'évolution par les récepteurs oxostéroïdiens et le récepteur aux œstrogènes, dont l'interface de dimérisation des DBDs est symétrique.

La comparaison des profils thermodynamiques obtenus par titration calorimétrique isotherme du DBD en présence de différents ERREs ou de l'élément composite nous a permis de mettre en avant la liaison coopérative du dimère sur cet élément de réponse. Les mesures effectuées à différentes températures ont permis de quantifier la capacité thermique des DBDs sur l'élément de réponse et de mettre en évidence leur mode de liaison site-spécifique avec un repliement local important des domaines au contact de l'élément composite. La liaison du dimère de DBDs sur l'élément composite s'effectue par l'intermédiaire d'une coopération entre les domaines via un processus fortement enthalpique. Les facteurs de températures de la structure cristallographique indiquent que la fixation du DBD au niveau du premier demi-site en 5', lequel possède l'extension « TCA » est plus stable que la fixation du DBD sur le second demi-site où l'extension de trois paires de bases est « TAA ». Les expériences de retard sur gel en conditions natives ont en plus permis d'identifier les extensions de trois paires de bases en 5' « TCA » et TGA » comme les plus favorables à la stabilisation de l'interaction d'un DBD de $ERR\alpha$ sur le premier demi-site en 5'. Il convient à ce stade de discuter du cas de la séquence $Tr\alpha$ ERRE testée. L'ITC révèle la possibilité pour deux monomères de DBDs de se fixer sur cet ERRE avec des affinités de l'ordre du nanomolaire et des événements de liaison fortement enthalpiques. Ces résultats ne sont pas consistant avec les expériences de retard sur gel menées avec différentes séquences de ERRE, dont $Tr\alpha$ ERRE. Les deux événements de fixation étant indépendants et l'affinité de fixation du « second » DBD étant du même ordre de grandeur que celui du dimère sur un élément composite, il est raisonnable de supposer qu'il soit possible que le second DBD soit capable d'interagir dynamiquement avec le second demi-site de fixation sur un ERRE ou qu'une autre façon de se lier soit possible pour le DBD sur ce type d'ERRE. Dans le cas d'une interaction dynamique, le DBD en 3' ne serait pas stabilisée du fait des paramètres de formes de l'ADN. Les seconds demi-sites de fixation des ERRE ne sont pas consensus et si, *in vivo*, le récepteur agit sous forme de dimère, il reste à élucider la façon dont le second demi-site est reconnu et fixé par le DBD de $ERR\alpha$ dans le cadre d'un ERRE.

Le second objectif de cette thèse consistait justement à considérer la façon dont le récepteur nucléaire entier se liait sur ses éléments de réponses, incluant l'élément composite, et à considérer par des mesures biochimiques et de spectrométrie de masse, l'impact de PGC-1 α sur son activité et sa conformation. Les mesures réalisées au FIDABio, quoique encore préliminaires et devant être répétées, montrent une augmentation significative de l'affinité pour la séquence $Tr\alpha$ ERRE en présence de PGC-1 α . L'affinité devient du même ordre de grandeur que celle du récepteur pour l'élément composite.

PGC-1 α semble également augmenter l'affinité du récepteur pour l'élément composite, bien que l'effet soit moindre par comparaison avec la séquence ERRE. Il faut bien entendu relativiser les valeurs de Kd mesurées pour ces différentes interactions et reproduire ces résultats avec, par exemple, une autre méthode de mesures telles que l'ITC ou encore la thermophorèse à micro-échelle afin de confirmer les résultats observés. L'effet général décrit par ces mesures est cependant en accord avec les modèles tout atomes calculés à partir des enveloppes SAXS du récepteur ERR γ en présence du coactivateur, obtenu par notre équipe. Ces données suggèrent que, sous l'effet de la liaison de PGC-1 α , le second DBD de ERR α qui serait « flottant » se fixerait sur le second demi-site de fixation en 3' par l'intermédiaire d'un effet allostérique.

Les résultats de XL-MS tendent à indiquer que PGC-1 α est spatialement proche d'au moins l'un des DBDs et qu'il pourrait interagir directement avec ce dernier. De plus, sous l'action de PGC-1 α , le nombre de réticulations intramoléculaires impliquant les DBDs augmentent, ce qui va dans le sens d'un rapprochement des DBDs et d'une potentielle dimérisation induite. Cette interaction avec PGC-1 α induit également un rapprochement des DBDs et des LBDs, ce qui traduit un changement global de conformation pour le récepteur. Les résultats des mesures de HDX indiquent que PGC-1 α induit des changements significatifs de la stabilité de certaines régions de ERR α . Au niveau des DBDs, PGC-1 α induit la déstabilisation dynamique de l'hélice H1 contenant la boîte P et du premier doigt de zinc. Ce phénomène pourrait être nécessaire à l'accommodation du second DBD sur le second demi-site de fixation pour lequel aucune séquence consensus n'a pour l'instant été identifiée. Les mesures réalisées ne permettent cependant pas de déterminer si l'un ou l'autre des DBDs est affecté. La protection induite par le coactivateur au niveau de la boîte D et de la boîte T montre que les résidus impliqués dans la dimérisation sont stabilisés et moins accessibles en solution. Cela va dans le sens d'une dimérisation induite par le coactivateur. Compte-tenu des résultats de XL-MS, la nature de cet effet, à savoir par le biais d'une interaction directe entre le coactivateur et les DBDs ou d'un effet allostérique, demeure à élucider. Les résultats de XL-MS permettent également d'apprécier l'effet de l'élément composite sur la conformation globale du complexe ERR α -PGC-1 α . Celui-ci semble induire le rapprochement d'au moins l'un des DBDs et d'une partie de la région N-terminale de l'un des LBDs. Les résultats de HDX montrent que le coactivateur induit la déstabilisation dynamique d'une large région du domaine N-terminal. Cette région est connue pour pouvoir être successivement phosphorylée au niveau des Ser19 et Ser22, puis sumoylée sur la Lys14 ce qui entraîne l'inhibition de l'activité transcriptionnelle associée à ERR α . Il a par ailleurs été montré que la mutation de ces résidus augmentait l'activité transcriptionnelle du récepteur en présence de PGC-1 α uniquement. Il est donc possible que l'inhibition de l'activité transcriptionnelle causée par la modification de ces résidus affecte l'interaction entre ERR α et le coactivateur PGC-1 α .

Le troisième objectif de cette thèse consistait à élucider la topologie du complexe ERR α -PGC-1 α sur l'élément composite. Les résultats de cryo-microscopie ont permis l'obtention d'une carte qui permet de replacer les domaines individuels de ERR α par rapport à l'élément composite. La forme globale du complexe correspond à la forme en « L » décrite pour les récepteurs USP/VDR, VDR/RXR, PPAR/RXR (SAXS), HNF-4 et RAR/RXR/Med-1. L'un des LBDs est « assis » sur le premier site de fixation en 5', lequel est le site de plus forte affinité des DBDs comme expliqué précédemment. L'autre ERR α dont le LBD est du côté 3' est plus distant de l'ADN. Il déplie sa région coudée de sorte à pouvoir interagir avec son DBD au niveau du second demi-site de fixation en 3'. Les DBDs peuvent être positionnés sur le côté de l'ADN, ce qui explique que le LBD en position 5' puisse être « si » proche du premier demi-site étendu de fixation. De plus, cela est en accord avec les résultats de XL-MS qui suggéraient que l'un des LBDs se rapprochait spatialement des DBDs lors de la formation du

complexe. Il n'est cependant pas possible, au vu de la résolution de la carte, de statuer sur d'éventuelles interactions entre les domaines.

Le coactivateur PGC-1 α n'est pas identifiable à cette résolution. Les résultats de XL-MS nous permettent cependant de supposer que l'interaction avec PGC-1 α est délimitée par une région de 51 résidus comprise entre les Lys184 et Lys235 du coactivateur. Les résultats de HDX suggèrent également que l'interaction au niveau de l'un des LBDs comprend l'hélice H12, mais aussi la boucle entre les hélices H11-H12. Il est difficile de statuer sur l'extension de cette interaction à la boucle comprise entre les hélices H1-H3. La protection de cette boucle pourrait être associée à sa structuration et à la formation d'une hélice H2 par effet allostérique. Il ne s'agit cependant que d'une hypothèse à ce stade. La topologie décrite de ERR α est la seconde connue à ce jour pour un récepteur homodimérique sur un ADN composé d'une séquence inversement répétée en considérant la publication récente de la carte du récepteur aux androgènes sur un élément de réponse ARE³⁵ à une résolution comparable à la nôtre (Wasmuth *et al.*, 2022).

Considérant qu'une carte de cryo-EM peut donner l'un ou l'autre des énantiomères d'une molécule imagée, et considérant que l'hélicité de l'ADN ne peut pas être retrouvée dans la carte obtenue, la topologie décrite n'est donc pas complète. L'affinement local de la région comprenant l'ADN et les LBDs pourrait permettre de retrouver l'hélicité de l'ADN et de conclure sur la chiralité de la carte obtenue si la résolution est suffisante. Il nous est cependant possible de comparer la carte obtenue avec celle d'autres récepteurs nucléaires. Tout d'abord, par comparaison avec le récepteur USP/ECR, lequel est un hétérodimère sur un élément de réponse de type IR1, soit une séquence inversement répétée séparée d'une seule paire de base, la carte de ERR α -PGC-1 α -BE42embedded dans sa chiralité initiale (**Fig. 57**) montre une topologie similaire (**Fig. 59**). La comparaison des positions des domaines individuels de ERR α en accord avec la carte obtenue avec la structure de l'hétérodimère USP/ECR en accord avec sa carte de cryo-microscopie montre que le LBD en 5' est le plus proche de l'ADN et le DBD de ce même LBD se fixe sur le demi-site de fixation en 5'. Néanmoins, l'orientation globale des LBDs est différente entre ces complexes. Le LBD dont le site de fixation est en 5' est tourné vers le côté 3' dans le cas de ERR α alors que ce dernier est tourné vers le côté 5' dans le cas du complexe USP/ECR. Cela induit une proximité spatiale plus importante entre le DBD en 5' de l'homodimère ERR α et ses LBDs que celle observée pour l'hétérodimère USP/ECR. Cette proximité est accentuée par le positionnement du DBD en 5' de ERR α sur le « côté » de l'ADN par rapport à l'axe formé par l'ADN et les LBDs alors que ce DBD est en « dessous » de l'ADN dans le cas du complexe USP/ECR. Il est intéressant de noter que dans ces deux complexes, le DBD se liant en 3' semble se situer à la même position. Dans le complexe USP/ECR, le coactivateur a été identifié comme se liant sur le LBD en position 3'. La carte que nous avons obtenue ne nous permet pas de positionner le coactivateur, bien que la région correspondant au LBP en position 3' soit effectivement plus grande que celle correspondant au LBD en 5'.

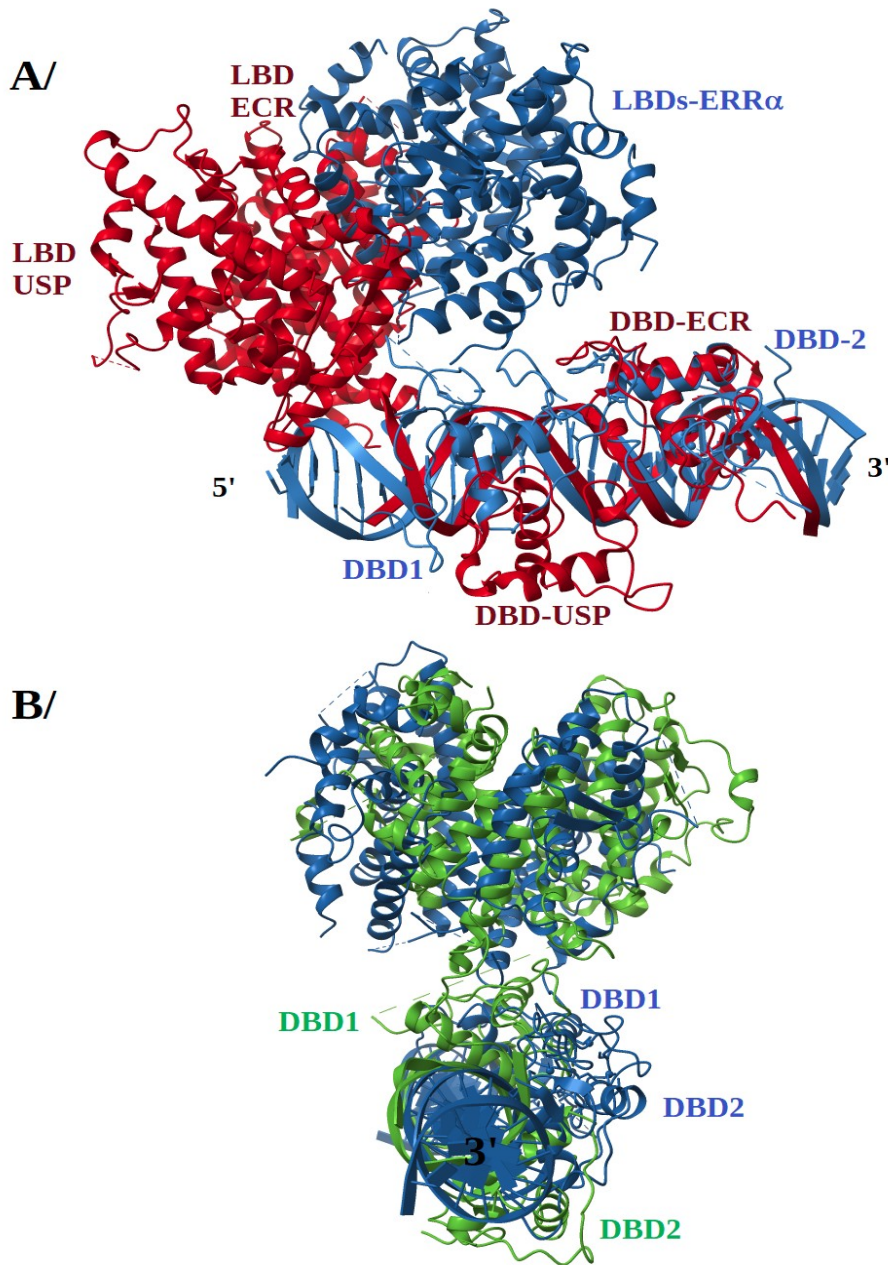


Figure 59: Comparaison de la topologie du complexe $ERR\alpha$ -PGC-1 α avec les récepteurs USP/ECR et HNF-4

Légende :

A/ Structure du récepteur USP/ECR en accord avec sa carte de cryo-EM montrée en rouge (PDB : 4UMM, Maletta et al., 2014) comparée à la position des domaines individuels de $ERR\alpha$ dans la carte résultant de l'affinement local en bleu (Fig. 57). L'alignement des structures s'est fait à la main en prenant la position des sillons de l'ADN comme référence dans le logiciel ChimeraX. **B/** Structure cristallographique de l'homodimère de HNF-4 (PDB : 4iqr, Chandra et al., 2013) en vert comparée à la position des domaines individuels de $ERR\alpha$ dans la carte résultant de l'affinement local après changement de la chiralité (Fig. 58). L'alignement des structures s'est fait à la main en prenant la position des sillons de l'ADN comme référence dans le logiciel ChimeraX.

Pour comparer les cartes du complexe $ERR\alpha$ -PGC-1 α -BE42embedded avec les récepteurs reconnaissant des éléments de réponse composés d'une répétition directe, la carte après changement de sa chiralité est plus adéquate (**Fig. 58**). La comparaison des positions des domaines individuels de $ERR\alpha$ en accord avec la carte obtenue et la structure de l'homodimère HNF-4 sur un élément de réponse composé d'une répétition directe séparée d'une paire de base révèle un positionnement globale des domaines très similaire (**Fig. 59-B**). Le récepteur dont le demi-site de fixation se trouve en 5' présente un LBD spatialement plus proche de son DBD que celui dont le demi-site de fixation est en position 3'. Le $ERR\alpha$ dont le demi-site de fixation se trouve en position 3' dispose d'un LBD plus éloigné de son DBD que celui de HNF-4. Cela s'explique probablement par la région coudée qui adopte, dans le cas de la carte de $ERR\alpha$ -PGC-1 α -BE42embedded, une conformation différente. La position relative des DBDs est également intéressante à comparer. Le DBD se liant en 5' est entre l'élément de réponse et son LBD dans le cas de HNF-4 alors que celui-ci est décalé dans le sens horaire, en observant l'ADN depuis son extrémité 3', dans le cas de $ERR\alpha$, venant donc davantage sur le côté de l'ADN. Le DBD se liant en position 3' est sur le côté de l'ADN pour les deux homodimères, mais celui de $ERR\alpha$ est cette fois déplacé dans le sens anti-horaire par comparaison avec celui de HNF-4 lorsque l'on regarde l'axe de l'ADN depuis son extrémité 3'. Ce rapprochement des DBDs de $ERR\alpha$ par comparaison avec ceux de HNF-4 est probablement le fait de la mise en place de leur interface de dimérisation. Du fait de la chiralité de la carte du complexe qui reste à confirmer, il n'est cependant pas encore possible de statuer sur ces comparaisons de façon certaine.

Le jeu de données acquis après vitrification au chameleon doit être poursuivi afin d'identifier la nature des deux autres classes 3D identifiées et de déterminer la chiralité de la carte du complexe par l'affinement de la région correspondant à l'élément de réponse et aux DBDs. La résolution globale des cartes obtenues demeure néanmoins un véritable frein à l'analyse de l'architecture du complexe, ainsi qu'à l'identification des éléments de structures secondaires ou du coactivateur.

