



**HAL**  
open science

# Dimérisation du domaine transmembranaire des récepteurs de la famille ErbB/HER. Etude par simulations de dynamique moléculaire.

Oumarou Samna Soumana

► **To cite this version:**

Oumarou Samna Soumana. Dimérisation du domaine transmembranaire des récepteurs de la famille ErbB/HER. Etude par simulations de dynamique moléculaire.. domain\_other. Université d'Orléans, 2007. Français. NNT: . tel-00329923

**HAL Id: tel-00329923**

**<https://theses.hal.science/tel-00329923>**

Submitted on 13 Oct 2008

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# THESE

PRESENTEE

A L'UNIVERSITE D'ORLEANS

POUR OBTENIR LE GRADE DE

DOCTEUR DE L'UNIVERSITE D'ORLEANS

Discipline :

**Biophysique Moléculaire**

PAR

**Oumarou SAMNA SOUMANA**

**Dimérisation du domaine transmembranaire des récepteurs de la famille ErbB/HER. Etude par simulations de dynamique moléculaire.**

Soutenue le 12 octobre 2007

## MEMBRES DU JURY :

Mme M. GENEST	Directeur de Recherche, CNRS	Directeur de Thèse
M. L. MORIN-ALLORY	Professeur, Université d'Orléans	Président du jury
Mme C. ETCHEBEST	Professeur, Université Paris 7	Rapporteur
M. S. CROUZY	Chargé de Recherche, CEA Grenoble	Rapporteur
M. J. STURGIS	Professeur, Université Aix Marseille 2	Examineur
Mme V. BESWICK	Maître de Conférence, Université d'Evry	Examineur
M. N. GARNIER	Maître de Conférence, Université d'Orléans	Examineur

## Remerciements :

*Le présent manuscrit est le fruit d'un travail réalisé au Centre de Biophysique Moléculaire, CNRS Orléans dirigé par le Docteur Jean-Claude Beloeil.*

*J'adresse tout d'abord mes sincères remerciements au Docteur Veronica Beswick, aux Professeurs James Sturgis et Luc Morin-Allory d'avoir accepté de faire partie de mon jury.*

*Que mes rapporteurs : le Professeur Catherine Etchebest et le Docteur Serge Couzry trouvent ici l'expression de ma profonde gratitude pour avoir accepté de juger ce travail (en plein mois d'août !!!).*

*Je tiens à exprimer ma profonde reconnaissance au Docteur Monique Genest pour m'avoir accueilli dans son équipe depuis la maîtrise, pour m'avoir accordé sa confiance. Merci pour sa rigueur scientifique et son entière disponibilité. Merci pour l'encadrement et pour tant d'autres choses tout au long de cette période.*

*Je remercie le Docteur Norbert Garnier qui passe de « CHARMMeur niveau 2 » à « Saint Norbert » pour son soutien, ses conseils et ses remarques et aussi pour avoir partagé son bureau pendant deux ans. Merci aussi pour l'encadrement du DEA.*

*Merci à tous les membres de l'équipe de simulation numérique et modélisation moléculaire avec qui j'ai passé d'agréables moments. Merci au Docteur Daniel Genest pour ses explications des concepts physiques, pour son humour et pour l'encadrement de la maîtrise ; à Alain Boyer, le monsieur informatique du groupe et à Sonia Mazier ou chocoquick pour avoir accepté de partager son bureau quand j'étais SBF, merci pour l'ambiance et pour ton aide. Bon courage pour ta dernière année.*

*Je remercie du fond du cœur les Docteurs Stéphane Bourg et Jean-Pierre Duneau pour leurs conseils et leur sympathie.*

*Merci à tous les membres du CBM (Bertrand Castaing, Françoise Vovelle, Françoise Keita-Paquet, Christian Damblon, Mélanie Spothem-Maurizot, Franck Coste,..., et j'en oublie plein d'autres) avec qui j'ai passé quatre agréables années. Merci pour les conseils, les causeries de tout genre.*

*Merci aux doctorants (ou anciens doctorants) : Samia Aci, Pierre Aller (3 ans ça passe très vite), Yann Bilbille, Nathalie Gillard, Olivier Lamiable, Yann-Vaï Le Bihan, Sonia Mazier, François Orange et Gilles Rautureau pour les moments de détente et d'intenses cogitations : pauses café, resto, discussions scientifiques, commérages, etc. Bon courage à tous.*

*A Yan-Vaï merci mille fois pour les soirées : repas copieux. On remet ça quand tu voudras.*

*Je remercie les amis : buyaux (particulièrement Pierre André Frot et sa Lili à qui je souhaite plein de bonnes choses); lexiens (ma seconde famille) particulièrement mes frangins et amis Bachir, SKH, Boubacar, Bacary, Sofia, Souley et BIZ.*