Il est possible que le coactivateur soit présent, mais qu'il soit difficile de l'identifier à cause de sa flexibilité lui permettant d'adopter un vaste ensemble de conformations différentes. L'analyse structurale du complexe reposerait alors sur l'analyse d'un grand nombre de données pour identifier au moins l'une de ces conformations. S'il est certain que la mise en place, durant la thèse, d'un microscope électronique de cryo-microscopie Glacios fonctionnant à une tension d'accélération de 200 kV pour obtenir un meilleur contraste et du système de vitrification Chameleon permettant l'obtention d'une épaisseur de la glace plus fine furent cruciaux pour l'obtention de cartes du complexe à une résolution d'environ 10 à 12 Å, le projet souffre encore d'un certain nombre de limites.

La vitrification avec le Chameleon n'est pas facilement reproductible, le temps de vitrification n'étant pas suffisamment précis pour permettre de reproduire une même vitrification un certain nombre de fois. En l'occurrence, pour un temps demandé de vitrification de 54 ms, l'appareil peut vitrifier une grille en 39 ou en 74 ms, ce qui peut suffire à impacter la qualité de la glace résultante sur la grille. Cela nécessite donc probablement l'amélioration des conditions de tampons afin de pouvoir vitrifier de façon reproductible et avec une glace fine l'échantillon. L'utilisation de grilles d'affinité est également une solution en essayant de capturer le complexe par le coactivateur PGC-1 α afin de s'assurer de sa présence. Néanmoins, le bruit de fond induit par le support suffit à limiter grandement l'application de cette méthode à un échantillon d'un poids moléculaire peu élevé dont le contraste est faible. L'utilisation d'agents réticulant comme le glutaraldéhyde peut être tenté dans l'optique de stabiliser une conformation du complexe et/ou du coactivateur. Les conditions d'une telle réticulation sont en cours d'optimisation.

Enfin, la reconstitution d'un complexe *in vitro* plus important ou la capture d'un complexe d'activation de la transcription depuis une souche cellulaire pourrait s'avérer plus propice à l'étude de l'interaction de ERR avec ses coactivateurs, l'échantillon étant probablement plus stable, plus facile à aligner, bien qu'il puisse être plus hétérogène. Certaines protéines sont connues pour interagir avec ERR α et/ou PGC-1 α . La protéine PROX1, par exemple, est connue pour se lier en amont de ERR α sur l'ADN et serait capable d'interagir avec PGC-1 α . Elle agirait comme un corépresseur du complexe et pourrait former un complexe de répression minimum de l'activité du récepteur. De la même manière, la protéine PCAF, que nous avons pu purifier au cours de cette thèse est un bon candidat. La construction purifiée ici ne comportait pas la région flexible présente en N-terminal. C'est peut-être un défaut, cette région ayant peut-être un rôle dans l'interaction avec le récepteur nucléaire compte-tenu de la présence de nombreuses régions intrinsèquement désordonnées chez l'ensemble des coactivateurs et corépresseurs des récepteurs nucléaires tels que les protéines de la famille SRC ou NCOR. La protéine p300 est également une protéine capable d'interagir avec les récepteurs nucléaires et leurs coactivateurs, et certaines expériences de co-immuno-précipitations tendent à indiquer qu'elle interagit avec PGC-1 α et ERR α . Les résultats de cette thèse laissent sous-entendre qu'il serait possible d'optimiser la purification du complexe formé avec le récepteur nucléaire et son coactivateur. Néanmoins, trop peu d'éléments demeurent connus sur les complexes entiers que forme ERR α et l'enchaînement des interactions entre les différents partenaires. Le séquençage de l'assemblage de ces complexes sera nécessaire pour procéder ensuite à l'étude structurale plus fine des mécanismes moléculaires mises en jeu. Cela apportera à la communauté scientifique une meilleure compréhension des étapes nécessaires à la mise en place de la transcription et permettra le développement de nouveaux outils scientifiques, technologiques et thérapeutiques.

Bibliographie

- Alaynick *et al.*, “ERRγ Directs and Maintains the Transition to Oxidative Metabolism in the Postnatal Heart,” *Cell Metabolism* 6, no. 1 (July 11, 2007): 13–24, <https://doi.org/10.1016/j.cmet.2007.06.007>.
- Alexander *et al.*, “The Concise Guide to PHARMACOLOGY 2013/14: Nuclear Hormone Receptors,” *British Journal of Pharmacology* 170, no. 8 (2013): 1652–75, <https://doi.org/10.1111/bph.12448>.
- Allegretti *et al.*, “Atomic Model of the F420-Reducing [NiFe] Hydrogenase by Electron Cryo-Microscopy Using a Direct Electron Detector,” ed. Wesley Sundquist, *ELife* 3 (February 25, 2014): e01963, <https://doi.org/10.7554/eLife.01963>.
- Almlöf *et al.*, “Role of Important Hydrophobic Amino Acids in the Interaction between the Glucocorticoid Receptor T1-Core Activation Domain and Target Factors,” *Biochemistry* 37, no. 26 (June 1, 1998): 9586–94, <https://doi.org/10.1021/bi973029x>.
- Anbalagan *et al.*, “Post-Translational Modifications of Nuclear Receptors and Human Disease,” *Nuclear Receptor Signaling* 10, no. 1 (January 1, 2012): nrs.10001, <https://doi.org/10.1621/nrs.10001>.
- Andre J. and Rochefort H., “Specific Effect of Estrogens on an Interaction between the Uterine Estradiol Receptor and DNA,” *FEBS Letters* 29, no. 2 (1973): 135–40, [https://doi.org/10.1016/0014-5793\(73\)80544-7](https://doi.org/10.1016/0014-5793(73)80544-7).
- Aranda & Pascual, “Nuclear Hormone Receptors and Gene Expression” *Physiological Reviews*, (July 1, 2001), <https://journals.physiology.org/doi/full/10.1152/physrev.2001.81.3.1269>
- Ariazi *et al.*, “Estrogen-Related Receptor α and Estrogen-Related Receptor γ Associate with Unfavorable and Favorable Biomarkers, Respectively, in Human Breast Cancer1 | Cancer Research | American Association for Cancer Research,” accessed August 17, 2022, <https://aacrjournals.org/cancerres/article/62/22/6510/509422/Estrogen-related-Receptor-and-Estrogen-related>.
- Bai *et al.*, “Ribosome Structures to Near-Atomic Resolution from Thirty Thousand Cryo-EM Particles,” ed. Werner Kühlbrandt, *ELife* 2 (February 19, 2013): e00461, <https://doi.org/10.7554/eLife.00461>.
- Baker, “Steroid Receptor Phylogeny and Vertebrate Origins,” *Molecular and Cellular Endocrinology* 135, no. 2 (December 12, 1997): 101–7, [https://doi.org/10.1016/S0303-7207\(97\)00207-4](https://doi.org/10.1016/S0303-7207(97)00207-4).
- Baker, Nelson, and Studer, “Origin of the Response to Adrenal and Sex Steroids: Roles of Promiscuity and Co-Evolution of Enzymes and Steroid Receptors,” *The Journal of Steroid Biochemistry and Molecular Biology*, Steroid/Sterol Signaling – Promiscuity and Diversity, 151 (July 1, 2015): 12–24, <https://doi.org/10.1016/j.jsbmb.2014.10.020>.
- Bammes *et al.*, “Radiation Damage Effects at Four Specimen Temperatures from 4 to 100K,” *Journal of Structural Biology* 169, no. 3 (March 1, 2010): 331–41, <https://doi.org/10.1016/j.jsb.2009.11.001>.
- Beato *et al.*, “Interaction of Glucocorticoids with Rat Liver Nuclei. II. Nature of the Cytosol Transfer Factor and the Nuclear Acceptor Site,” research-article, ACS Publications (American Chemical Society, May 1, 2002), world, <https://doi.org/10.1021/bi00742a002>.

- Beinsteiner *et al.*, “A structural signature motif enlightens the origin and diversification of nuclear receptors.” *PLoS genetics* vol. 17,4 e1009492. (April 21, 2021), <https://doi.org/10.1371/journal.pgen.1009492>
- Beinsteiner, Brice *et al.*, “A Novel Nuclear Receptor Subfamily Enlightens the Origin of Heterodimerization.” *BMC Biology* 20 (October 5, 2022): 217. <https://doi.org/10.1186/s12915-022-01413-0>.
- Bepler *et al.*, “Positive-Unlabeled Convolutional Neural Networks for Particle Picking in Cryo-Electron Micrographs,” *Research in Computational Molecular Biology : ... Annual International Conference, RECOMB ... : Proceedings. RECOMB (Conference : 2005-)* 10812 (April 2018): 245–47.
- Billas *et al.*, “Allosteric Controls of Nuclear Receptor Function in the Regulation of Transcription.” *Journal of Molecular Biology, Allosteric Interactions and Biological Regulation (Part II)*, 425, no. 13 (July 10, 2013): 2317–29. <https://doi.org/10.1016/j.jmb.2013.03.017>.
- Boland, Chang, and Barford, “The Potential of Cryo-Electron Microscopy for Structure-Based Drug Design,” ed. Rob L.M. van Montfort and Paul Workman, *Essays in Biochemistry* 61, no. 5 (November 8, 2017): 543–60, <https://doi.org/10.1042/EBC20170032>.
- Bookout *et al.*, “Anatomical Profiling of Nuclear Receptor Expression Reveals a Hierarchical Transcriptional Network,” *Cell* 126, no. 4 (August 25, 2006): 789–99, <https://doi.org/10.1016/j.cell.2006.06.049>.
- Bourguet *et al.*, “Crystal Structure of the Ligand-Binding Domain of the Human Nuclear Receptor RXR- α | Nature,” accessed August 18, 2022, <https://www.nature.com/articles/375377a0>.
- Bourguet, Germain, and Gronemeyer, “Nuclear Receptor Ligand-Binding Domains: Three-Dimensional Structures, Molecular Interactions and Pharmacological Implications,” *Trends in Pharmacological Sciences* 21, no. 10 (October 1, 2000): 381–88, [https://doi.org/10.1016/S0165-6147\(00\)01548-0](https://doi.org/10.1016/S0165-6147(00)01548-0).
- Brelivet *et al.*, “Signature of the Oligomeric Behaviour of Nuclear Receptors at the Sequence and Structural Level,” *EMBO Reports* 5, no. 4 (April 2004): 423–29, <https://doi.org/10.1038/sj.embor.7400119>.
- Brilot *et al.*, “Beam-Induced Motion of Vitrified Specimen on Holey Carbon Film,” *Journal of Structural Biology* 177, no. 3 (March 1, 2012): 630–37, <https://doi.org/10.1016/j.jsb.2012.02.003>.
- Brown *et al.*, “Estrogen-Related Receptors Mediate the Adaptive Response of Brown Adipose Tissue to Adrenergic Stimulation,” *IScience* 2 (April 27, 2018): 221–37, <https://doi.org/10.1016/j.isci.2018.03.005>.
- Burris *et al.*, “Nuclear Receptors and Their Selective Pharmacologic Modulators,” ed. Dianne M. Perez, *Pharmacological Reviews* 65, no. 2 (April 1, 2013): 710–78, <https://doi.org/10.1124/pr.112.006833>.
- Cai *et al.*, “Regulation of Glycolysis and the Warburg Effect by Estrogen-Related Receptors,” *Oncogene* 32, no. 16 (April 2013): 2079–86, <https://doi.org/10.1038/onc.2012.221>.
- Campbell *et al.*, “Movies of Ice-Embedded Particles Enhance Resolution in Electron Cryo-Microscopy,” *Structure* 20, no. 11 (November 7, 2012): 1823–28, <https://doi.org/10.1016/j.str.2012.08.026>.
- Cartoni *et al.*, “Mitofusins 1/2 and ERR α Expression Are Increased in Human Skeletal Muscle after Physical Exercise,” *The Journal of Physiology* 567, no. 1 (2005): 349–58, <https://doi.org/10.1113/jphysiol.2005.092031>.

- Cavallini *et al.*, “Oestrogen Receptor-Related Receptor Alpha (ERR α) and Oestrogen Receptors (ER α and ER β) Exhibit Different Gene Expression in Human Colorectal Tumour Progression,” *European Journal of Cancer* 41, no. 10 (July 1, 2005): 1487–94, <https://doi.org/10.1016/j.ejca.2005.04.008>.
- Chandra *et al.*, “Structure of the Intact PPAR- γ -RXR- α Nuclear Receptor Complex on DNA,” *Nature* 456, no. 7220 (November 2008): 350–56, <https://doi.org/10.1038/nature07413>.
- Chandra *et al.*, “Multidomain Integration in the Structure of the HNF-4 α Nuclear Receptor Complex,” *Nature* 495, no. 7441 (March 2013): 394–98, <https://doi.org/10.1038/nature11966>.
- Chandra *et al.*, “The Quaternary Architecture of RAR β -RXR α Heterodimer Facilitates Domain-Domain Signal Transmission,” *Nature Communications* 8, no. 1 (October 11, 2017): 868, <https://doi.org/10.1038/s41467-017-00981-y>.
- Charest-Marcotte *et al.*, “The Homeobox Protein Prox1 Is a Negative Modulator of ERR α /PGC-1 α Bioenergetic Functions,” *Genes & Development* 24, no. 6 (March 15, 2010): 537–42, <https://doi.org/10.1101/gad.1871610>.
- Chaveroux *et al.*, “Molecular and Genetic Crosstalks between MTOR and ERR α Are Key Determinants of Rapamycin-Induced Nonalcoholic Fatty Liver,” *Cell Metabolism* 17, no. 4 (April 2, 2013): 586–98, <https://doi.org/10.1016/j.cmet.2013.03.003>.
- Chan *et al.*, “Rates of Induction of Specific Translatable Messenger RNAs for Ovalbumin and Avidin by Steroid Hormones | PNAS,” accessed August 18, 2022, <https://www.pnas.org/doi/abs/10.1073/pnas.70.6.1870>.
- Chari *et al.*, “ProteoPlex - Stability Optimization of Macromolecular Complexes by Sparse-Matrix Screening of Chemical Space.” *Nature Methods* 12, no. 9 (September 2015): 859–65. <https://doi.org/10.1038/nmeth.3493>.
- Chen and Evans, “A Transcriptional Co-Repressor That Interacts with Nuclear Hormone Receptors,” *Nature* 377, no. 6548 (October 1995): 454–57, <https://doi.org/10.1038/377454a0>.
- Chen *et al.*, “Identification of Two HERR2-Related Novel Nuclear Receptors Utilizing Bioinformatics and Inverse PCR,” *Gene* 228, no. 1 (March 4, 1999): 101–9, [https://doi.org/10.1016/S0378-1119\(98\)00619-2](https://doi.org/10.1016/S0378-1119(98)00619-2).
- Cho *et al.*, “Peroxisome Proliferator-Activated Receptor γ Coactivator 1 (PGC-1)- and Estrogen-Related Receptor (ERR)-Induced Regulator in Muscle 1 (PERM1) Is a Tissue-Specific Regulator of Oxidative Capacity in Skeletal Muscle Cells *,” *Journal of Biological Chemistry* 288, no. 35 (August 30, 2013): 25207–18, <https://doi.org/10.1074/jbc.M113.489674>.
- Comstock *et al.*, “Estrogen-Induced Changes in Translation, and Specific Messenger RNA Levels during Oviduct Differentiation,” *Proceedings of the National Academy of Sciences* 69, no. 9 (September 1972): 2377–80, <https://doi.org/10.1073/pnas.69.9.2377>.
- Crowther, R. A., Linda A. Amos, J. T. Finch, D. J. De Rosier, and A. Klug. “Three Dimensional Reconstructions of Spherical Viruses by Fourier Synthesis from Electron Micrographs.” *Nature* 226, no. 5244 (May 1970): 421–25. <https://doi.org/10.1038/226421a0>.
- Dahlman-Wright and McEwan, “Structural Studies of Mutant Glucocorticoid Receptor Transactivation Domains Establish a Link between Transactivation Activity in Vivo and α -Helix-Forming Potential in Vitro,” *Biochemistry* 35, no. 4 (January 1, 1996): 1323–27, <https://doi.org/10.1021/bi952409k>.
- Dang *et al.*, “Cryo-EM Structures of the TMEM16A Calcium-Activated Chloride Channel,” *Nature* 552, no. 7685 (December 2017): 426–29, <https://doi.org/10.1038/nature25024>.
- Darimont *et al.*, “Structure and Specificity of Nuclear Receptor-Coactivator Interactions,” *Genes & Development* 12, no. 21 (November 1, 1998): 3343–56, <https://doi.org/10.1101/gad.12.21.3343>.