*A M. Monniaux et sa famille, à M. et Mme Rougon, à M. Jean Ramstein et à la famille de PAF, je tiens à leur exprimer ma profonde amitié et toute ma sympathie. Merci à TATA.*

*Je ne saurais terminer sans remercier ma famille (clin d'œil à Marie toujours disponible, à ma grande sœur Mme Gado Hadizatou qui ne ménage aucun effort pour que ses petits frères et sœurs réussissent, à Rabi et la petite Aïda, à Habî, Chaïbou, Daouda, ..., je ne saurais tous vous citer), mes beaux parents et ma*

*famille et amis d'Orléans, Agada, Aïcha, Farid, Leïla (Miss j'oublie tout!!!!), Sidi, etc. A tous ceux qui m'ont soutenu tout au long de mes études.*

*A mes parents qui malgré les 3500 km qui nous séparent ont suivi avec tendresse l'évolution de ce travail.  
A toi ma mère « Merci pour tout ce que tu fis pour moi, ton fils,  
Si loin, si près de toi. »*

*A toi ma Rayna, tu vois j'y suis arrivé et je le dois en grande partie à ton indéfectible soutien et ton amour.  
Je te dédie ce manuscrit.*

*A nous !  
A l'avenir !*

## ABREVIATIONS

**AR** : amphiréguline  
**ATP** : adénosine triphosphate  
**BTC** :  $\beta$ -celluline  
**CRC** : cancer colorectal  
**CRD** : domaine riche en cystéine  
**EGFD** : domaine de type facteur de croissance de l'épiderme  
**EGF** : facteur de croissance de l'épiderme  
**EGFR** : récepteur au facteur de croissance de l'épiderme  
**ErbB/HER** : récepteur de la famille de l'EGFR  
**EPO** : érythropoïétine  
**EPOR** : récepteur à l'érythropoïétine  
**EPR** : épiréguline  
**FGF** : facteur de croissance des fibroblastes  
**FGFR** : récepteur au facteur de croissance des fibroblastes  
**FNIII** : domaine type fibronectine III  
**GpA** : glycophorine A  
**GH** : hormone de croissance  
**GHR** : récepteur à l'hormone de croissance  
**HB-EGF** : facteur de croissance type EGF fixant l'héparine  
**HGF** : facteur de croissance des hépatocytes  
**HGFR** : récepteur au facteur de croissance des hépatocytes  
**HRG** : héréguline  
**IgD** : domaine type immunoglobuline  
**IGF** : facteur de croissance de l'insuline  
**IR ou InsulinR** : récepteur à l'insuline  
**Met** : récepteur au facteur de croissance des hépatocytes  
**NDF** : facteur de différenciation de la protéine Neu  
**Neu\* ou Neu<sup>E</sup>** : récepteur Neu portant la mutation Val664/Glu664 dans le segment transmembranaire  
**Neu ou Neu<sup>V</sup>** : récepteur Neu sauvage homologue de ErbB2 chez le rat  
**NGF** : facteur de croissance du tissu nerveux  
**NGFR** : récepteur au facteur de croissance tissu nerveux  
**NRG** : neuréguline  
**NSCLC** : cancers du poumon n'étant pas à petites cellules  
**PDGF** : facteur de croissance des dérivés plaquettaires  
**PDGFR** : récepteur au facteur de croissance des dérivés plaquettaires  
**PI3K** : phosphatidyl inositol 3-kinase  
**PTK** : protéine tyrosine kinase  
**RTK** : récepteur tyrosine kinase  
**TGF $\alpha$**  : facteur de croissance transformant  $\alpha$   
**TKI** : inhibiteur de tyrosine kinase  
**TM** : domaine ou segment transmembranaire  
**VEGF** : facteur de croissance endothélial vasculaire  
**VEGFR** : récepteur au facteur de croissance endothélial vasculaire  
**PTB** : domaine de fixation des phosphotyrosines  
**SH2** : domaine homologue à src type 2  
**SH3** : domaine homologue à src type 3  
**DCPC** : dicaproyl-phosphatidylcholine

**DMPC** : dimyristoyl-phosphatidylcholine  
**DMPS** : dimyristoyl-phosphatidylsérine  
**DOPC** : dioleoyl-phosphatidylcholine  
**DOPG** : dioleoyl-phosphatidylglycérol  
**DPC** : dodécyl-phosphocholine  
**DPPC** : dipalmitoyl-phosphatidylcholine  
**POPC** : palmitoyloléyl-phosphatidylcholine  
**POPE** : palmitoyloléyl-phosphatidyléthanolamine  
**SDS** : sodium dodécyl sulfate  
**TFE** : trifluoroéthanol  
**RMN** : Résonance Magnétique Nucléaire  
 $E_{elec}$  : énergie électrostatique  
 $E_{vdw}$  : énergie de van der Waals  
**LH** : liaison hydrogène  
**TM** : segment (ou domaine) transmembranaire  
**RMSD** : écart quadratique moyen