- Darrow *et al.*, “Chameleon: Next Generation Sample Preparation for CryoEM Based on Spotiton,” *Microscopy and Microanalysis* 25, no. S2 (August 2019): 994–95, <https://doi.org/10.1017/S1431927619005701>.
- Deblois *et al.*, “Genome-Wide Identification of Direct Target Genes Implicates Estrogen-Related Receptor α as a Determinant of Breast Cancer Heterogeneity,” *Cancer Research* 69, no. 15 (July 29, 2009): 6149–57, <https://doi.org/10.1158/0008-5472.CAN-09-1251>.
- Deblois *et al.*, “ERR α Mediates Metabolic Adaptations Driving Lapatinib Resistance in Breast Cancer,” *Nature Communications* 7, no. 1 (July 12, 2016): 12156, <https://doi.org/10.1038/ncomms12156>.
- Drulyte *et al.*, “Approaches to Altering Particle Distributions in Cryo-Electron Microscopy Sample Preparation,” *Acta Crystallographica Section D: Structural Biology* 74, no. 6 (June 1, 2018): 560–71, <https://doi.org/10.1107/S2059798318006496>.
- Dubochet *et al.*, “Cryo-Electron Microscopy of Vitrified Specimens,” *Quarterly Reviews of Biophysics* 21, no. 2 (May 1988): 129–228, <https://doi.org/10.1017/S0033583500004297>.
- Dufour *et al.*, “Genome-Wide Orchestration of Cardiac Functions by the Orphan Nuclear Receptors ERR α and γ ,” *Cell Metabolism* 5, no. 5 (May 9, 2007): 345–56, <https://doi.org/10.1016/j.cmet.2007.03.007>.
- Dufour *et al.*, “Genomic Convergence among ERR α , PROX1, and BMAL1 in the Control of Metabolic Clock Outputs,” *PLOS Genetics* 7, no. 6 (June 23, 2011): e1002143, <https://doi.org/10.1371/journal.pgen.1002143>.
- Egea, Mitschler, and Moras, “Molecular Recognition of Agonist Ligands by RXRs,” *Molecular Endocrinology* 16, no. 5 (May 1, 2002): 987–97, <https://doi.org/10.1210/mend.16.5.0823>.
- Eichner and Giguère, “Estrogen Related Receptors (ERRs): A New Dawn in Transcriptional Control of Mitochondrial Gene Networks,” *Mitochondrion* 11, no. 4 (July 1, 2011): 544–52, <https://doi.org/10.1016/j.mito.2011.03.121>.
- Evrard *et al.*, “DADIMODO: a program for refining the structure of multidomain proteins and complexes against small-angle scattering data and NMR-derived restraints.” (2011) *J. Appl. Cryst.*, 44: 1264-1271. <https://doi.org/10.1107/S0021889811035758>
- Fan and Evans, “PPARs and ERRs: Molecular Mediators of Mitochondrial Metabolism,” *Current Opinion in Cell Biology*, Cell regulation, 33 (April 1, 2015): 49–54, <https://doi.org/10.1016/j.ceb.2014.11.002>.
- Fan *et al.*, “ERR γ Promotes Angiogenesis, Mitochondrial Biogenesis, and Oxidative Remodeling in PGC1 α/β -Deficient Muscle,” *Cell Reports* 22, no. 10 (March 6, 2018): 2521–29, <https://doi.org/10.1016/j.celrep.2018.02.047>.
- Faruqi and McMullan, “Electronic Detectors for Electron Microscopy,” *Quarterly Reviews of Biophysics* 44, no. 3 (August 2011): 357–90, <https://doi.org/10.1017/S0033583511000035>.
- Feng *et al.*, “A Fast and Effective Microfluidic Spraying-Plunging Method for High-Resolution Single-Particle Cryo-EM,” *Structure (London, England : 1993)* 25, no. 4 (April 4, 2017): 663–670.e3, <https://doi.org/10.1016/j.str.2017.02.005>.
- Feng *et al.*, “Cryo-EM Studies of TMEM16F Calcium-Activated Ion Channel Suggest Features Important for Lipid Scrambling,” *Cell Reports* 28, no. 2 (July 9, 2019): 567-579.e4, <https://doi.org/10.1016/j.celrep.2019.06.023>.
- Francis *et al.*, “Nuclear Receptors and the Control of Metabolism,” *Annual Review of Physiology* 65 (2003): 261–311, <https://doi.org/10.1146/annurev.physiol.65.092101.142528>.
- Frank, *Three-Dimensional Electron Microscopy of Macromolecular Assemblies: Visualization of Biological Molecules in Their Native State* (Oxford University Press, 2006).

- Freedman *et al.*, “The Function and Structure of the Metal Coordination Sites within the Glucocorticoid Receptor DNA Binding Domain,” *Nature* 334, no. 6182 (August 1988): 543–46, <https://doi.org/10.1038/334543a0>.
- Fu *et al.*, “P300 and P300/CAMP-Response Element-Binding Protein-Associated Factor Acetylate the Androgen Receptor at Sites Governing Hormone-Dependent Transactivation *,” *Journal of Biological Chemistry* 275, no. 27 (July 7, 2000): 20853–60, <https://doi.org/10.1074/jbc.M000660200>.
- Fujimoto and Sato, “Clinical Implication of Estrogen-Related Receptor (ERR) Expression in Uterine Endometrial Cancers,” *The Journal of Steroid Biochemistry and Molecular Biology* 116, no. 1 (August 1, 2009): 71–75, <https://doi.org/10.1016/j.jsbmb.2009.04.012>.
- Fujimura *et al.*, “Increased Expression of Estrogen-Related Receptor α (ERR α) Is a Negative Prognostic Predictor in Human Prostate Cancer,” *International Journal of Cancer* 120, no. 11 (2007): 2325–30, <https://doi.org/10.1002/ijc.22363>.
- Gampe *et al.*, “Asymmetry in the PPAR γ /RXR α Crystal Structure Reveals the Molecular Basis of Heterodimerization among Nuclear Receptors,” *Molecular Cell* 5, no. 3 (March 1, 2000): 545–55, [https://doi.org/10.1016/S1097-2765\(00\)80448-7](https://doi.org/10.1016/S1097-2765(00)80448-7).
- Gantner *et al.*, “Complementary Roles of Estrogen-Related Receptors in Brown Adipocyte Thermogenic Function,” *Endocrinology* 157, no. 12 (December 1, 2016): 4770–81, <https://doi.org/10.1210/en.2016-1767>.
- Gao *et al.*, “The Discovery of a Novel Series of Potential ERR α Inverse Agonists Based on P-Nitrobenzenesulfonamide Template for Triple-Negative Breast Cancer in Vivo,” *Journal of Enzyme Inhibition and Medicinal Chemistry* 37, no. 1 (December 31, 2022): 125–34, <https://doi.org/10.1080/14756366.2021.1995728>.
- Gearhart *et al.*, “Monomeric Complex of Human Orphan Estrogen Related Receptor-2 with DNA: A Pseudo-Dimer Interface Mediates Extended Half-Site Recognition,” *Journal of Molecular Biology* 327, no. 4 (April 4, 2003): 819–32, [https://doi.org/10.1016/S0022-2836\(03\)00183-9](https://doi.org/10.1016/S0022-2836(03)00183-9).
- Gehring Ulrich and Tomkins Gordon M., “A New Mechanism for Steroid Unresponsiveness: Loss of Nuclear Binding Activity of a Steroid Hormone Receptor,” *Cell* 3, no. 3 (November 1, 1974): 301–6, [https://doi.org/10.1016/0092-8674\(74\)90145-7](https://doi.org/10.1016/0092-8674(74)90145-7).
- Giguère *et al.*, “Identification of a New Class of Steroid Hormone Receptors,” *Nature* 331, no. 6151 (January 1988): 91–94, <https://doi.org/10.1038/331091a0>.
- Giguère, “Orphan Nuclear Receptors: From Gene to Function*,” *Endocrine Reviews* 20, no. 5 (October 1, 1999): 689–725, <https://doi.org/10.1210/edrv.20.5.0378>.
- Glaeser, “Limitations to Significant Information in Biological Electron Microscopy as a Result of Radiation Damage,” *Journal of Ultrastructure Research* 36, no. 3 (August 1, 1971): 466–82, [https://doi.org/10.1016/S0022-5320\(71\)80118-1](https://doi.org/10.1016/S0022-5320(71)80118-1).
- Glaeser, Han, and Cate, “Streptavidin Monolayer-Crystal Affinity Grids: A Step Toward Controlling What Happens During Cryo-EM Sample Preparation,” *Microscopy and Microanalysis* 23, no. S1 (July 2017): 820–21, <https://doi.org/10.1017/S1431927617004767>.
- Glaeser, “Proteins, Interfaces, and Cryo-EM Grids,” *Current Opinion in Colloid & Interface Science, Microscopy Methods: You Can Observe a Lot by Watching*, 34 (March 1, 2018): 1–8, <https://doi.org/10.1016/j.cocis.2017.12.009>.
- Glaeser *et al.*, “Defocus-Dependent Thon-Ring Fading,” *Ultramicroscopy* 222 (March 1, 2021): 113213, <https://doi.org/10.1016/j.ultramic.2021.113213>.

- Glascock R. F. and Hoekstra W. G., “Selective Accumulation of Tritium-Labelled Hexoestrol by the Reproductive Organs of Immature Female Goats and Sheep,” *Biochemical Journal* 72, no. 4 (August 1959): 673–82.
- Glasser, Chytil, and Spelsberg, “Early Effects of Oestradiol-17 β on the Chromatin and Activity of the Deoxyribonucleic Acid-Dependent Ribonucleic Acid Polymerases (I and II) of the Rat Uterus,” *Biochemical Journal* 130, no. 4 (December 1, 1972): 947–57, <https://doi.org/10.1042/bj1300947>.
- Grant and Grigorieff, “Measuring the Optimal Exposure for Single Particle Cryo-EM Using a 2.6 Å Reconstruction of Rotavirus VP6,” ed. Wesley I Sundquist, *ELife* 4 (May 29, 2015): e06980, <https://doi.org/10.7554/eLife.06980>.
- Grassucci, Derek J. Taylor, and Joachim Frank, “Preparation of Macromolecular Complexes for Cryo-Electron Microscopy,” *Nature Protocols* 2, no. 12 (December 2007): 3239–46, <https://doi.org/10.1038/nprot.2007.452>.
- Greschik *et al.*, “Structural and Functional Evidence for Ligand-Independent Transcriptional Activation by the Estrogen-Related Receptor 3,” *Molecular Cell* 9, no. 2 (February 1, 2002): 303–13, [https://doi.org/10.1016/S1097-2765\(02\)00444-6](https://doi.org/10.1016/S1097-2765(02)00444-6).
- Greshik, H., Flaig, R., Renaud, J. P., & Moras, D. (2004). PROTEIN STRUCTURE AND FOLDING-Structural Basis for the Deactivation of the Estrogen-related Receptor γ by Diethylstilbestrol or 4-Hydroxytamoxifen and Determinants of Selectivity. *Journal of Biological Chemistry*, 279(32), 33639-33646.
- Greschik *et al.*, “Communication between the ERR α Homodimer Interface and the PGC-1 α Binding Surface via the Helix 8–9 Loop*,” *Journal of Biological Chemistry* 283, no. 29 (July 18, 2008): 20220–30, <https://doi.org/10.1074/jbc.M801920200>.
- Gronemeyer and Moras, “How to Finger DNA,” *Nature* 375, no. 6528 (May 1995): 190–91, <https://doi.org/10.1038/375190a0>.
- Gronemeyer *et al.*, “Principles for Modulation of the Nuclear Receptor Superfamily,” *Nature Reviews Drug Discovery* 3, no. 11 (November 2004): 950–64, <https://doi.org/10.1038/nrd1551>.
- Härd *et al.*, “Solution Structure of the Glucocorticoid Receptor DNA-Binding Domain,” *Science* 249, no. 4965 (July 13, 1990): 157–60, <https://doi.org/10.1126/science.2115209>.
- Haelens *et al.*, “The Hinge Region Regulates DNA Binding, Nuclear Translocation, and Transactivation of the Androgen Receptor,” *Cancer Research* 67, no. 9 (May 4, 2007): 4514–23, <https://doi.org/10.1158/0008-5472.CAN-06-1701>.
- Han *et al.*, “Long Shelf-Life Streptavidin Support-Films Suitable for Electron Microscopy of Biological Macromolecules,” *Journal of Structural Biology* 195, no. 2 (August 1, 2016): 238–44, <https://doi.org/10.1016/j.jsb.2016.06.009>.
- He B. *et al.*, “Electrostatic Modulation in Steroid Receptor Recruitment of LXXLL and FXXLF Motifs | Molecular and Cellular Biology,” accessed August 18, 2022, <https://journals.asm.org/doi/full/10.1128/MCB.23.6.2135-2150.2003>.
- Heard *et al.*, “Human ERR γ , a Third Member of the Estrogen Receptor-Related Receptor (ERR) Subfamily of Orphan Nuclear Receptors: Tissue-Specific Isoforms Are Expressed during Development and in the Adult,” *Molecular Endocrinology* 14, no. 3 (March 1, 2000): 382–92, <https://doi.org/10.1210/mend.14.3.0431>.
- Heery *et al.*, “A Signature Motif in Transcriptional Co-Activators Mediates Binding to Nuclear Receptors,” *Nature* 387, no. 6634 (June 1997): 733–36, <https://doi.org/10.1038/42750>.
- Helsen *et al.*, “Structural Basis for Nuclear Hormone Receptor DNA Binding,” *Molecular and Cellular Endocrinology*, Nuclear Receptor Structure: Dynamics and Function, 348, no. 2 (January 30, 2012): 411–17, <https://doi.org/10.1016/j.mce.2011.07.025>.

- Helsen and Claessens. “Looking at Nuclear Receptors from a New Angle.” *Molecular and Cellular Endocrinology* 382, no. 1 (January 25, 2014): 97–106. <https://doi.org/10.1016/j.mce.2013.09.009>.
- Henderson R. and Unwin P. N. T., “Three-Dimensional Model of Purple Membrane Obtained by Electron Microscopy,” *Nature* 257, no. 5521 (September 1975): 28–32, <https://doi.org/10.1038/257028a0>.
- Hollenberg *et al.*, “Primary Structure and Expression of a Functional Human Glucocorticoid Receptor CDNA,” *Nature* 318, no. 6047 (December 1985): 635–41, <https://doi.org/10.1038/318635a0>.
- Hu X. and Lazar Mitchell A., “The CoRNR Motif Controls the Recruitment of Corepressors by Nuclear Hormone Receptors,” *Nature* 402, no. 6757 (November 1999): 93–96, <https://doi.org/10.1038/47069>.
- Huang, Chandra, and Rastinejad, “Structural Overview of the Nuclear Receptor Superfamily: Insights into Physiology and Therapeutics,” *Annual Review of Physiology* 72 (2010): 247, <https://doi.org/10.1146/annurev-physiol-021909-135917>.
- Hudson, Youn, and Ortlund, “Crystal Structure of the Mineralocorticoid Receptor DNA Binding Domain in Complex with DNA,” *PLOS ONE* 9, no. 9 (September 4, 2014): e107000, <https://doi.org/10.1371/journal.pone.0107000>.
- Huss *et al.*, “Constitutive activities of estrogen-related receptors: Transcriptional regulation of metabolism by the ERR pathways in health and disease,” *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, Volume 1852, Issue 9, (2015) : 1912-1927, <https://doi.org/10.1016/j.bbadis.2015.06.016>.
- Jensen EV, Jacobson HI “Basic guides to the mechanism of estrogen action.” *Recent Prog Horm Res* (1962) 18:387–414
- Jensen *et al.*, “A Two-Step Mechanism for the Interaction of Estradiol with Rat Uterus.,” *Proceedings of the National Academy of Sciences* 59, no. 2 (February 1968): 632–38, <https://doi.org/10.1073/pnas.59.2.632>.
- Johnson *et al.*, “Ligand-Induced Stabilization of PPAR γ Monitored by NMR Spectroscopy: Implications for Nuclear Receptor Activation” Edited by P. E. Wright,” *Journal of Molecular Biology* 298, no. 2 (April 28, 2000): 187–94, <https://doi.org/10.1006/jmbi.2000.3636>.
- Kallen *et al.*, “Evidence for Ligand-Independent Transcriptional Activation of the Human Estrogen-Related Receptor α (ERR α): CRYSTAL STRUCTURE OF ERR α LIGAND BINDING DOMAIN IN COMPLEX WITH PEROXISOME PROLIFERATOR-ACTIVATED RECEPTOR COACTIVATOR-1 α *,” *Journal of Biological Chemistry* 279, no. 47 (November 19, 2004): 49330–37, <https://doi.org/10.1074/jbc.M407999200>.
- Kenney, Wicks, and Greenman, “Hydrocortisone Stimulation of RNA Synthesis in Induction of Hepatic Enzymes,” *Journal of Cellular and Comparative Physiology* 66, no. S1 (1965): 125–36, <https://doi.org/10.1002/jcp.1030660412>.
- Kaledhonkar *et al.*, “Time-Resolved Cryo-Electron Microscopy Using a Microfluidic Chip,” in *Protein Complex Assembly: Methods and Protocols*, ed. Joseph A. Marsh, Methods in Molecular Biology (New York, NY: Springer, 2018), 59–71, https://doi.org/10.1007/978-1-4939-7759-8_4.
- Khatter *et al.*, “Structure of the Human 80S Ribosome,” *Nature* 520, no. 7549 (April 2015): 640–45, <https://doi.org/10.1038/nature14427>.
- Khorasanizadeh and Rastinejad, “Nuclear-Receptor Interactions on DNA-Response Elements,” *Trends in Biochemical Sciences* 26, no. 6 (June 1, 2001): 384–90, [https://doi.org/10.1016/S0968-0004\(01\)01800-X](https://doi.org/10.1016/S0968-0004(01)01800-X).