### CODE DES ACIDES AMINES

Acide aspartique	Asp	<b>D</b>
Acide glutamique	Glu	<b>E</b>
Alanine	Ala	<b>A</b>
Arginine	Arg	<b>R</b>
Asparagine	Asn	<b>N</b>
Cystéine	Cys	<b>C</b>
Glutamine	Gln	<b>Q</b>
Glycine	Gly	<b>G</b>
Histidine	His	<b>H</b>
Isoleucine	Ile	<b>I</b>
Leucine	Leu	<b>L</b>
Lysine	Lys	<b>K</b>
Méthionine	Met	<b>M</b>
Phénylalanine	Phe	<b>F</b>
Proline	Pro	<b>P</b>
Sérine	Ser	<b>S</b>
Thréonine	Thr	<b>T</b>
Tryptophane	Trp	<b>W</b>
Tyrosine	Tyr	<b>Y</b>
Valine	Val	<b>V</b>

## Sommaire

# CHAPITRE I : LES RECEPTEURS A ACTIVITE TYROSINE KINASE (RTK)..... 12

<b>I</b>	<b>GENERALITES SUR LES RTK.....</b>	<b>13</b>
A	ARCHITECTURE DES RTK.....	14
1	<i>Mutations au sein du domaine TM des RTK.....</i>	<i>16</i>
B	ACTIVATION ET SIGNALISATION DES RTK.....	18
1	<i>L'activation.....</i>	<i>18</i>
2	<i>Les voies de signalisation des RTK.....</i>	<i>19</i>
<b>II</b>	<b>LES RECEPTEURS DE LA FAMILLE ERBB/HER.....</b>	<b>22</b>
A	STRUCTURE.....	23
1	<i>Domaine extracellulaire.....</i>	<i>26</i>
2	<i>Domaine intracellulaire.....</i>	<i>29</i>
3	<i>Le segment transmembranaire.....</i>	<i>32</i>
4	<i>Le récepteur entier.....</i>	<i>35</i>
B	ACTIVATION.....	38
1	<i>Les ligands.....</i>	<i>38</i>
2	<i>L'activation des récepteurs ErbB.....</i>	<i>40</i>
a_	<i>L'homodimérisation.....</i>	<i>42</i>
b_	<i>L'hétérodimérisation.....</i>	<i>42</i>
c_	<i>Hiérarchie des hétérodimères.....</i>	<i>45</i>
d_	<i>Dégradation des dimères.....</i>	<i>47</i>

# CHAPITRE II : METHODOLOGIE..... 49

<b>I</b>	<b>INTRODUCTION.....</b>	<b>50</b>
<b>II</b>	<b>PRINCIPE DE LA DYNAMIQUE MOLECULAIRE ET MECANIQUE MOLECULAIRE.....</b>	<b>52</b>

A	LE CHAMP DE FORCES .....	52
1	<i>Interactions entre atomes liés</i> .....	53
a_	Elongation des liaisons covalentes .....	53
b_	Distorsion des angles de valence.....	54
c_	Angle de torsion .....	54
d_	Angle dièdre impropre .....	54
2	<i>Interactions entre atomes non liés</i> .....	55
a_	Interactions de van der Waals .....	55
b_	Interactions électrostatiques .....	56
c_	Liaisons hydrogène .....	57
3	<i>Termes de contraintes</i> .....	58
a_	Les contraintes harmoniques de positions atomiques .....	58
b_	Les contraintes harmoniques de distances interatomiques.....	59
c_	Les contraintes d'angle de valence et d'angles dièdres $\phi$ .....	59
d_	Les contraintes de liaison rigide (SHAKE).....	59
B	ALGORITHMES DE DYNAMIQUE MOLECULAIRE.....	60
1	<i>Algorithme de Verlet</i> .....	61
2	<i>Algorithme de Leap-Frog</i> .....	62
C	METHODES DE MINIMISATION D'ENERGIE .....	63
<b>III</b>	<b>LES LOGICIELS UTILISES .....</b>	<b>65</b>
A	LE LOGICIEL CHARMM .....	65
B	LES LOGICIELS DE VISUALISATION.....	66
1	<i>VMD</i> .....	66
2	<i>MolMol</i> .....	66
<b>IV</b>	<b>PROTOCOLE DE RECHERCHE DES MODES D'ASSOCIATION .....</b>	<b>66</b>
A	CONSTRUCTION DU SYSTEME A ETUDIER.....	67
1	<i>Chauffe</i> .....	68
2	<i>Equilibration - Production</i> .....	69
3	<i>Analyses</i> .....	69
a_	Longueur des hélices .....	70
b_	Arrangement des hélices .....	70
c_	Angles d'inclinaison et de croisement .....	71