- Kim *et al.*, “ESRRA (Estrogen Related Receptor Alpha) Is a Critical Regulator of Intestinal Homeostasis through Activation of Autophagic Flux via Gut Microbiota,” *Autophagy* 17, no. 10 (n.d.): 2856–75, <https://doi.org/10.1080/15548627.2020.1847460>.
- Kim and Lazar, “Transcriptional Control of Circadian Rhythms and Metabolism: A Matter of Time and Space,” *Endocrine Reviews* 41, no. 5 (October 1, 2020): 707–32, <https://doi.org/10.1210/endrev/bnaa014>.
- King *et al.*, “THE INTRANUCLEAR LOCALIZATION OF [6,7-3H]-OESTRADIOL-17 β IN DIMETHYLBENZANTHRACENE-INDUCED RAT MAMMARY ADENOCARCINOMA AND OTHER TISSUES,” *Journal of Endocrinology* 36, no. 2 (October 1, 1966): 139-NP, <https://doi.org/10.1677/joe.0.0360139>.
- King RJ, Gordon J. Involvement of DNA in the acceptor mechanism for uterine oestradiol receptor. *Nat New Biol.* 1972 Dec 6;240(101):185-7. doi: 10.1038/newbio240185a0. PMID: 4343951.
- Klebl *et al.*, “Sample Deposition onto Cryo-EM Grids: From Sprays to Jets and Back,” *Acta Crystallographica Section D: Structural Biology* 76, no. 4 (April 1, 2020): 340–49, <https://doi.org/10.1107/S2059798320002958>.
- Krust *et al.*, “The Chicken Oestrogen Receptor Sequence: Homology with v-ErbA and the Human Oestrogen and Glucocorticoid Receptors,” *The EMBO Journal* 5, no. 5 (May 1986): 891–97, <https://doi.org/10.1002/j.1460-2075.1986.tb04300.x>.
- Kuijper *et al.*, “FEI’s Direct Electron Detector Developments: Embarking on a Revolution in Cryo-TEM,” *Journal of Structural Biology*, Recent Advances in Detector Technologies and Applications for Molecular TEM, 192, no. 2 (November 1, 2015): 179–87, <https://doi.org/10.1016/j.jsb.2015.09.014>.
- Kumar and Thompson, “Transactivation Functions of the N-Terminal Domains of Nuclear Hormone Receptors: Protein Folding and Coactivator Interactions,” *Molecular Endocrinology* 17, no. 1 (January 1, 2003): 1–10, <https://doi.org/10.1210/me.2002-0258>.
- Kunath *et al.*, “Time-Resolved Low-Dose Microscopy of Glutamine Synthetase Molecules,” *Ultramicroscopy* 13, no. 3 (January 1, 1984): 241–52, [https://doi.org/10.1016/0304-3991\(84\)90203-1](https://doi.org/10.1016/0304-3991(84)90203-1).
- Lavery and Mcewan, “Structure and Function of Steroid Receptor AF1 Transactivation Domains: Induction of Active Conformations,” *Biochemical Journal* 391, no. 3 (October 25, 2005): 449–64, <https://doi.org/10.1042/BJ20050872>.
- Lathe Richard and Houston Douglas R., “Fatty-Acylation Target Sequence in the Ligand-Binding Domain of Vertebrate Steroid Receptors Demarcates Evolution from Estrogen-Related Receptors,” *The Journal of Steroid Biochemistry and Molecular Biology*, The Evolution of Steroid-liganded Nuclear Receptors, 184 (November 1, 2018): 20–28, <https://doi.org/10.1016/j.jsbmb.2018.07.010>.
- le Maire *et al.*, “Regulation of RXR-RAR Heterodimers by RXR- and RAR-Specific Ligands and Their Combinations.” *Cells*. 2019;8(11):1392. Published 2019 Nov 5. <https://doi.org/10.3390/cells8111392>
- Li, Lambert, and Xu, “Activation of Nuclear Receptors: A Perspective from Structural Genomics,” *Structure* 11, no. 7 (July 1, 2003): 741–46, [https://doi.org/10.1016/S0969-2126\(03\)00133-3](https://doi.org/10.1016/S0969-2126(03)00133-3).
- Li *et al.*, “Electron Counting and Beam-Induced Motion Correction Enable near-Atomic-Resolution Single-Particle Cryo-EM,” *Nature Methods* 10, no. 6 (June 2013): 584–90, <https://doi.org/10.1038/nmeth.2472>.

- Li *et al.*, “ERR α Activates SHMT2 Transcription to Enhance the Resistance of Breast Cancer to Lapatinib via Modulating the Mitochondrial Metabolic Adaption,” *Bioscience Reports* 40, no. 1 (January 17, 2020): BSR20192465, <https://doi.org/10.1042/BSR20192465>.
- Lou *et al.*, “Structure of the Retinoid X Receptor α –Liver X Receptor β (RXR α –LXR β) Heterodimer on DNA,” *Nature Structural & Molecular Biology* 21, no. 3 (March 2014): 277–81, <https://doi.org/10.1038/nsmb.2778>.
- Luisi *et al.*, “Crystallographic Analysis of the Interaction of the Glucocorticoid Receptor with DNA,” *Nature* 352, no. 6335 (August 1991): 497–505, <https://doi.org/10.1038/352497a0>.
- Ma, Qin, and Li, “Role of STAT3 Signaling Pathway in Breast Cancer,” *Cell Communication and Signaling* 18, no. 1 (February 28, 2020): 33, <https://doi.org/10.1186/s12964-020-0527-z>.
- Mäeots *et al.*, “Modular Microfluidics Enables Kinetic Insight from Time-Resolved Cryo-EM,” *Nature Communications* 11 (July 10, 2020): 3465, <https://doi.org/10.1038/s41467-020-17230-4>.
- Mainwaring W. I., “Interaction of Androgenic Steroids with Prostatic Chromatin.,” *Biochemical Journal* 124, no. 5 (October 1971): 42P–43P.
- Mainwaring W. I. P. and Peterken Brenda M., “A Reconstituted Cell-Free System for the Specific Transfer of Steroid–Receptor Complexes into Nuclear Chromatin Isolated from Rat Ventral Prostate Gland,” *Biochemical Journal* 125, no. 1 (November 1971): 285–95.
- Maletta *et al.*, “The Palindromic DNA-Bound USP/EcR Nuclear Receptor Adopts an Asymmetric Organization with Allosteric Domain Positioning,” *Nature Communications* 5, no. 1 (June 19, 2014): 4139, <https://doi.org/10.1038/ncomms5139>.
- Mangelsdorf *et al.*, “The Nuclear Receptor Superfamily: The Second Decade,” *Cell* 83, no. 6 (December 15, 1995): 835–39.
- Mark Danielsen, “Bioinformatics of Nuclear Receptors,” in *Steroid Receptor Methods: Protocols and Assays*, ed. Benjamin A. Lieberman, *Methods in Molecular Biology*TM (Totowa, NJ: Humana Press, 2001), 3–22, <https://doi.org/10.1385/1-59259-115-9:3>.
- Markov and Laudet, “Origin and Evolution of the Ligand-Binding Ability of Nuclear Receptors,” *Molecular and Cellular Endocrinology*, *Evolution of Nuclear Hormone Receptors*, 334, no. 1 (March 1, 2011): 21–30, <https://doi.org/10.1016/j.mce.2010.10.017>.
- Markov *et al.*, “Origin of an Ancient Hormone/Receptor Couple Revealed by Resurrection of an Ancestral Estrogen | Science Advances,” accessed August 18, 2022, <https://www.science.org/doi/full/10.1126/sciadv.1601778>.
- Maurer H. Rainer and Chalkley G. Roger, “Some Properties of a Nuclear Binding Site of Estradiol,” *Journal of Molecular Biology* 27, no. 3 (August 14, 1967): 431–41, [https://doi.org/10.1016/0022-2836\(67\)90049-6](https://doi.org/10.1016/0022-2836(67)90049-6).
- May, “Novel Drugs That Target the Estrogen-Related Receptor Alpha: Their Therapeutic Potential in Breast Cancer,” *Cancer Management and Research* 6 (May 23, 2014): 225–52, <https://doi.org/10.2147/CMAR.S35024>.
- McCorquodale D.J. and Mueller Gerald C., “Preparation and Properties of Amino Acid-Adenylic Acid Mixed Anhydrides,” *Archives of Biochemistry and Biophysics* 77, no. 1 (September 1958): 13–19, [https://doi.org/10.1016/0003-9861\(58\)90035-3](https://doi.org/10.1016/0003-9861(58)90035-3).
- McMullan *et al.*, “Detective Quantum Efficiency of Electron Area Detectors in Electron Microscopy,” *Ultramicroscopy* 109, no. 9 (August 1, 2009): 1126–43, <https://doi.org/10.1016/j.ultramic.2009.04.002>.
- McMullan *et al.*, “Comparison of Optimal Performance at 300keV of Three Direct Electron Detectors for Use in Low Dose Electron Microscopy,” *Ultramicroscopy* 147 (December 1, 2014): 156–63, <https://doi.org/10.1016/j.ultramic.2014.08.002>.

- Means A R and Hamilton T H , “Evidence for Depression of Nuclear Protein Synthesis and Concomitant Stimulation of Nuclear RNA Synthesis during Early Estrogen Action.,” *Proceedings of the National Academy of Sciences* 56, no. 2 (August 1966): 686–93, <https://doi.org/10.1073/pnas.56.2.686>.
- Means *et al.*, “Ovalbumin Messenger RNA of Chick Oviduct: Partial Characterization, Estrogen Dependence, and Translation In Vitro,” *Proceedings of the National Academy of Sciences* 69, no. 5 (May 1972): 1146–50, <https://doi.org/10.1073/pnas.69.5.1146>.
- Mindell and Grigorieff, “Accurate Determination of Local Defocus and Specimen Tilt in Electron Microscopy,” *Journal of Structural Biology* 142, no. 3 (June 1, 2003): 334–47, [https://doi.org/10.1016/S1047-8477\(03\)00069-8](https://doi.org/10.1016/S1047-8477(03)00069-8).
- Mizwicki *et al.*, “Identification of an Alternative Ligand-Binding Pocket in the Nuclear Vitamin D Receptor and Its Functional Importance in $1\alpha,25(\text{OH})_2$ -Vitamin D₃ Signaling | PNAS,” accessed August 18, 2022, <https://www.pnas.org/doi/abs/10.1073/pnas.0403606101>.
- Mohideen-Abdul *et al.*, “Importance of the Sequence-Directed DNA Shape for Specific Binding Site Recognition by the Estrogen-Related Receptor,” *Frontiers in Endocrinology* 8 (2017), <https://www.frontiersin.org/articles/10.3389/fendo.2017.00140>.
- Moras Dino and Gronemeyer Hinrich, “The Nuclear Receptor Ligand-Binding Domain: Structure and Function,” *Current Opinion in Cell Biology* 10, no. 3 (June 1, 1998): 384–91, [https://doi.org/10.1016/S0955-0674\(98\)80015-X](https://doi.org/10.1016/S0955-0674(98)80015-X).
- Mullican, DiSpirito, and Lazar, “The Orphan Nuclear Receptors at Their 25-Year Reunion,” *Journal of Molecular Endocrinology* 51, no. 3 (December 1, 2013): T115–40, <https://doi.org/10.1530/JME-13-0212>.
- Musille *et al.*, “Unexpected Allosteric Network Contributes to LHRH-1 Co-Regulator Selectivity *,” *Journal of Biological Chemistry* 291, no. 3 (January 15, 2016): 1411–26, <https://doi.org/10.1074/jbc.M115.662874>.
- Musliner Thomas A. and Chader Gerald J., “Estradiol-Receptors of the Rat Uterus: Interaction of the Cytoplasmic Estrogen-Receptor with DNA in Vitro,” *Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis* 262, no. 3 (March 24, 1972): 256–63, [https://doi.org/10.1016/0005-2787\(72\)90261-4](https://doi.org/10.1016/0005-2787(72)90261-4).
- Nagy *et al.*, “Mechanism of Corepressor Binding and Release from Nuclear Hormone Receptors,” *Genes & Development* 13, no. 24 (December 15, 1999): 3209–16.
- Nakane and Scheres. “Multi-Body Refinement of Cryo-EM Cryo-Electron Microscopy (Cryo-EM) Images in RELION.” In *CryoEM: Methods and Protocols*, edited by Tamir Gonen and Brent L. Nannenga, 145–60. Methods in Molecular Biology. New York, NY: Springer US, 2021. https://doi.org/10.1007/978-1-0716-0966-8_7.
- Naydenova, Peet, and Russo, “Multifunctional Graphene Supports for Electron Cryomicroscopy,” *Proceedings of the National Academy of Sciences* 116, no. 24 (June 11, 2019): 11718–24, <https://doi.org/10.1073/pnas.1904766116>.
- Nie Y. and Wong C., “Suppressing the Activity of ERR α in 3T3-L1 Adipocytes Reduces Mitochondrial Biogenesis but Enhances Glycolysis and Basal Glucose Uptake,” *Journal of Cellular and Molecular Medicine* 13, no. 9b (2009): 3051–60, <https://doi.org/10.1111/j.1582-4934.2008.00382.x>.
- Niesen *et al.*, “The Use of Differential Scanning Fluorimetry to Detect Ligand Interactions That Promote Protein Stability.” *Nature Protocols* 2, no. 9 (September 2007): 2212–21. <https://doi.org/10.1038/nprot.2007.321>.
- Nicolette and Gorski, “Cortisol Effects on the Uterine Response to Estrogen₁,” *Endocrinology* 74, no. 6 (June 1, 1964): 955–59, <https://doi.org/10.1210/endo-74-6-955>.

- Nils Marechal *et al.*, “Formation of Thyroid Hormone Revealed by a Cryo-EM Structure of Native Bovine Thyroglobulin,” *Nature Communications* 13, no. 1 (May 2, 2022): 2380, <https://doi.org/10.1038/s41467-022-30082-4>.
- Nomenclature, R. A. R., heterodimer partner NR0B2, S. S., Peroxisome proliferator-activated receptor- γ , P., NR1D1, N. R. C. R. E. A. R. E., NR1D2, R. E. R. E., & NR1H1, E. E. R. (1999). Nuclear Receptors Nomenclature Committee. A Unified Nomenclature System for the Nuclear Receptor Superfamily. *Cell* 97, 161-163 (1999). *Cell*, 97, 161-163.
- O’Malley B W and McGuire W L , “Studies on the Mechanism of Estrogen-Mediated Tissue Differentiation: Regulation of Nuclear Transcription and Induction of New RNA Species.,” *Proceedings of the National Academy of Sciences* 60, no. 4 (August 1968): 1527–34, <https://doi.org/10.1073/pnas.60.4.1527>.
- O’Malley, B.W. (1971). Mechanisms of Action of Steroid Hormones. *N. Engl. J. Med.* 284, 370–377.
- O’Malley and Tsai, “Molecular Pathways of Steroid Receptor Action,” *Biology of Reproduction* 46, no. 2 (February 1, 1992): 163–67, <https://doi.org/10.1095/biolreprod46.2.163>.
- Orlov *et al.*, “Structure of the Full Human RXR/VDR Nuclear Receptor Heterodimer Complex with Its DR3 Target DNA,” *The EMBO Journal* 31, no. 2 (January 18, 2012): 291–300, <https://doi.org/10.1038/emboj.2011.445>.
- Osz, Judit, Maxim V. Pethoukhov, Serena Sirigu, Dmitri I. Svergun, Dino Moras, and Natacha Rochel. “Solution Structures of PPAR γ /RXR α Complexes.” *PPAR Research* 2012 (2012): 701412. <https://doi.org/10.1155/2012/701412>.
- Panda Satchidananda , “Circadian Physiology of Metabolism,” *Science* 354, no. 6315 (November 25, 2016): 1008–15, <https://doi.org/10.1126/science.aah4967>.
- Parks *et al.*, “‘True Agonadism’: A Misnomer?,” *The Journal of Pediatrics* 84, no. 3 (March 1, 1974): 375–80, [https://doi.org/10.1016/S0022-3476\(74\)80720-1](https://doi.org/10.1016/S0022-3476(74)80720-1).
- Paulino *et al.*, “Activation Mechanism of the Calcium-Activated Chloride Channel TMEM16A Revealed by Cryo-EM,” *Nature* 552, no. 7685 (December 2017): 421–25, <https://doi.org/10.1038/nature24652>.
- Pawlak, Lefebvre, and Staels, “General Molecular Biology and Architecture of Nuclear Receptors,” *Current Topics in Medicinal Chemistry* 12, no. 6 (March 1, 2012): 486–504, <https://doi.org/10.2174/156802612799436641>.
- Peng L. *et al.*, “Identification of Pyrido[1,2- α]Pyrimidine-4-Ones as New Molecules Improving the Transcriptional Functions of Estrogen-Related Receptor α ,” *Journal of Medicinal Chemistry* 54, no. 21 (November 10, 2011): 7729–33, <https://doi.org/10.1021/jm200976s>.
- Perissi V. *et al.*, “Molecular Determinants of Nuclear Receptor–Corepressor Interaction,” *Genes & Development* 13, no. 24 (December 15, 1999): 3198–3208.
- Pratt *et al.*, “Role of Molecular Chaperones in Steroid Receptor Action,” ed. Iain J. McEwan, *Essays in Biochemistry* 40 (June 1, 2004): 41–58, <https://doi.org/10.1042/bse0400041>.
- Punjani, Ali *et al.*, “CryoSPARC: Algorithms for Rapid Unsupervised Cryo-EM Structure Determination.” *Nature Methods* 14, no. 3 (March 2017): 290–96. <https://doi.org/10.1038/nmeth.4169>.
- Punjani, Ali, Haowei Zhang, and David J. Fleet. “Non-Uniform Refinement: Adaptive Regularization Improves Single-Particle Cryo-EM Reconstruction.” *Nature Methods* 17, no. 12 (December 2020): 1214–21. <https://doi.org/10.1038/s41592-020-00990-8>.
- Punjani, Ali, and David J. Fleet. “3D Variability Analysis: Resolving Continuous Flexibility and Discrete Heterogeneity from Single Particle Cryo-EM.” *Journal of Structural Biology* 213, no. 2 (June 1, 2021): 107702. <https://doi.org/10.1016/j.jsb.2021.107702>.