d_	Distance inter-hélices .....	72
e_	Choix des critères géométriques .....	73
<b>BIBLIOGRAPHIE CHAPITRES I ET II.....</b>		<b>75</b>
<b>CHAPITRE III : RESULTATS .....</b>		<b>91</b>
<b>INTRODUCTION.....</b>		<b>92</b>
<b>1<sup>ERE</sup> PARTIE: METHODOLOGIE DE PREDICTION DE L'ASSOCIATION DES HELICES TRANSMEMBRANAIRES. VALIDATION DE LA RECHERCHE DE STRUCTURES.....</b>		<b>94</b>
<b>INTRODUCTION.....</b>		<b>95</b>
<b>ARTICLE EBJ .....</b>		<b>97</b>
<b>CONCLUSION.....</b>		<b>110</b>
<b>2<sup>E</sup> PARTIE : ASSOCIATION HETERODIMERIQUE DES HELICES TRANSMEMBRANAIRES DES RECEPTEURS HER : ETUDE PAR SIMULATION DE DYNAMIQUE MOLECULAIRE EN L'ABSENCE D'ENVIRONNEMENT MEMBRANAIRE. ....</b>		<b>111</b>
<b>INTRODUCTION.....</b>		<b>112</b>
<b>ARTICLE.....</b>		<b>113</b>
<b>CONCLUSION.....</b>		<b>148</b>
<b>CONCLUSION GENERALE .....</b>		<b>150</b>
<b>ANNEXES .....</b>		<b>152</b>
<b>ANNEXE 1 : ROLE DES RECEPTEURS ERBB DANS LES PROCESSUS CANCEREUX .....</b>		<b>153</b>
I	<b>INTRODUCTION .....</b>	<b>154</b>
II	<b>ALTERATION DES GENES ERBB .....</b>	<b>156</b>
III	<b>LES RECEPTEURS ERBB ET LEURS LIGANDS DANS LES CANCERS HUMAINS.....</b>	<b>158</b>
VI	<b>LES RECEPTEURS ERBB ET LA THERAPIE .....</b>	<b>164</b>
<b>BIBLIOGRAPHIE ANNEXE 1 .....</b>		<b>169</b>

<b>ANNEXE 2 : ETUDE DU DOMAINE TM DU RECEPTEUR NEU PROTO- ONCOGENIQUE ET ONCOGENIQUE (<i>ARTICLE JBSD 2005</i>).....</b>	<b>178</b>
<b>INTRODUCTION.....</b>	<b>179</b>
<b>CONCLUSION.....</b>	<b>190</b>

# INTRODUCTION

*Les récepteurs à activité tyrosine kinase ou RTK constituent une superfamille de protéines impliquées dans d'importants processus physiologiques. Ils sont également impliqués dans plusieurs pathologies telles le diabète, le cancer, etc. Dès lors, ils font l'objet de nombreuses études expérimentales et théoriques afin d'essayer de comprendre leur mode de fonctionnement dans le but de pouvoir proposer des thérapies contre les maladies dans lesquelles ces protéines sont impliquées.*

*Les RTK présentent une architecture commune : un domaine extracellulaire suivi d'une unique partie transmembranaire (TM) faisant la jonction avec le domaine intracellulaire.*

*Au cours de ces travaux de thèse, je me suis intéressé aux récepteurs de la famille de l'EGFR ou ErbB (ou HER) qui forment la classe I des RTK. Cette famille comprend quatre membres : ErbB1, ErbB2, ErbB3 et ErbB4.*

*Mon travail de thèse porte sur l'étude de l'association des domaines TM de ces récepteurs ErbB. Il s'agit plus particulièrement de comprendre le mode de formation des hétérodimères au niveau membranaire afin de caractériser les structures préférentiellement formées et de rendre compte du rôle de ce domaine TM dans l'activation de ces récepteurs.*

*Mon manuscrit se subdivise en trois chapitres : les notions générales sur les RTK, le principe de la technique de simulation de dynamique moléculaire et les résultats obtenus au cours de cette thèse.*