- Rangwala *et al.*, “Estrogen-Related Receptor γ Is a Key Regulator of Muscle Mitochondrial Activity and Oxidative Capacity,” *Journal of Biological Chemistry* 285, no. 29 (July 16, 2010): 22619–29, <https://doi.org/10.1074/jbc.M110.125401>.
- Rastinejad, “Structure of the RXR–RAR DNA-Binding Complex on the Retinoic Acid Response Element DR1,” *The EMBO Journal* 19, no. 5 (March 2000): 1045–54, <https://doi.org/10.1093/emboj/19.5.1045>.
- Redmon *et al.*, “You Only Look Once: Unified, Real-Time Object Detection,” 2016, 779–88, https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Redmon_You_Only_Look_CVPR_2016_paper.html.
- Renaud *et al.*, “Crystal Structure of the RAR- γ Ligand-Binding Domain Bound to All-Trans Retinoic Acid,” *Nature* 378, no. 6558 (December 1995): 681–89, <https://doi.org/10.1038/378681a0>.
- Ringold *et al.*, “Dexamethasone-Mediated Induction of Mouse Mammary Tumor Virus RNA: A System for Studying Glucocorticoid Action,” *Cell* 6, no. 3 (November 1, 1975): 299–305, [https://doi.org/10.1016/0092-8674\(75\)90181-6](https://doi.org/10.1016/0092-8674(75)90181-6).
- Ringold *et al.*, “Glucocorticoid-Stimulated Accumulation of Mouse Mammary Tumor Virus RNA: Increased Rate of Synthesis of Viral RNA.,” *Proceedings of the National Academy of Sciences* 74, no. 7 (July 1977): 2879–83, <https://doi.org/10.1073/pnas.74.7.2879>.
- Rochel *et al.*, “Common Architecture of Nuclear Receptor Heterodimers on DNA Direct Repeat Elements with Different Spacings,” *Nature Structural & Molecular Biology* 18, no. 5 (May 2011): 564–70, <https://doi.org/10.1038/nsmb.2054>.
- Roemer *et al.*, “Structure of the Progesterone Receptor-Deoxyribonucleic Acid Complex: Novel Interactions Required for Binding to Half-Site Response Elements,” *Molecular Endocrinology* 20, no. 12 (December 1, 2006): 3042–52, <https://doi.org/10.1210/me.2005-0511>.
- Rohou and Grigorieff, “CTFFIND4: Fast and Accurate Defocus Estimation from Electron Micrographs,” *Journal of Structural Biology*, Recent Advances in Detector Technologies and Applications for Molecular TEM, 192, no. 2 (November 1, 2015): 216–21, <https://doi.org/10.1016/j.jsb.2015.08.008>.
- Rosenfeld *et al.*, “INDUCTION OF SPECIFIC TRANSLATABLE MESSENGER RNA’S BY OESTROGEN AND PROGESTERONE,” *European Journal of Endocrinology* 71, no. 2_Suppla (December 1, 1972): S381–95, <https://doi.org/10.1530/acta.0.071S0381>.
- Ruskin, Yu, and Grigorieff, “Quantitative Characterization of Electron Detectors for Transmission Electron Microscopy,” *Journal of Structural Biology* 184, no. 3 (December 1, 2013): 385–93, <https://doi.org/10.1016/j.jsb.2013.10.016>.
- Russo and Passmore, “Controlling Protein Adsorption on Graphene for Cryo-EM Using Low-Energy Hydrogen Plasmas,” *Nature Methods* 11, no. 6 (June 2014): 649–52, <https://doi.org/10.1038/nmeth.2931>.
- Schmid, “Chapter 2 - Single-Particle Electron Cryotomography (CryoET),” in *Advances in Protein Chemistry and Structural Biology*, ed. Steven J. Ludtke and B. V. Venkataram Prasad, vol. 82, Recent Advances in Electron Cryomicroscopy, Part B (Academic Press, 2011), 37–65, <https://doi.org/10.1016/B978-0-12-386507-6.00002-6>.
- Scholtes C. and Giguère V., “Transcriptional Regulation of ROS Homeostasis by the ERR Subfamily of Nuclear Receptors,” *Antioxidants* 10, no. 3 (March 2021): 437, <https://doi.org/10.3390/antiox10030437>.
- Schreiber *et al.*, “The Estrogen-Related Receptor α (ERR α) Functions in PPAR γ Coactivator 1 α (PGC-1 α)-Induced Mitochondrial Biogenesis | PNAS,” accessed August 17, 2022, <https://www.pnas.org/doi/abs/10.1073/pnas.0308686101>.

- Schwabe *et al.*, “The Crystal Structure of the Estrogen Receptor DNA-Binding Domain Bound to DNA: How Receptors Discriminate between Their Response Elements,” *Cell* 75, no. 3 (November 5, 1993): 567–78, [https://doi.org/10.1016/0092-8674\(93\)90390-C](https://doi.org/10.1016/0092-8674(93)90390-C).
- Scheres, Sjors H. W., Mikel Valle, Rafael Nuñez, Carlos O. S. Sorzano, Roberto Marabini, Gabor T. Herman, and Jose-Maria Carazo. “Maximum-Likelihood Multi-Reference Refinement for Electron Microscopy Images.” *Journal of Molecular Biology* 348, no. 1 (April 1, 2005): 139–49. <https://doi.org/10.1016/j.jmb.2005.02.031>.
- Scheres, Sjors H.W., Mikel Valle, and José-María Carazo. “Fast Maximum-Likelihood Refinement of Electron Microscopy Images.” *Bioinformatics* 21, no. suppl_2 (January 1, 2005): ii243–44. <https://doi.org/10.1093/bioinformatics/bti1140>.
- Scheres, Sjors H.W., and Shaoxia Chen. “Prevention of Overfitting in Cryo-EM Structure Determination.” *Nature Methods* 9, no. 9 (September 2012): 853–54. <https://doi.org/10.1038/nmeth.2115>.
- Scheres, Sjors H. W. “RELION: Implementation of a Bayesian Approach to Cryo-EM Structure Determination.” *Journal of Structural Biology* 180, no. 3 (December 1, 2012): 519–30. <https://doi.org/10.1016/j.jsb.2012.09.006>.
- Scheres, “Beam-Induced Motion Correction for Sub-Megadalton Cryo-EM Particles,” *ELife* 3 (August 13, 2014): e03665, <https://doi.org/10.7554/eLife.03665>.
- Scolnick *et al.*, “Rifampin, Oral Contraceptives, and Pregnancy | JAMA | JAMA Network,” accessed August 18, 2022, <https://jamanetwork.com/journals/jama/article-abstract/348317>.
- Shinozuka *et al.*, “Discovery of a Novel Class of ERR α Agonists,” *ACS Medicinal Chemistry Letters* 12, no. 5 (May 13, 2021): 817–21, <https://doi.org/10.1021/acsmchemlett.1c00100>.
- Shyamala and Gorski, “Estrogen Receptors in the Rat Uterus: Studies on the Interaction of Cytosol and Nuclear Binding Sites,” *Journal of Biological Chemistry* 244, no. 5 (March 10, 1969): 1097–1103, [https://doi.org/10.1016/S0021-9258\(18\)91813-2](https://doi.org/10.1016/S0021-9258(18)91813-2).
- Sies *et al.*, “Defining Roles of Specific Reactive Oxygen Species (ROS) in Cell Biology and Physiology,” *Nature Reviews Molecular Cell Biology* 23, no. 7 (July 2022): 499–515, <https://doi.org/10.1038/s41580-022-00456-z>.
- Sigworth, “A Maximum-Likelihood Approach to Single-Particle Image Refinement,” *Journal of Structural Biology* 122, no. 3 (January 1, 1998): 328–39, <https://doi.org/10.1006/jsbi.1998.4014>.
- Singh *et al.*, “Thyroid Hormone Receptor and ERR α Coordinately Regulate Mitochondrial Fission, Mitophagy, Biogenesis, and Function | Science Signaling,” accessed August 17, 2022, <https://www.science.org/doi/abs/10.1126/scisignal.aam5855>.
- Sjögren *et al.*, “Branched-Chain Amino Acid Metabolism Is Regulated by ERR α in Primary Human Myotubes and Is Further Impaired by Glucose Loading in Type 2 Diabetes,” *Diabetologia* 64, no. 9 (September 1, 2021): 2077–91, <https://doi.org/10.1007/s00125-021-05481-9>.
- Sladek, Bader, and Giguère, “The Orphan Nuclear Receptor Estrogen-Related Receptor Alpha Is a Transcriptional Regulator of the Human Medium-Chain Acyl Coenzyme A Dehydrogenase Gene,” *Molecular and Cellular Biology* 17, no. 9 (September 1997): 5400–5409, <https://doi.org/10.1128/MCB.17.9.5400>.
- Solomon *et al.*, “Crystal Structure of the Human LRH-1 DBD–DNA Complex Reveals Ftz-F1 Domain Positioning Is Required for Receptor Activity,” *Journal of Molecular Biology* 354, no. 5 (December 16, 2005): 1091–1102, <https://doi.org/10.1016/j.jmb.2005.10.009>.

- Sonoda *et al.*, “Nuclear Receptor ERR α and Coactivator PGC-1 β Are Effectors of IFN- γ -Induced Host Defense,” *Genes & Development* 21, no. 15 (August 1, 2007): 1909–20, <https://doi.org/10.1101/gad.1553007>.
- Soriano *et al.*, “Evidence for a Mitochondrial Regulatory Pathway Defined by Peroxisome Proliferator–Activated Receptor- γ Coactivator-1 α , Estrogen-Related Receptor- α , and Mitofusin 2,” *Diabetes* 55, no. 6 (June 1, 2006): 1783–91, <https://doi.org/10.2337/db05-0509>.
- Spear *et al.*, “The Influence of Frame Alignment with Dose Compensation on the Quality of Single Particle Reconstructions,” *Journal of Structural Biology*, Recent Advances in Detector Technologies and Applications for Molecular TEM, 192, no. 2 (November 1, 2015): 196–203, <https://doi.org/10.1016/j.jsb.2015.09.006>.
- Spelsberg Thomas C. , Steggles Alan W., and Bert W. O’Malley, “Progesterone-Binding Components of Chick Oviduct,” *Journal of Biological Chemistry* 246, no. 13 (July 1971): 4188–97, [https://doi.org/10.1016/S0021-9258\(18\)62071-X](https://doi.org/10.1016/S0021-9258(18)62071-X).
- Steggles *et al.*, “Soluble Complexes between Steroid Hormones and Target-Tissue Receptors Bind Specifically to Target-Tissue Chromatin,” accessed August 17, 2022, <https://doi.org/10.1073/pnas.68.7.1479>.
- Sun *et al.*, “An Estrogen Receptor α -Dependent Regulation of Estrogen Receptor–Related Receptor α in the Proliferation of Endometrial Carcinoma Cells,” *International Journal of Gynecological Cancer* 16, no. s2 (2006): 564–68, <https://doi.org/10.1111/j.1525-1438.2006.00697.x>.
- Svensson *et al.*, “Crystal Structure of the Heterodimeric Complex of LXR α and RXR β Ligand-Binding Domains in a Fully Agonistic Conformation,” *The EMBO Journal* 22, no. 18 (September 15, 2003): 4625–33, <https://doi.org/10.1093/emboj/cdg456>.
- Svensson *et al.*, “Skeletal Muscle PGC-1 α Modulates Systemic Ketone Body Homeostasis and Ameliorates Diabetic Hyperketonemia in Mice,” *The FASEB Journal* 30, no. 5 (2016): 1976–86, <https://doi.org/10.1096/fj.201500128>.
- Taylor and Glaeser, “Electron Diffraction of Frozen, Hydrated Protein Crystals,” *Science* 186, no. 4168 (December 13, 1974): 1036–37, <https://doi.org/10.1126/science.186.4168.1036>.
- Tenbaum Stephan and Baniahmad Aria, “Nuclear Receptors: Structure, Function and Involvement in Disease,” *The International Journal of Biochemistry & Cell Biology* 29, no. 12 (December 1, 1997): 1325–41, [https://doi.org/10.1016/S1357-2725\(97\)00087-3](https://doi.org/10.1016/S1357-2725(97)00087-3).
- Teng C. S. and Hamilton T. H., “The Role of Chromatin in Estrogen Action in the Uterus, I. The Control of Template Capacity and Chemical Composition and the Binding of H3-Estradiol-17 Beta.,” *Proceedings of the National Academy of Sciences* 60, no. 4 (August 1968): 1410–17, <https://doi.org/10.1073/pnas.60.4.1410>.
- Thornton “Evolution of Vertebrate Steroid Receptors from an Ancestral Estrogen Receptor by Ligand Exploitation and Serial Genome Expansions | PNAS,” accessed August 18, 2022, <https://www.pnas.org/doi/abs/10.1073/pnas.091553298>.
- Toft, “The Interaction of Uterine Estrogen Receptors with DNA,” *Journal of Steroid Biochemistry* 3, no. 3 (April 1, 1972): 515–22, [https://doi.org/10.1016/0022-4731\(72\)90098-2](https://doi.org/10.1016/0022-4731(72)90098-2).
- Toft D. and Gorski J., “A Receptor Molecule for Estrogens: Isolation from the Rat Uterus and Preliminary Characterization.,” *Proceedings of the National Academy of Sciences* 55, no. 6 (June 1966): 1574–81, <https://doi.org/10.1073/pnas.55.6.1574>.
- Tremblay *et al.*, “Phosphorylation-Dependent Sumoylation Regulates Estrogen-Related Receptor- α and - γ Transcriptional Activity through a Synergy Control Motif,” *Molecular Endocrinology* 22, no. 3 (March 1, 2008): 570–84, <https://doi.org/10.1210/me.2007-0357>.

- Van Heel, “Angular Reconstitution: A Posteriori Assignment of Projection Directions for 3D Reconstruction,” *Ultramicroscopy* 21, no. 2 (January 1, 1987): 111–23, [https://doi.org/10.1016/0304-3991\(87\)90078-7](https://doi.org/10.1016/0304-3991(87)90078-7).
- Vanacker *et al.*, “Transcriptional Targets Shared by Estrogen Receptor-Related Receptors (ERRs) and Estrogen Receptor (ER) α , but Not by ER β ,” *The EMBO Journal* 18, no. 15 (August 2, 1999): 4270–79, <https://doi.org/10.1093/emboj/18.15.4270>.
- Vega and Kelly, “A Role for Estrogen-Related Receptor α in the Control of Mitochondrial Fatty Acid β -Oxidation during Brown Adipocyte Differentiation *,” *Journal of Biological Chemistry* 272, no. 50 (December 12, 1997): 31693–99, <https://doi.org/10.1074/jbc.272.50.31693>.
- Veesler *et al.*, “Maximizing the Potential of Electron Cryomicroscopy Data Collected Using Direct Detectors,” *Journal of Structural Biology* 184, no. 2 (November 1, 2013): 193–202, <https://doi.org/10.1016/j.jsb.2013.09.003>.
- Vernier *et al.*, “Estrogen-Related Receptors Are Targetable ROS Sensors,” *Genes & Development* 34, no. 7–8 (April 1, 2020): 544–59, <https://doi.org/10.1101/gad.330746.119>.
- Vu, Kraus, and Mertz, “Phosphorylation-Dependent Sumoylation of Estrogen-Related Receptor A1,” *Biochemistry* 46, no. 34 (August 1, 2007): 9795–9804, <https://doi.org/10.1021/bi700316g>.
- Wade, “A Brief Look at Imaging and Contrast Transfer,” *Ultramicroscopy* 46, no. 1 (October 1, 1992): 145–56, [https://doi.org/10.1016/0304-3991\(92\)90011-8](https://doi.org/10.1016/0304-3991(92)90011-8).
- WADE and FRANK, “ELECTRON MICROSCOPE TRANSFER FUNCTIONS FOR PARTIALLY COHERENT AXIAL ILLUMINATION AND CHROMATIC DEFOCUS SPREAD.,” *ELECTRON MICROSCOPE TRANSFER FUNCTIONS FOR PARTIALLY COHERENT AXIAL ILLUMINATION AND CHROMATIC DEFOCUS SPREAD.*, 1977.
- Wagner *et al.*, “SPHIRE-CrYOLO Is a Fast and Accurate Fully Automated Particle Picker for Cryo-EM,” *Communications Biology* 2, no. 1 (June 19, 2019): 1–13, <https://doi.org/10.1038/s42003-019-0437-z>.
- Walberg *et al.*, “Recruitment of the SWI-SNF Chromatin Remodeling Complex as a Mechanism of Gene Activation by the Glucocorticoid Receptor T1 Activation Domain | Molecular and Cellular Biology,” accessed August 18, 2022, <https://journals.asm.org/doi/full/10.1128/MCB.20.6.2004-2013.2000>.
- Wang *et al.*, “Estrogen-Related Receptor α (ERR α) and ERR γ Are Essential Coordinators of Cardiac Metabolism and Function,” *Molecular and Cellular Biology* 35, no. 7 (April 2015): 1281–98, <https://doi.org/10.1128/MCB.01156-14>.
- Wang, Ounjai, and Sigworth, “Streptavidin Crystals as Nanostructured Supports and Image-Calibration References for Cryo-EM Data Collection,” *Journal of Structural Biology* 164, no. 2 (November 1, 2008): 190–98, <https://doi.org/10.1016/j.jsb.2008.07.008>.
- Wang *et al.*, “Amino and PEG-Amino Graphene Oxide Grids Enrich and Protect Samples for High-Resolution Single Particle Cryo-Electron Microscopy,” *Journal of Structural Biology* 209, no. 2 (February 1, 2020): 107437, <https://doi.org/10.1016/j.jsb.2019.107437>.
- Wärnmark *et al.*, “Activation Functions 1 and 2 of Nuclear Receptors: Molecular Strategies for Transcriptional Activation,” *Molecular Endocrinology* 17, no. 10 (October 1, 2003): 1901–9, <https://doi.org/10.1210/me.2002-0384>.
- Wasmuth *et al.*, “Modulation of Androgen Receptor DNA Binding Activity through Direct Interaction with the ETS Transcription Factor ERG,” *Proceedings of the National Academy of Sciences* 117, no. 15 (April 14, 2020): 8584–92, <https://doi.org/10.1073/pnas.1922159117>.
- Wasmuth *et al.*, “Allosteric Interactions Prime Androgen Receptor Dimerization and Activation,” *Molecular Cell* 82, no. 11 (June 2, 2022): 2021-2031.e5, <https://doi.org/10.1016/j.molcel.2022.03.035>.