*Le chapitre I concerne les généralités sur les récepteurs RTK. Dans cette partie un état actuel des connaissances sur ce type de récepteurs est présenté, notamment sur leur structure, leur activation et leur mode de signalisation. J'ai aussi fait le bilan des données sur les récepteurs de la famille ErbB.*

*Le chapitre II porte sur la méthodologie de simulation de dynamique moléculaire adoptée pour mener nos études sur les domaines TM des récepteurs de la famille ErbB. Ici, j'ai d'abord développé les notions principales de cette technique et présenté quelques logiciels*

*utilisés au cours de cette thèse. J'ai aussi présenté le protocole de recherche des modes d'association des hélices TM.*

*Dans le chapitre III, je présente mes travaux sur les hétérodimères du domaine TM des récepteurs ErbB. Ce chapitre est scindé en deux parties (sous forme d'articles).*

*Le premier article (1<sup>ère</sup> PARTIE) présente la méthode que j'ai développée pour prédire des modèles structuraux de l'association des domaines TM. Ce travail a permis de mettre au point une stratégie de sélection de meilleures structures dimériques. J'ai ensuite validé la méthode sur le dimère du domaine TM de la glycophorine A (GpA) dont la structure a été résolue par RMN (MacKenzie et al. 1997). Ces résultats sont publiés dans le journal : European Biophysics Journal (Eur. Biophys. J.) : (Samna Soumana et al. 2007).*

*Le deuxième article (2<sup>e</sup> PARTIE) est consacré à l'application de cette méthode à l'ensemble des hétérodimères des domaines TM des récepteurs ErbB. Cette méthode combine la mécanique moléculaire, la dynamique moléculaire en absence d'environnement membranaire explicite et l'application de critères de sélection des structures modélisées. Cet article vient d'être soumis à publication dans Biopolymers.*

*Le manuscrit est complété par deux annexes, la première portant sur le rôle des récepteurs ErbB dans les cancers humains et la seconde est un article publié dans JBSD en 2005.*

# **CHAPITRE I : LES RECEPTEURS A ACTIVITE TYROSINE KINASE (RTK)**

## I Généralités sur les RTK

La première phosphorylation de protéine fut découverte dans les années 1950 par Krebs et Fischer (Krebs and Fischer 1956) alors que l'existence de protéines kinases (PKs) a été mise en évidence chez les bactéries par Wang et collaborateurs en 1978 (Wang and Koshland 1978). Les processus de phosphorylation/déphosphorylation mis en place par les kinases sont impliqués dans la plupart des mécanismes de régulation intracellulaires (Hunter 1995; Taylor et al. 1995). Certaines de ces kinases jouent un rôle important dans le métabolisme des glucides, des nucléotides, des vitamines, dans la régulation des gènes, dans les mécanismes de contraction musculaire, etc. Les kinases sont de nos jours les enzymes les plus étudiées.

Actuellement, grâce au séquençage du génome humain, plus de 500 protéines kinases sont connues. Les protéines kinases, impliquées dans plusieurs voies de transduction membranaires du signal, correspondent à l'une des plus importantes familles de protéines (Noble et al. 2004). Près d'une centaine de gènes codent pour ces types de protéines (Blume-Jensen and Hunter 2001).

Les protéines kinases peuvent être classées en fonction de leur structure ou de leur séquence. Elles peuvent aussi être classées selon leur localisation dans la cellule ou selon le type d'acide aminé qu'elles phosphorylent. Il existe les sérine/thréonine kinases et les tyrosine kinases. Les PKs ont deux localisations possibles. On distingue les PKs cytosoliques et les PKs membranaires. Comme PKs cytosoliques nous pouvons citer la famille des protéines :

- Src (impliquées dans la régulation de la différenciation et de la prolifération cellulaire, dans la transduction du signal *via* les récepteurs aux facteurs de croissance, les intégrines, etc.). Ce sont les premières protéines tyrosine kinases connues (Parsons and Parsons 1997).

- Abl, pour Abelson, intervient dans les processus de régulation du cycle cellulaire et dans le couplage direct à l'ADN (Smith and Mayer 2002).

- Jak ou Janus kinase (rôle dans les voies de transduction du signal *via* les cytokines et dans l'hématopoïèse).