- Webb *et al.*, “The Nuclear Receptor Corepressor (N-CoR) Contains Three Isoleucine Motifs (I/LXXII) That Serve as Receptor Interaction Domains (IDs),” *Molecular Endocrinology* 14, no. 12 (December 1, 2000): 1976–85, <https://doi.org/10.1210/mend.14.12.0566>.
- Wei *et al.*, (2018). “Optimizing "self-wicking" nanowire grids. *Journal of structural biology*”, 202(2), 170–174, <https://doi.org/10.1016/j.jsb.2018.01.001>
- Weikum *et al.*, “A Structural Investigation into Oct4 Regulation by Orphan Nuclear Receptors, Germ Cell Nuclear Factor (GCNF), and Liver Receptor Homolog-1 (LRH-1),” *Journal of Molecular Biology* 428, no. 24, Part B (December 4, 2016): 4981–92, <https://doi.org/10.1016/j.jmb.2016.10.025>.
- Weikum *et al.*, “Glucocorticoid Receptor Control of Transcription: Precision and Plasticity via Allostery,” *Nature Reviews. Molecular Cell Biology* 18, no. 3 (March 2017): 159–74, <https://doi.org/10.1038/nrm.2016.152>.
- Weissenberger, Rene J. M. Henderikx, and Peter J. Peters, “Understanding the Invisible Hands of Sample Preparation for Cryo-EM,” *Nature Methods* 18, no. 5 (May 2021): 463–71, <https://doi.org/10.1038/s41592-021-01130-6>.
- Wende *et al.*, “PGC-1 α Coactivates PDK4 Gene Expression via the Orphan Nuclear Receptor ERR α : A Mechanism for Transcriptional Control of Muscle Glucose Metabolism,” *Molecular and Cellular Biology* 25, no. 24 (December 15, 2005): 10684–94, <https://doi.org/10.1128/MCB.25.24.10684-10694.2005>.
- Williams, S., Sigler, P. “Atomic structure of progesterone complexed with its receptor.” *Nature* 393, 392–396 (1998). <https://doi.org/10.1038/30775>
- Wilson *et al.*, “An Acetylation Switch Modulates the Transcriptional Activity of Estrogen-Related Receptor α ,” *Molecular Endocrinology* 24, no. 7 (July 1, 2010): 1349–58, <https://doi.org/10.1210/me.2009-0441>.
- Wurtz *et al.*, “A Canonical Structure for the Ligand-Binding Domain of Nuclear Receptors,” *Nature Structural Biology* 3, no. 1 (January 1996): 87–94, <https://doi.org/10.1038/nsb0196-87>.
- Wurtz *et al.*, “Three-Dimensional Models of Estrogen Receptor Ligand Binding Domain Complexes, Based on Related Crystal Structures and Mutational and Structure–Activity Relationship Data,” *Journal of Medicinal Chemistry* 41, no. 11 (May 1, 1998): 1803–14, <https://doi.org/10.1021/jm970406v>.
- Xu Eric H. and Lambert Millard H., “Structural Insights into Regulation of Nuclear Receptors by Ligands,” *Nuclear Receptor Signaling* 1 (June 15, 2003): e004, <https://doi.org/10.1621/nrs.01004>.
- Xu Eric H., “Family Reunion of Nuclear Hormone Receptors: Structures, Diseases, and Drug Discovery,” *Acta Pharmacologica Sinica* 36, no. 1 (January 2015): 1–2, <https://doi.org/10.1038/aps.2014.140>.
- Xu *et al.*, “Nuclear Receptor ERR α Contributes to Castration-Resistant Growth of Prostate Cancer via Its Regulation of Intratumoral Androgen Biosynthesis,” *Theranostics* 10, no. 9 (March 4, 2020): 4201–16, <https://doi.org/10.7150/thno.35589>.
- Yamamoto Keith R. and Alberts Bruce, “On the Specificity of the Binding of the Estradiol Receptor Protein to Deoxyribonucleic Acid,” *Journal of Biological Chemistry* 249, no. 22 (November 25, 1974): 7076–86, [https://doi.org/10.1016/S0021-9258\(19\)42077-2](https://doi.org/10.1016/S0021-9258(19)42077-2).
- Yamamoto Keith R. and Alberts Bruce, “The Interaction of Estradiol-Receptor Protein with the Genome: An Argument for the Existence of Undetected Specific Sites,” *Cell* 4, no. 4 (April 1, 1975): 301–10, [https://doi.org/10.1016/0092-8674\(75\)90150-6](https://doi.org/10.1016/0092-8674(75)90150-6).
- Yang *et al.*, “Nuclear Receptor Expression Links the Circadian Clock to Metabolism,” *Cell* 126, no. 4 (August 25, 2006): 801–10, <https://doi.org/10.1016/j.cell.2006.06.050>.

- Yao *et al.*, “Structural Insights into the Specificity of Ligand Binding and Coactivator Assembly by Estrogen-Related Receptor β ,” *Journal of Molecular Biology* 432, no. 19 (September 4, 2020): 5460–72, <https://doi.org/10.1016/j.jmb.2020.08.007>.
- Yi *et al.*, “Structure of a Biologically Active Estrogen Receptor-Coactivator Complex on DNA,” *Molecular Cell* 57, no. 6 (March 19, 2015): 1047–58, <https://doi.org/10.1016/j.molcel.2015.01.025>.
- Yi *et al.*, “Structural and Functional Impacts of ER Coactivator Sequential Recruitment,” *Molecular Cell* 67, no. 5 (September 7, 2017): 733–743.e4, <https://doi.org/10.1016/j.molcel.2017.07.026>.
- Yip *et al.*, “Atomic-Resolution Protein Structure Determination by Cryo-EM,” *Nature* 587, no. 7832 (November 2020): 157–61, <https://doi.org/10.1038/s41586-020-2833-4>.
- Yu *et al.*, “Orphan Nuclear Receptor Estrogen-Related Receptor- β Suppresses in Vitro and in Vivo Growth of Prostate Cancer Cells via P21WAF1/CIP1 Induction and as a Potential Therapeutic Target in Prostate Cancer,” *Oncogene* 27, no. 23 (May 2008): 3313–28, <https://doi.org/10.1038/sj.onc.1210986>.
- Young *et al.*, “Steroid Induction of Mouse Mammary Tumor Virus: Effect upon Synthesis and Degradation of Viral RNA,” *Journal of Virology* 21, no. 1 (January 1977): 139–46.
- Zhang *et al.*, “Estrogen-Related Receptors Stimulate Pyruvate Dehydrogenase Kinase Isoform 4 Gene Expression*,” *Journal of Biological Chemistry* 281, no. 52 (December 29, 2006): 39897–906, <https://doi.org/10.1074/jbc.M608657200>.
- Zhang *et al.*, “3.3 Å Cryo-EM Structure of a Nonenveloped Virus Reveals a Priming Mechanism for Cell Entry,” *Cell* 141, no. 3 (April 30, 2010): 472–82, <https://doi.org/10.1016/j.cell.2010.03.041>.
- Zhang *et al.*, “An Improved Cryo-FIB Method for Fabrication of Frozen Hydrated Lamella,” *Journal of Structural Biology* 194, no. 2 (May 1, 2016): 218–23, <https://doi.org/10.1016/j.jsb.2016.02.013>.
- Zhao *et al.*, “Nuclear Receptors Rock around the Clock,” *EMBO Reports* 15, no. 5 (May 2014): 518–28, <https://doi.org/10.1002/embr.201338271>.

ANNEXES

Résultats et commentaires sur l'expression et la purification des protéines

Les test d'expression pour les protéines produites par coinfection de cellules d'insectes sont réalisés en faisant varier le taux d'infection de chaque cellule pour chacune des constructions.

Test de coexpression du complexe mouse $ERR\alpha$ 1-422-PGC-1 α 1-472

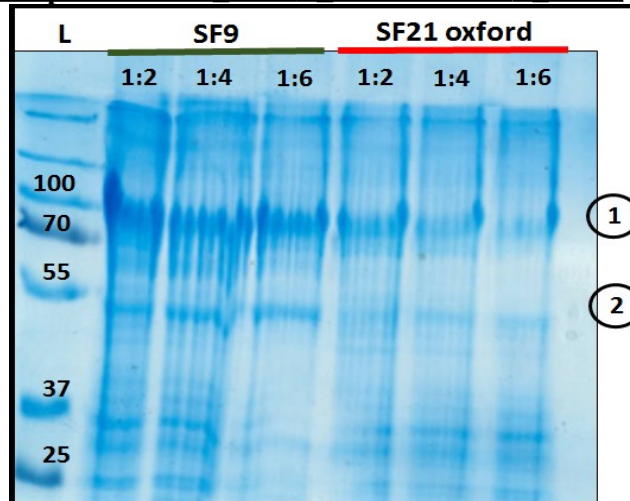


Figure S1 : Test d'expression du complexe mouse $ERR\alpha$ 1-422-PGC-1 α 1-472

Légende :

Les pistes correspondent aux protéines purifiées par affinité en cellules SF9 (vert) ou SF21 Oxford (rouge). Les ratios représentent la quantité du virus porteur de l'ADN codant pour la construction mouse_ $ERR\alpha$ 1-422 sur la quantité de virus porteur de l'ADN codant pour la construction PGC-1 α 1-472 par cellule infectée.

1/ PGC-1 α 1-472 a la particularité de migrer, quelque soit la construction, plus haut que son poids moléculaire attendu, ici 100 kDa. 2/ mouse_ $ERR\alpha$ migre à 46 kDa environ.

La condition correspondant à une infection de 72 h est nécessaire à l'expression de PGC-1 α dans le cadre du complexe (**Fig. S1**). De plus, la souche sf9 présente une expression supérieure du complexe par comparaison à la souche sf21 oxford. La condition « sf9-1:2 » est donc celle qui est conservée puisqu'elle permet d'atteindre le meilleur niveau de production des deux protéines en économisant du virus.

Test d'expression des constructions PGC-1 α 1-635 et 1-798

Les test d'expression des deux protéines ont été réalisés en faisant varier la durée d'infection de 48 à 72h.

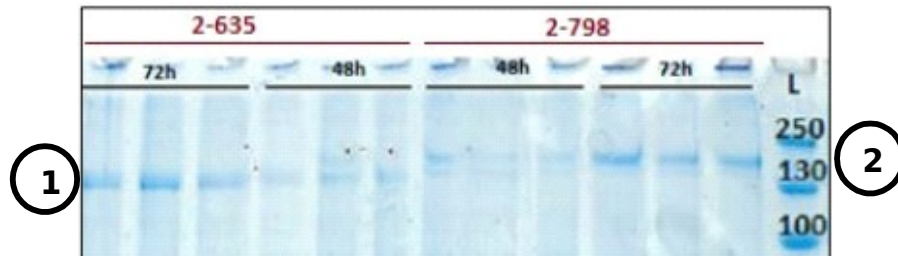


Figure S2 : Test d'expression des constructions PGC-1 α 1-635 et PGC-1 α 1-798

Légende :

1/ Construction PGC-1 α 1-635 exprimée après 48 et 72h d'infection. Elle migre à environ 150 kDa.

2/ Construction PGC-1 α 1-798 exprimée après 48 et 72h d'infection. Elle migre à environ 200 kDa.

Un temps d'incubation de 72h est retenu compte-tenu des résultats du test d'expression. Les fragments sont dans un premier temps produit seuls, mais la quantité produite étant très faible, il est choisi de coproduire ces fragments avec le récepteur nucléaire ERR α , comme ce fut le cas pour la construction précédente.

Test d'expression de la protéine PCAF

La protéine PCAF est produite en faisant varier le taux d'infection de 30 virions par cellules jusqu'à 90 virions par cellule sur une période de 48 à 72h.

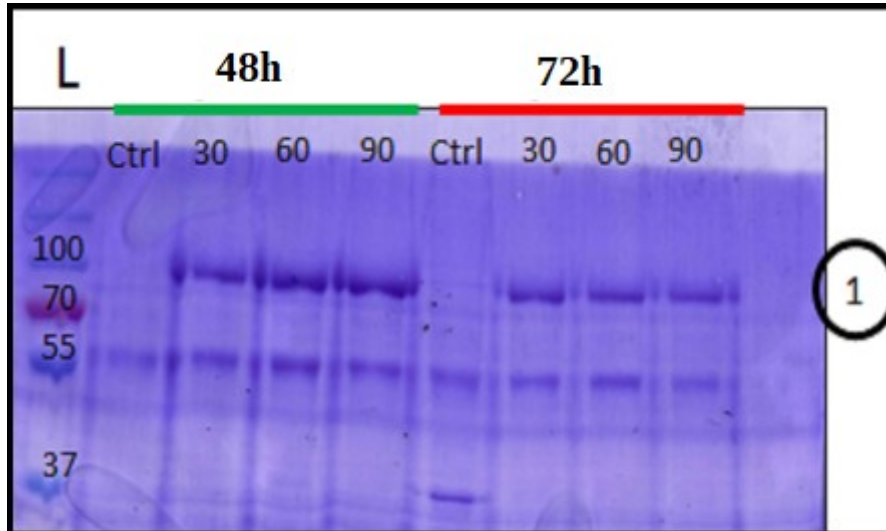


Figure S3 : Test d'expression de la protéine PCAF

Légende :

1/ Protéine PCAF migrant à 90 kDa.

48 h d'expression permettent une plus haute expression de la construction PCAF. Une dose de 30 virions par cellule semble suffisante. Un contaminant est observé à 50 kDa (Fig. S3). Ce dernier n'a pas encore été identifié par spectrométrie de masse.

Test d'expression de la protéine p300

La protéine p300 est produite en faisant varier le taux d'infection de 1 à 5 virions par cellule.

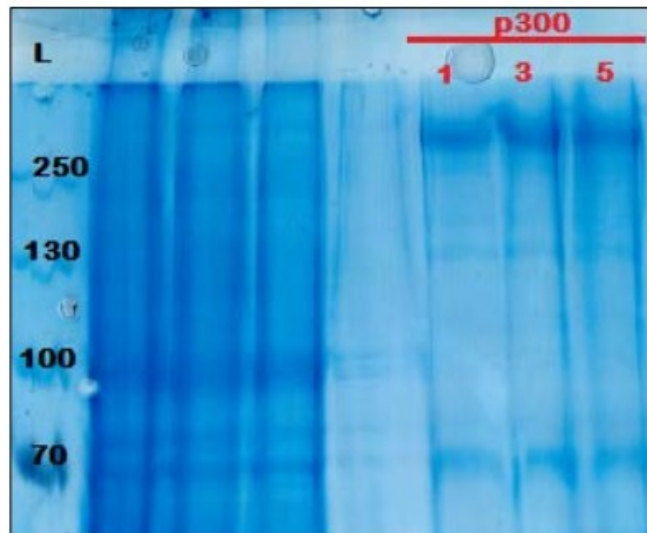


Figure S4 : Test d'expression de la protéine p300

Légende :

La protéine p300 est exprimée grâce à des rapports virions/cellules allant de 1 à 5. La bande observée à 265 kDa correspond à la protéine p300.

Au vu des résultats de la migration sur gel de polyacrylamides réalisée en conditions dénaturantes par l'utilisation de sodium dodecyl sulfate (SDS), une infection utilisant un rapport de 3 virions par cellule d'insecte est choisie pour une durée d'infection de 48 h.

Purification du complexe mouse ERR α 1-422-PGC-1 α 1-472

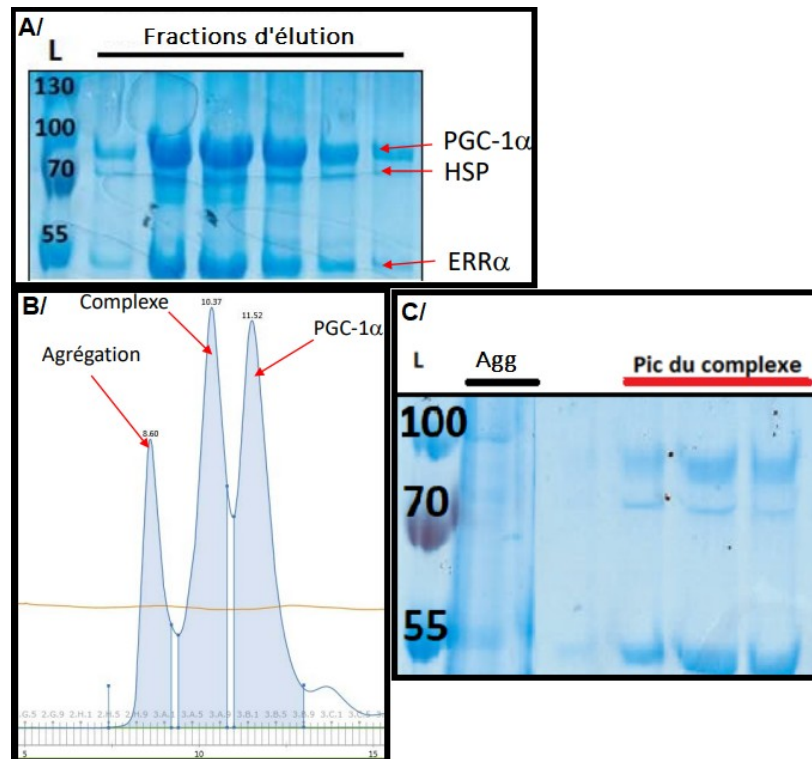


Figure S5 : Purification du complexe mouse ERR α 1-422 et PGC-1 α 1-472 obtenu par co-infection par baculovirus de cellules d'insectes

Légende :

A/ Migration sur gel SDS des fractions après purification sur colonne StrepTrapTM HP 5 mL. Les protéines PGC et ERR sont indiquées et migrent respectivement à environ 90 et 46 kDa. Une bande correspondant à de la contamination par la protéine de choc thermique de masse moléculaire 70 kDa (HSP70) est observable à 70 kDa, cette dernière étant capable de se lier aux PGC-1 α en excès d'après nos analyses en spectrométrie de masse native.

B/ Chromatogramme résultant de la SEC du complexe. Un premier pic d'agrégats est observable à un volume de rétention de 8.6 mL. Il correspond à une partie du complexe agrégé et contaminé par de l'ADN cellulaire d'après analyse au Nanodrop. Le complexe intact est élué à 10.37 mL et n'est pas contaminé par de l'ADN après analyse au Nanodrop en observant le ratio A260/A280, lequel est inférieur à 0.7. Le PGC en excès élué à 11.52 mL.

C/ Migration sur gel SDS du pic d'agrégats (Agg) et du pic correspondant au complexe. Cette fois encore, ERR et PGC sont observables ainsi qu'une bande correspondant au contaminant HSP70.

Le pic d'agrégats observé peut-être retiré en amont de la SEC par une centrifugation de l'éluât issu de l'étape de purification précédente réalisée à 4°C et 11 000 g. Cela permet d'éviter un recouvrement des pics lorsque la concentration de complexes non agrégés est faible et permet l'obtention d'un meilleur échantillon qualitativement. Les tentatives de réduire ce pic d'agrégation en changeant de souches cellulaires ou en jouant sur les paramètres de lyse, notamment par l'augmentation de la concentration et du temps d'action de la benzonase nucléase ou la variation du pH, n'ont pas permis de réduire

significativement son importance. L'augmentation de la quantité de cellules lysées en même temps à cependant produit un résultat intéressant puisque, proportionnellement, le pic de complexes fonctionnels est plus important par comparaison avec celui des protéines agrégées. Enfin, une purification à un pH 6.5 en tampon sodium phosphate dissocie le complexe et permet de ne purifier que PGC-1 α 1-472 lors de l'étape de purification par affinité.

Purification de PGC-1 α 1-635

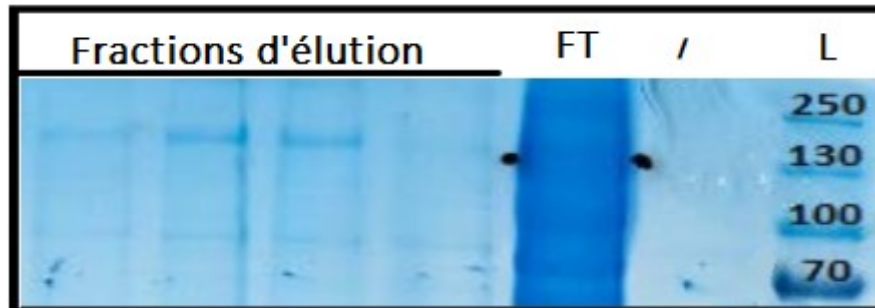


Figure S6 : Purification de la construction PGC-1 α 1-635

Légende :

« FT » correspond aux fractions non retenues sur colonne alors que « fractions d'élution » correspond aux fractions éluées par compétition à l' α -desthiobiotine. PGC-1 α est observé dans les fractions d'élution à environ 150 kDa.

La purification de PGC-1 α 1-635 par son étiquette streptavidine permet d'obtenir la protéine en faible quantité, de l'ordre de quelques microlitres à une concentration de 0,07 mg/mL (**Fig. 33**). Les tests de colyse en présence de mouse_ERR α _1-422 ne permettront pas de réduire la rapide perte de PGC-1 α 1-635 par agrégation au cours du temps. Il est possible que le tampon ne soit pas adapté dès la lyse ou que la coproduction soit nécessaire pour éviter l'agrégation *in cellulo*. Néanmoins, les essais de coproduction effectués n'ont pas augmenté significativement la quantité produite par opposition aux constructions plus petites. Des résultats similaires ont été obtenus avec la construction PGC-1 α 1-798.

Purification de PCAF

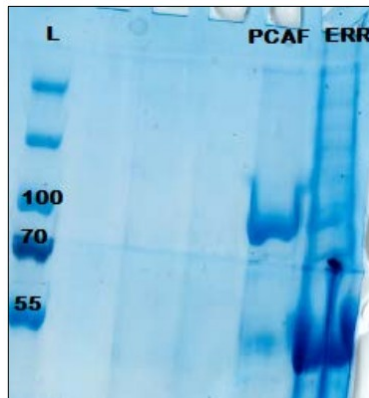


Figure S7 : Purification de PCAF

Légende :

« L » correspond au marqueur de poids moléculaire. PCAF migre à 85 kDa après sa purification par chromatographie d'affinité au peptide FLAG. La piste la plus à droite correspond à la protéine mouse_*ERR* α _1-422 également purifiée le même jour par IMAC.

Purification de PGC-1 α 74-219 et PGC-1 α 138-219

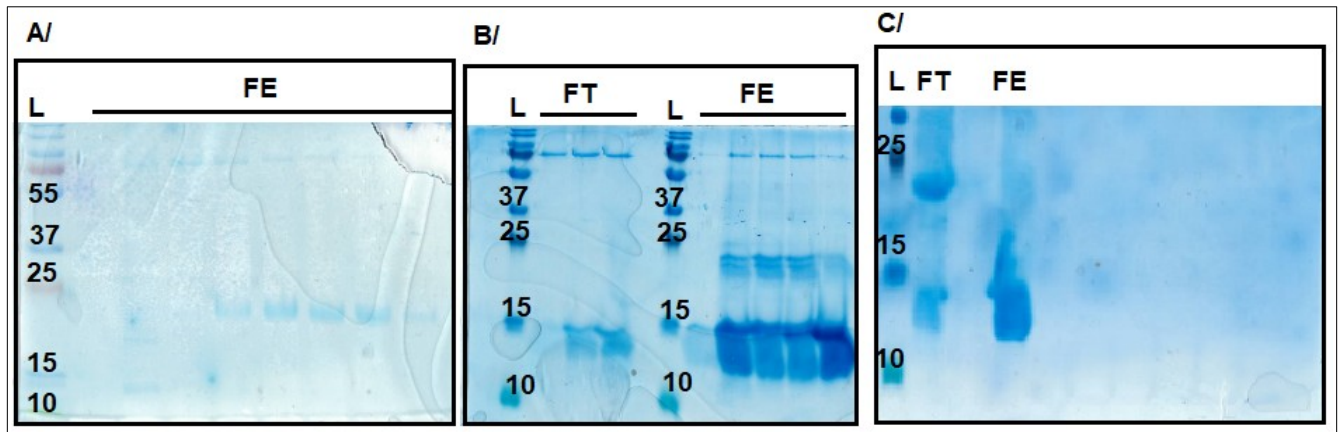


Figure S8 : Purification de PGC-1 α 74-219 & PGC-1 α 138-219

Légende :

« FT » correspond aux fractions non retenues sur colonne. « FE » correspond aux fractions éluées. **A/** IMAC de la protéine PGC-1 α 138-219. Cette dernière est observable à environ 20 kDa. **B/** RIMAC de PGC-1 α 138-219. La protéine est observable à environ 14 kDa dans les fractions non éluées. Les fractions éluées révèlent une bande correspond à la protéine de fusion Fh8 à environ 14 kDa et la protéine non coupée PGC-1 α 138-219 à environ 24 kDa. **C/** RIMAC de la protéine PGC-1 α 74-219. La protéine de fusion Fh8 est visible à environ 14 kDa dans la fraction d'élution alors que la protéine PGC-1 α 74-219 est visible dans la fraction non retenue sur colonne à environ 24 kDa.

Purification des protéines human $ERR\alpha$ 1-423 et PGC-1 α 1-285

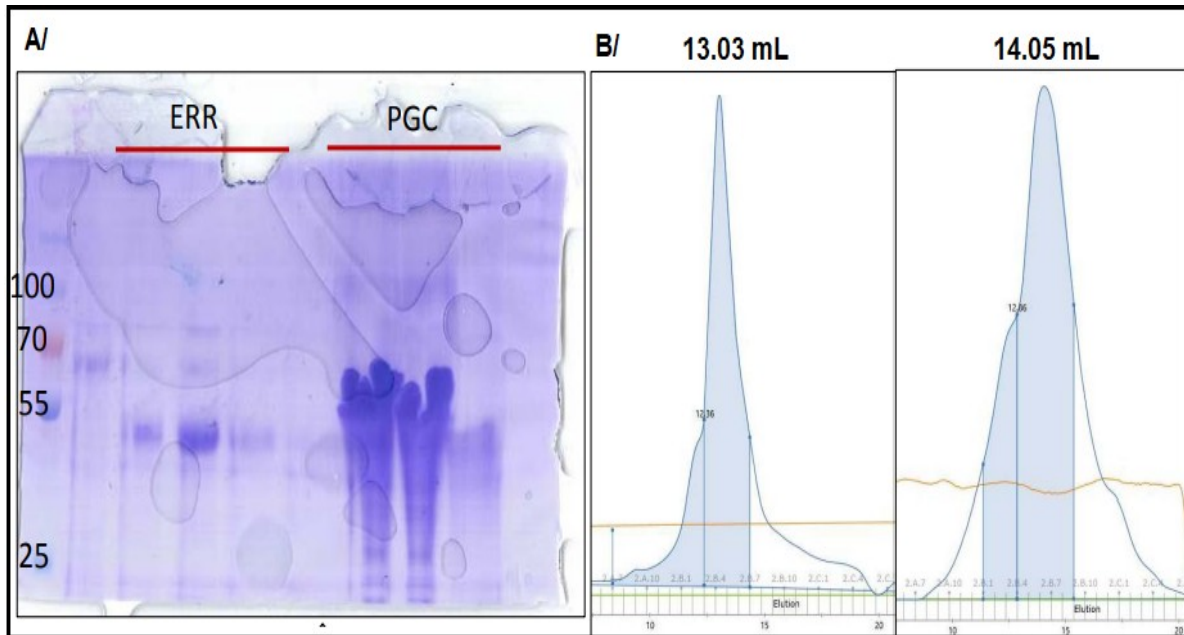


Figure S9 : Purification de human $ERR\alpha$ 1-423 et PGC-1 α 1-285

Légende :

A/ Ce gel présente human $ERR\alpha$ 1-423 avant sa coupure par la protéase du TEV dans la piste non annotée. Les trois fractions suivantes, annotées « ERR » correspondent à la protéine après coupure et chromatographie échangeuse d'ions. Les pistes annotées « PGC » correspondent à la protéine PGC-1 α 1-285 après passage sur chromatographie échangeuse d'ions. Ces échantillons n'ont pas subi de dénaturation thermique avant dépôt sur gel, car cela tend à précipiter les protéines et empêche leur dépôt de façon efficace dans les puits du gel SDS.

B/ Chromatogrammes obtenus après passage des protéines human $ERR\alpha$ 1-423 (à gauche) et PGC-1 α 1-285 (à droite) sur colonne de SEC. Les volumes de rétention de chacune des protéines sont indiqués.

Purification de p300

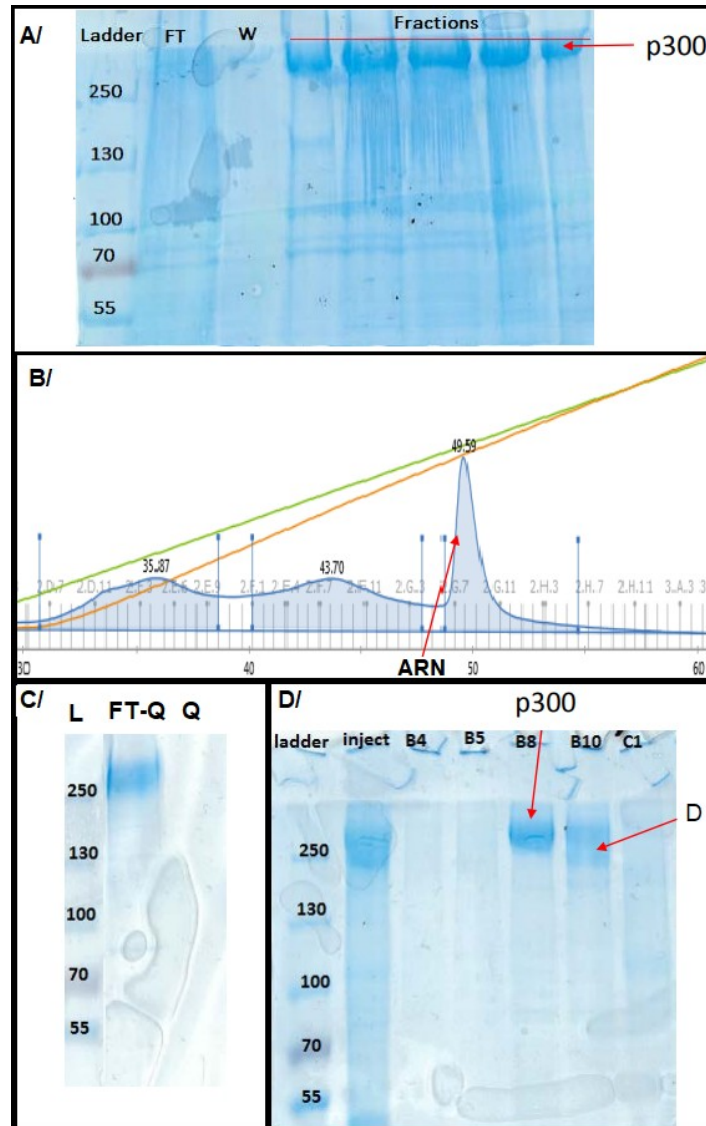


Figure S10 : Purification de p300

« L » et « Ladder » correspondent au marqueur de poids moléculaire, « FT » correspond à fraction non retenue sur colonne, « W » correspond à lavage, « Fractions » correspond aux fractions éluées, « FT-Q » correspond à la fraction non retenue sur colonne lors du passage sur colonne échangeuse d'ions, « Q » correspond à la fraction retenue sur colonne échangeuse d'ions, « D » correspond à des produits de dégradation.

Légende :

A/ Étape d'IMAC de p300. La protéine migre vers 265 kDa.

B/ Chromatogramme de l'étape d'échange d'ions. La protéine ne se fixe pas sur la colonne, mais une fraction marquée « ARN » est retenue sur la colonne et correspond, d'après l'analyse du ratio de l'absorbance à 230 et 260 nm au Nanodrop à un ARN contaminant. Ce dernier n'a pas pu être identifié à ce jour.

C/ Gel montrant p300 après son passage sur colonne échangeuse d'ions. P300 n'a pas été retenue sur la colonne.

D/ p300 après SEC. On observe que certaines fractions contiennent des produit de dégradation. Cette dernière étape nous permet de nous en débarrasser.

La purification de p300 couplée aux expériences de spectrométrie de masse à partir des bandes découpées sur gel révèle que la protéine est très sensible à la dégradation. Les tests d'optimisation de tampon réalisés montrent que p300 demeure stable plus longtemps en tampon Tris HCl ou Imidazole à un pH compris entre 7.0 et 8.0. Dans les autres tampons testés, la protéine tend à s'agréger rapidement. De plus, la protéine ne supporte pas les concentrations trop basses en sel (inférieure à 220 mM). L'étape de SEC permet d'isoler de façon consistante la construction voulue. La nature des acides ribonucléiques qui se lient à la protéine n'a pas pu être identifiée à ce jour malgré des tentatives de les isoler et de les faire migrer sur gel d'agarose.

Formation du complexe human $ERR\alpha$ -1-423-PGC-1 α 1-285-BE42embedded

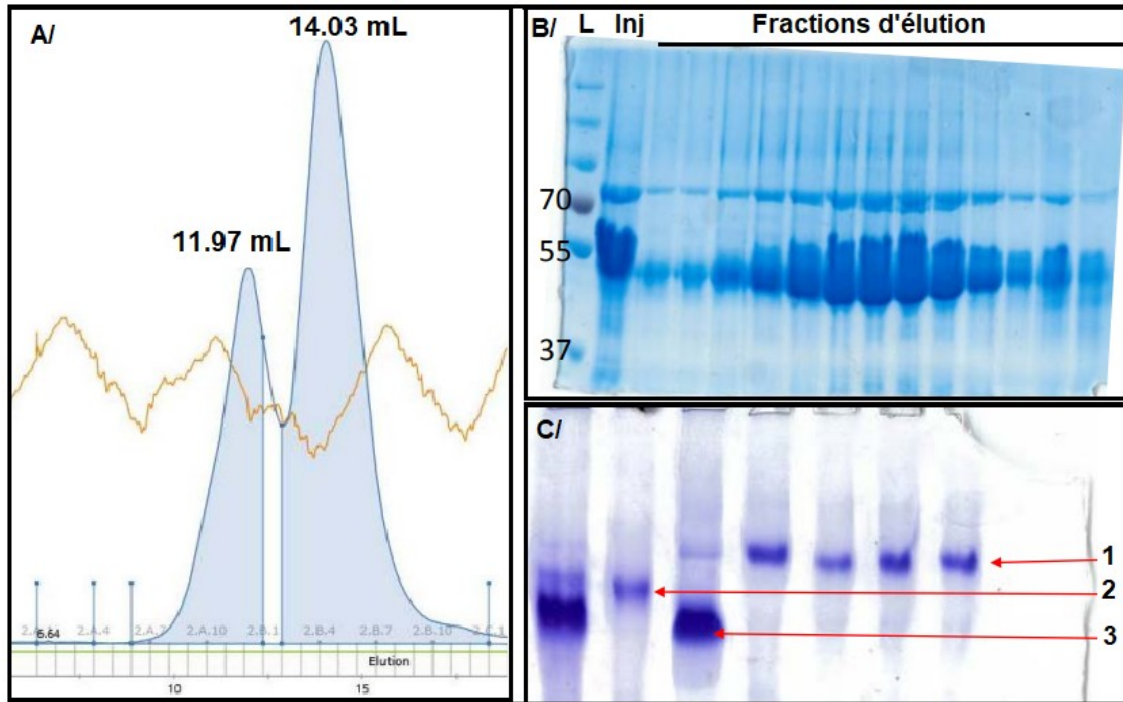


Figure S11 : Formation du complexe ERR-PGC-BE42embedded

Légende :

A/ Chromatogramme issue de la SEC du complexe human $ERR\alpha$ -1-423-PGC-1 α 1-285. Le pic du complexe est à 11,97 mL par comparaison avec ERR dont le volume de rétention est à 13 mL et 14 mL dans le cas de PGC dont l'excès est visible sur ce chromatogramme.