Parmi les PKs membranaires il existe principalement deux classes à savoir les récepteurs à activité tyrosine kinase (ou RTK) et les sérine/thréonine kinases. Dans le génome humain, les

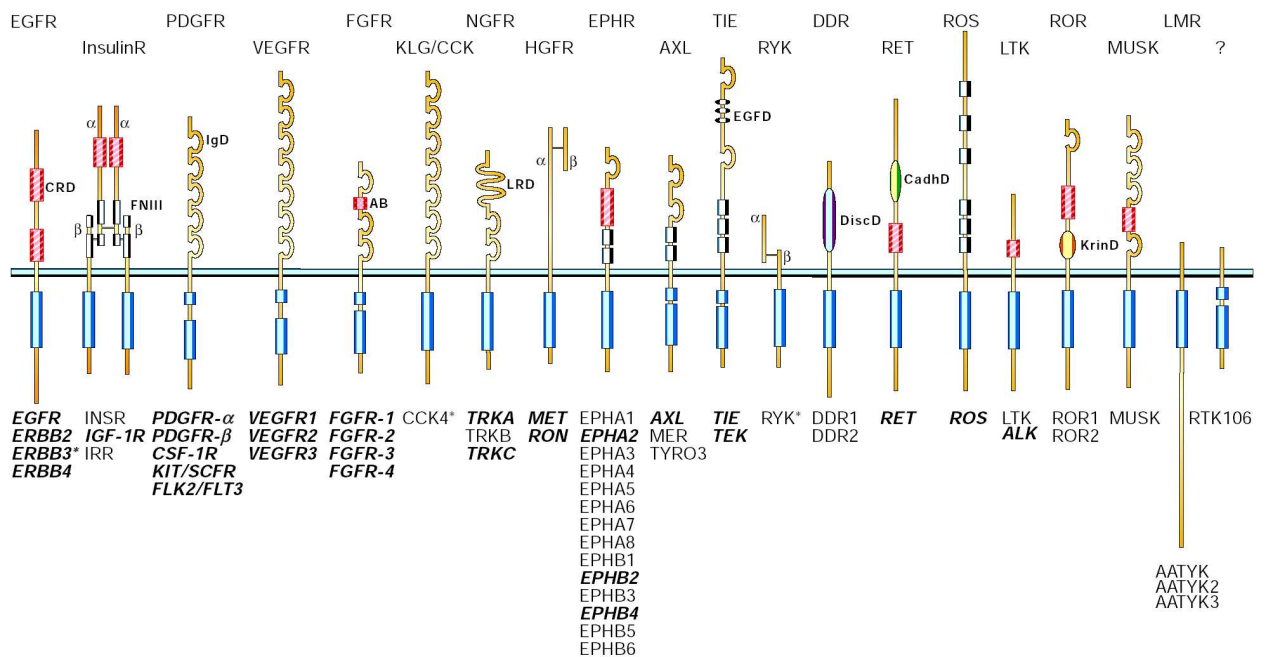
récepteurs à activité tyrosine kinase sont codés par 58 gènes (Blume-Jensen and Hunter 2001).

Au cours de ces travaux de thèse, nous nous sommes intéressés principalement aux récepteurs à activité tyrosine kinase ou RTK.

## **A Architecture des RTK**

Les récepteurs à activité tyrosine kinase ont été découverts dans les années 1980 par Hunter et Sefton (Hunter and Sefton 1980). Ces récepteurs sont des glycoprotéines présentant une organisation commune à savoir : un domaine extracellulaire portant le site de fixation d'un ligand polypeptidique, un unique segment transmembranaire et une portion cytoplasmique siège de l'activité tyrosine kinase et portant des sites d'interaction protéine-protéine. Les RTK possèdent une activité TK intrinsèque stimulée par la liaison d'un ligand. Cette activité enzymatique est localisée dans le cytoplasme et permet le transfert du phosphate  $\gamma$  de l'ATP vers l'hydroxyle des tyrosines des protéines cibles et/ou du récepteur lui-même. Ce processus est appelé autophosphorylation (Hunter 1998). Il peut aussi y avoir une trans-autophosphorylation des résidus tyrosine des récepteurs. Les RTK permettent la transmission d'un signal de l'extérieur vers l'intérieur de la cellule et jouent ainsi un rôle important dans le contrôle de nombreux processus biologiques, tels que le cycle cellulaire, la migration cellulaire, le métabolisme, la prolifération et la différenciation cellulaire (Ullrich and Schlessinger 1990).

Ils sont classés en une vingtaine de familles (ou classes), présentées sur la **figure 1.1**, (Blume-Jensen and Hunter 2001; Hubbard 2002) selon la structure de leurs domaines extracellulaire et intracellulaire. Parmi ces différentes familles, on distingue les récepteurs aux facteurs de croissance de l'épiderme (EGFR), les récepteurs à l'insuline (InsulinR), les récepteurs aux facteurs de croissance des plaquettes (PDGFR), les récepteurs aux facteurs de croissance de l'endothélium vasculaire (VEGFR) et les récepteurs aux facteurs de croissance des fibroblastes (FGFR).