B/ Gel SDS des fractions de gel filtration. « inj » correspond à la fraction injectée sur la colonne avant la SEC. On observe effectivement les deux protéines ERR et PGC qui coéluent en formant le complexe. On observe également les traces de HSP70 qui demeurent parfois à cette étape de purification.

C/ Gel natif montrant en « 1 » le complexe ERR-PGC-be42embedded, en « 2 » le complexe ERR-PGC et en 3 la protéine « ERR ». Il y apparaît clairement que le complexe est formé et qu'il n'y a pas de dissociation de ce dernier en encore avec le gel.

Formation du complexe mouse $ERR\alpha$ 1-422-PGC-1 α 1-472-Be42embedded

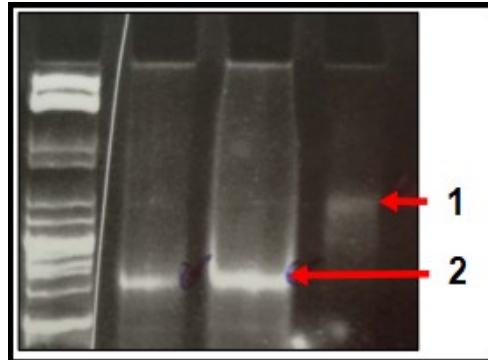


Figure S12 : Contrôle de la formation du complexe mouse $ERR\alpha$ 1-422-PGC-1 α 1-472-Be42embedded

Légende :

Gel natif de polyacrylamide (6%) Tris CAPS à pH 8.2 coloré au BET.

La bande annotée « 1 » correspond au complexe mouse_ $ERR\alpha$ _1-422-PGC1-472-Be42embedded pour lequel un retard est observé, alors que les bandes annotées « 2 » correspondent au complexe mouse_ $ERR\alpha$ _1-422-Be42embedded.

Validation statistique des expériences de HDX pour ERR α

Différence d'incorporation (en %) entre ERR α + PGC1 α et ERR α pour chaque peptide et temps de deuteration :

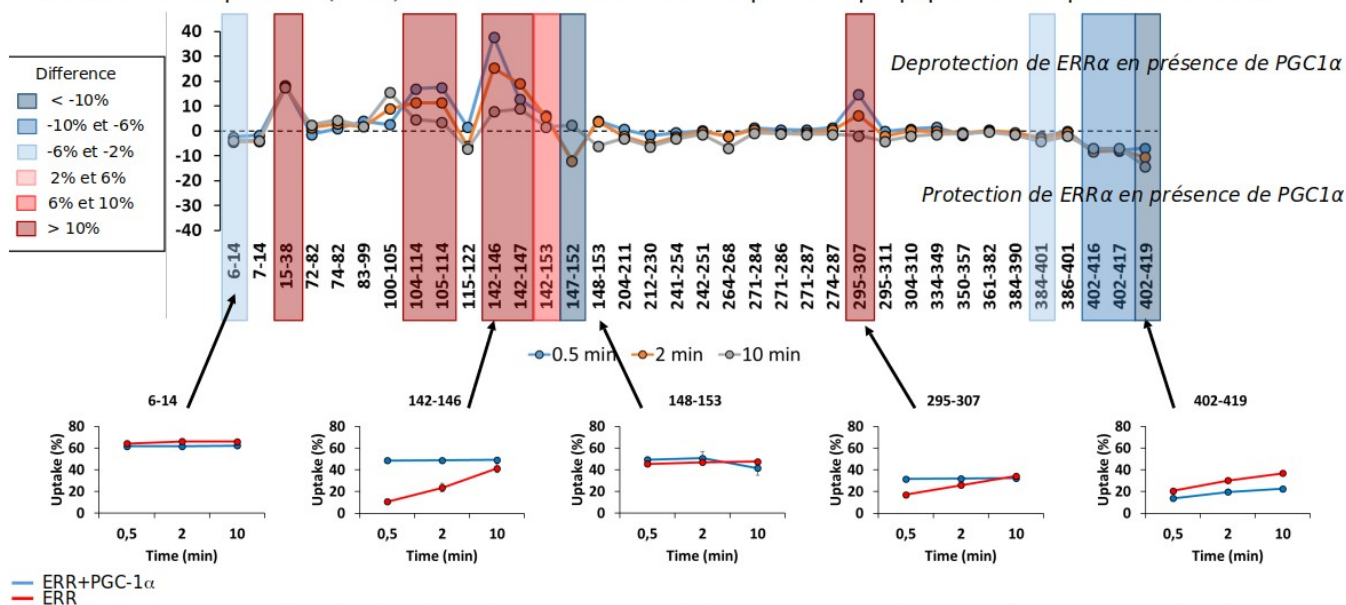


Figure S13 : Validation statistique des résultats de HDX

Légende :

Les peptides analysés par spectrométrie de masse sont indiqués par les numéros de résidus dans la séquence avec le pourcentage d'incorporation du deutérium au cours du temps. Les régions colorées indiquent des résultats statistiquement significatifs en terme de protection (en bleu) et de déprotection (en rouge) en présence du coactivateur PGC-1 α . L'évolution de certains peptides au cours du temps dans les deux conditions est indiquée par des courbes respectivement bleue pour le complexe et rouge pour ERR α .

Récapitulatif des réticulations observées par XL-MS

Dans l'ensemble des tableaux suivants, les réticulations obtenues en duplicats sont indiquées en bleu et celle en triplicats en orange.

protein 1	protein 2	site 1	site 2	Replicate (DSAU)		
				1	2	3
ERR	ERR	K53	K191	x	x	
ERR	ERR	K78	K191	x		x
ERR	ERR	K106	K106	x	x	x
ERR	ERR	K106	K131	x		x
ERR	ERR	K106	K140	x	x	
ERR	ERR	K131	K131	x	x	x
ERR	ERR	K140	K140	x	x	x
ERR	ERR	K148	K148	x		x
ERR	ERR	K191	K191	x	x	

Table S1 : Résultats de XL-MS à pH 8.5 pour ERR α

protein 1	protein 2	site 1	site 2	replicate		
				1	2	3
ERR	ERR	S21	K106	x		x
ERR	ERR	S21	K131	x	x	x
ERR	ERR	S21	K148	x	x	x
ERR	ERR	S21	K191	x	x	x
ERR	ERR	K53	K148	x	x	x
ERR	ERR	K53	K191	x	x	x
ERR	ERR	K70	K127	x	x	
ERR	ERR	K70	K191	x	x	x
ERR	ERR	K78	K78	x	x	
ERR	ERR	K78	K148	x	x	x
ERR	ERR	K78	K191	x	x	x
ERR	ERR	K106	K106	x	x	x
ERR	ERR	K106	K140	x	x	x
ERR	ERR	K106	K148	x	x	x
ERR	ERR	K127	K131	x		x
ERR	ERR	K127	K191	x		x
ERR	ERR	K131	K131	x	x	x
ERR	ERR	K131	K148	x	x	x
ERR	ERR	K131	K191	x	x	x
ERR	ERR	K131	T192	x		x
ERR	ERR	K140	K140	x	x	x
ERR	ERR	K140	K191	x	x	x
ERR	ERR	K148	K148	x	x	x
ERR	ERR	K148	K191	x	x	x
ERR	ERR	K148	K396	x		x
ERR	ERR	K148	K409	x	x	x
ERR	ERR	K191	K191	x	x	x

Table S2 : Résultats de XL-MS à pH 6.5 pour ERR α

protein 1	protein 2	site 1	site 2	Replicate (DSAU)		
				1	2	3
PGC1	PGC1	K289	K329		x	x

Table S3 : Résultats de XL-MS à pH 8.5 pour PGC-1 α

protein 1	protein 2	site 1	site 2	Replicate (DSBU)			Replicate (DSAU)		
				1	2	3	1	2	3
ERR	ERR	S21	K78				x	x	
ERR	ERR	S28/S29	K78	x	x			x	x
ERR	ERR	K131	K131				x	x	x

Table S4 : Résultats de XL-MS à pH 8.5 pour ERR α -ADN

protein 1	protein 2	site 1	site 2	Replicate (DSAU)		
				1	2	3
ERR	ERR	K53	K191	x	x	
ERR	ERR	K78	K191	x	x	x
ERR	ERR	K106	K148	x	x	
ERR	ERR	K131	K131	x	x	x
ERR	ERR	K131	K148	x	x	
ERR	ERR	K131	K409	x	x	
ERR	ERR	K148	K191	x	x	x
ERR	ERR	K191	K191	x	x	
PGC1	ERR	K289	K405	x	x	x
PGC1	ERR	K297	K396	x	x	
PGC1	ERR	K297	K405	x	x	x
PGC1	ERR	S300	K131		x	x
PGC1	ERR	K306	K106		x	x
PGC1	ERR	K306	K405	x	x	x
PGC1	ERR	K329	K191	x	x	x
PGC1	ERR	K340	K106	x		x

Table S5 : Résultats de XL-MS à pH 8.5 pour ERR α -PGC-1 α

protein 1	protein 2	site 1	site 2	pH 8.5			pH 6.5		
				replicate			replicate		
				1	2	3	1	2	3
PGC1	PGC1	K289	K306	x	x	x	x	x	x
PGC1	PGC1	K289	K329	x	x	x	x		x
PGC1	PGC1	K306	K329	x	x	x	x	x	x
PGC1	PGC1	K329	K329				x	x	x
PGC1	PGC1	K329	K340		x	x		x	
ERR	ERR	S21	K148				x	x	x
ERR	ERR	K78	K148	x			x	x	x
ERR	ERR	K78	K191	x		x	x	x	x
ERR	ERR	Y91	K191					x	x
ERR	ERR	K106	K106	x			x	x	
ERR	ERR	K106	K131	x	x	x	x		
ERR	ERR	K106	K140	x	x	x			x
ERR	ERR	K106	K148	x		x	x	x	x
ERR	ERR	K127	K191	x		x	x	x	
ERR	ERR	K131	K131	x		x	x	x	x
ERR	ERR	K131	K148	x		x	x	x	x
ERR	ERR	K131	K191	x	x	x	x	x	x
ERR	ERR	K131	T192	x		x			
ERR	ERR	K131	K396	x			x	x	x
ERR	ERR	K131	K405	x		x	x	x	x
ERR	ERR	K131	K409	x		x		x	x
ERR	ERR	K140	K140				x	x	x
ERR	ERR	K140	K191				x	x	x
ERR	ERR	K148	K148	x			x	x	x
ERR	ERR	K148	K191	x		x	x		x
ERR	ERR	K148	K396				x	x	
ERR	ERR	K148	K409	x			x	x	x
ERR	ERR	K191	K191	x	x	x	x	x	x
ERR	ERR	K191	K332				x		x
PGC1	ERR	K289	K191	x	x	x	x		
PGC1	ERR	K289	T192			x			
PGC1	ERR	K289	K405	x	x	x			
PGC1	ERR	K289	K409				x	x	x
PGC1	ERR	S293	K409			x			
PGC1	ERR	K297	K405	x		x			
PGC1	ERR	K297	K409	x		x			
PGC1	ERR	K299	K405	x	x	x			
PGC1	ERR	S300/K306	K405	x	x				
PGC1	ERR	K306	K131					x	x
PGC1	ERR	K306	K148				x	x	
PGC1	ERR	K306	K396	x		x			
PGC1	ERR	K306	K405	x	x	x	x		x
PGC1	ERR	K306	K409	x	x	x			x
PGC1	ERR	K329	K78	x			x		x
PGC1	ERR	K329	K131				x		x
PGC1	ERR	K329	K191	x		x	x	x	x
PGC1	ERR	K329	T192	x		x			

Table S6 : Résultats de XL-MS pour ERR α -PGC-1 α , seconde expérience

protein 1	protein 2	site 1	site 2	Replicate (DSBU)			Replicate (DSAU)		
				1	2	3	1	2	3
ERR	ERR	K53	K148	x	x	x		x	
ERR	ERR	K53	K191	x	x	x	x	x	x
ERR	ERR	K78	K191	x	x	x		x	
ERR	ERR	S89	K191	x	x	x			
ERR	ERR	K106	K140	x	x	x			
ERR	ERR	K106	K148	x	x		x		
ERR	ERR	K127	K191	x		x		x	
ERR	ERR	K131	K131	x	x	x	x	x	x
ERR	ERR	K131	K191		x	x			x
ERR	ERR	K140	K140	x	x	x			
ERR	ERR	K148	K191	x	x	x	x	x	x
ERR	ERR	K191	K191	x	x	x	x	x	
PGC1	ERR	K289	K191	x	x	x	x	x	
PGC1	ERR	K289	K405				x		x
PGC1	ERR	K289	K409				x	x	
PGC1	ERR	K297	K405				x	x	x
PGC1	ERR	K299	K409				x	x	x
PGC1	ERR	S300	K405				x	x	x
PGC1	ERR	S300	K409				x	x	x
PGC1	ERR	K306	K396			x	x	x	x
PGC1	ERR	K306	K405	x		x	x	x	x
PGC1	ERR	K306	K409				x	x	x
PGC1	ERR	K329	K78	x	x	x	x		x
PGC1	ERR	K329	K191	x				x	x
PGC1	ERR	K340	K405				x	x	x
PGC1	PGC1	K306	K329					x	x

Table S7 : Résultats de XL-MS à pH 8.5 pour ERR α -PGC-1 α -ADN

Résumé

Les récepteurs nucléaires sont une famille de facteurs de transcription impliqués dans la régulation et l'adaptation de l'expression génétique. Le récepteur apparenté au récepteur aux œstrogènes alpha ($ERR\alpha$) est un récepteur nucléaire orphelin constitutivement actif régulant le métabolisme énergétique et dont la dérégulation est liée à la prolifération tumorale.

Cette thèse caractérise l'interaction de son domaine de liaison à l'ADN (DBD) avec ses éléments de réponses (ERRE). Elle montre que les DBDs présentent une interface de dimerisation asymétrique stable et ancestrale que ne possède pas les récepteurs oxostéroïdiens. Elle montre aussi la façon dont la forme de l'ADN dicte la liaison de ce domaine. De plus, elle s'intéresse au rôle du coactivateur-1-alpha du récepteur gamma activé par les proliférateurs du peroxysoxe ($PGC-1\alpha$) dans l'activité transcriptionnelle de $ERR\alpha$. Elle développe le fait que $PGC-1\alpha$ permet l'assise du second DBD sur les ERRE. Elle montre par des expériences d'échange hydrogène-deutérium, de réticulations couplées à de la spectrométrie de masse et des analyses de dispersions induites par flux que le coactivateur permet le renforcement de la liaison à l'ADN par des effets allostériques sur les LBDs et DBDs. Enfin, l'étude structurale par cryo-microscopie électronique (cryo-EM) met en lumière l'architecture globale du récepteur nucléaire, ainsi que l'orientation de ce dernier sur l'ADN. Elle met également en avant l'asymétrie des deux sous-unités de $ERR\alpha$ au sein du dimère actif.

Mots-clefs : Récepteurs nucléaires, $ERR\alpha$, $PGC1-\alpha$, DBD, structure, spectrométrie de masse, diffraction aux rayons X, cryo-microscopie

Nuclear receptors are a family of transcription factors involved in genetic regulation and adaptability. The estrogen-related receptor alpha ($ERR\alpha$) is a constitutively active orphan nuclear receptor which regulates the energetic metabolism and for which dysregulation leads to cancer proliferation.

This thesis characterizes the DNA binding domain (DBD) interaction with response elements (ERRE). It shows DBD have a stable and ancestral asymmetric dimerization interface which is not encounter in oxo-steroid receptors. It shows that DNA' shape dictates the domain binding. This thesis get interested by the role of the peroxisome proliferator-activated receptor γ coactivator 1α ($PGC-1\alpha$) in $ERR\alpha$ transcriptional activity. It is shown that $PGC-1\alpha$ helps the binding of the second DBD on the ERRE. By Deuterium-Hydrogen exchanges, cross-linked mass spectrometry experiments and Flow-Induced Dispersion Analysis, it is shown that $PGC-1\alpha$ increases the DNA binding affinity of the receptor by allosteric effects on the LBDs and DBDs. Finally, the cryo-electron-microscopy (cryo-EM) structural analysis put in light the global architecture of the nuclear receptor and its orientation compared to the DNA. The two sub-units of $ERR\alpha$ appears to be asymmetric inside the active dimer.

Key-words : Nuclear receptors, $ERR\alpha$, $PGC-1\alpha$, DBD, structure, mass spectrometry, X-ray diffraction, cryo-electron-microscopy