### Fi

**caractéristiques structurales.** De haut en bas : le nom du représentant caractéristique de chaque famille de RTK, sa représentation schématique, et les membres de chaque famille. AB : boîte acide, CadhD : domaine de type cadhérine, CRD : domaine riche en cystéines, DiscD : domaine de type discoidine, EGFD : domaine de type EGF (facteur de croissance épidermique), FNIII : domaine de type fibronectine type III, IgD : domaine de type immunoglobuline, KrinD : domaine de type krinD, LRD : domaine riche en leucine. D'après (Blume-Jensen and Hunter 2001).

La majorité des RTK est exprimée à la surface cellulaire sous forme d'une chaîne polypeptidique unique et est monomérique en l'absence de ligand (**Figure 1.1**). Cependant certains RTK ont des structures particulières : c'est le cas des récepteurs de la sous famille HGFR et ceux de la sous famille des récepteurs à l'insuline (InsulinR). Le récepteur Met (facteur de croissance des hépatocytes) membre de la sous famille HGFR est constitué par une courte chaîne  $\alpha$  qui est liée par un pont disulfure à une chaîne  $\beta$  membranaire. Pour la sous-famille des récepteurs de l'insuline et de l'IGF1 (*Insulin Growth Factor 1*), la structure consiste en deux chaînes  $\alpha$  extracellulaires liées par des ponts disulfure à deux chaînes  $\beta$  membranaires. Les deux chaînes  $\alpha$  sont également liées par des ponts disulfure. Cela aboutit à la formation d'un hétérodimère  $\alpha_2\beta_2$  (Van Obberghen 1994).

Selon le récepteur, la partie extracellulaire est formée d'une association de plusieurs domaines de type soit immunoglobuline (IgD), soit fibronectine III (FNIII), soit riche en cystéine (CRD) et de type facteur de croissance épidermique (EGFD). La fixation du ligand implique un seul de ces domaines. Dans la plupart des cas, la fonction des autres domaines n'est pas connue,



sauf pour certains RTK, où ils pourraient moduler la liaison du ligand. A l'inverse le domaine intracellulaire présente une organisation assez simple.

En effet, la partie cytoplasmique est composée de 3 domaines à savoir, une région juxta-membranaire suivie par le domaine à activité tyrosine kinase et terminée par une queue carboxy-terminale. Le domaine à activité tyrosine kinase des récepteurs de la famille PDGF présentent une insertion d'environ 100 résidus. Selon les RTK, les domaines juxta-membranaire et carboxy-terminal présentent une variabilité dans leurs longueurs et leurs séquences. Ces domaines contiennent des sites de phosphorylation de tyrosine et de sérine/thréonine, qui peuvent être phosphorylés par le récepteur même (autophosphorylation) ou par des protéines kinases hétérologues (Hubbard 2002).

Quant au domaine transmembranaire (TM) servant de lien entre les parties extra- et intracellulaire, il est constitué de 23 à 26 résidus hydrophobes dont la séquence présente peu d'homologie d'une protéine à l'autre. Pour assurer l'arrêt de la translocation lors de la synthèse peptidique, le domaine TM a de nombreux résidus chargés dans sa partie C-terminale. Ces résidus permettent aussi l'ancrage cytoplasmique de la protéine en prévenant tout mouvement vertical de la protéine dans la membrane.

Ce domaine est une composante principale dans le fonctionnement de ces récepteurs. Son rôle a été suggéré suite à la découverte en 1986 par Bargmann et collaborateurs (Bargmann et al. 1986) d'une forme oncogénique du récepteur Neu. Cette forme induit une activation constitutive du récepteur et est caractérisée par une mutation Val664/Glu.

## 1 Mutations au sein du domaine TM des RTK

Des mutations comparables à Val664/Glu ont été réalisées dans des segments transmembranaires d'autres RTK et, comme dans Neu muté, une activation constitutive de la kinase est associée à la mutation dans la plupart des cas. Le **tableau 1.1** résume les données collectées sur ces mutations dans les segments transmembranaires des RTK.

Modifications au niveau du segment TM	Conséquences	Références
Neu : mutation Val664→Glu	Activation constitutive	(Bargmann and Weinberg 1988)
ErbB2 : mutation Val659→Glu	Dimérisation Activation constitutive	(Weiner et al. 1989) (Segatto et al. 1988)
EGFR : mutation Val627→Glu	Activation constitutive	(Miloso et al. 1995)
EGFR drosophile : mutation Val828→Glu	Activation constitutive	(Wides et al. 1990)
Récepteur insuline : mutation Val938→Asp	Activation constitutive ⤵ taux de synthèse du pro-récepteur	(Longo et al. 1992) (Yamada et al. 1992)
IR-1 : mutation Val922→Glu	Activation constitutive	(Takahashi et al. 1995)
PDGFR : chimère avec segment TM de Neu V→E	Activation constitutive	(Petti et al. 1998)
Récepteur insuline : chimère avec segment TM de Neu V→E	Activation constitutive	(Cheatham et al. 1993)
Récepteur insuline : chimère avec segment TM de Neu	Activation constitutive ⤵ internalisation	(Yamada et al. 1992) (Yamada et al. 1995)
Neu : expression de peptides correspondant au segment TM de Neu	Inhibition de la transformation par Neu V→E Inhibition de la dimérisation	(Lofts et al. 1993)
Récepteur insuline : inversion du segment TM (Lys917-Gln944)	Problème de biosynthèse Forte ⤵ activité tyrosine kinase in vitro et in vivo	(Yamada et al. 1995)
Récepteur insuline : Gly933-Pro934→Ala-Ala	Internalisation ↗ x2	(Goncalves et al. 1993)
Récepteur insuline : chimère avec les domaines extracellulaire et TM du récepteur aspartate	Activation par aspartate Spécificité du substrat ≠/récepteur insuline	(Moe et al. 1989)
v-ros : délétion de 3 acides aminés du segment TM	Perte du pouvoir transformant en modifiant la spécificité du substrat	(Zong and Wang 1994)
EGFR : chimères avec cytoplasmique de c-ros : -ER1 avec segment TM de EGFR -ER2 avec segment TM de c-ros	ER1 inhibe et ER2 stimule la croissance cellulaire, signalisation ≠	(Xiong et al. 1996)
EGFR : chimères avec domaine cytoplasmique de NGFR : -A avec segment TM de EGFR -B avec segment TM de NGFR	Différenciation cellulaire en réponse à EGF pour B uniquement	(Yan et al. 1991)
FGFR3 : mutation Gly 380→Arg	Achondroplasie (forme de nanisme) Activation constitutive	(Shiang et al. 1994) (Rousseau et al. 1994) (Webster and Donoghue 1996) (Monsonogo-Ornan et al., 2000)
FGFR3 : mutation Gly 369→Cys	Achondroplasie (forme de nanisme)	(Chen et al. 1999)
FGFR3 : mutation Ala 391→Glu	Syndrome de Crouzon avec Acanthosis nigricans	(Meyers et al. 1995)
FGFR3 : mutation Lys 650→Met	Displasie squelettique avec retard du développement et Acanthosis nigricans	(Tavormina et al. 1999)
FGFR3 murin : mutation Gly 374→Arg Et FGFR2 murin : mutation Cys 382→Arg	Activation constitutive	(Li et al. 1997)
Neu : chimère avec segment TM de FGFR3 Gly 380→Arg	Activation constitutive	(Webster and Donoghue 1996)

**Tableau 1.1 :** Tableau récapitulatif des effets d'altérations au niveau du domaine TM sur les récepteurs à activité tyrosine kinase. Extrait de la thèse de Pierre Aller, 2004.

## **B      Activation et signalisation des RTK**

Les modes d'activation et de signalisation des RTK sont représentés sur la **figure 1.2** (Holbro and Hynes 2004) correspondant au mécanisme d'activation et de signalisation des récepteurs ErbB.

### **1      L'activation**

Plusieurs étapes caractérisent l'activation des RTK. Nous avons d'abord la fixation d'un ligand qui induit ensuite une dimérisation ou une oligomérisation non covalente des récepteurs.

Cette dimérisation facilite la trans- autophosphorylation (entre les deux monomères) des résidus tyrosine situés dans la partie cytoplasmique du dimère, (Heldin 1995; Jiang and Hunter 1999; Lemmon and Schlessinger 1994; Schlessinger 1988; Ullrich and Schlessinger 1990).

L'autophosphorylation est nécessaire à deux fonctions distinctes dans le processus d'activation des RTK : la stimulation de l'activité catalytique du récepteur tyrosine kinase et ce à travers une stabilisation de la configuration active, et la création de sites de liaison pour des protéines de signalisation. La phosphorylation des résidus tyrosine lève l'auto-inhibition permettant l'accès au site actif à l'ATP et aux protéines effectrices. Elle permet également la création de sites de liaison pour des protéines de la voie de signalisation contenant des domaines de reconnaissance de phosphotyrosine (Hubbard 2002).

Parmi ces sites de liaison, on trouve des résidus phosphotyrosine à l'intérieur des séquences spécifiques qui servent de sites de recrutement pour des protéines contenant des modules de reconnaissance de phosphotyrosines comme les domaines SH2 ou PTB (Pawson 1995).

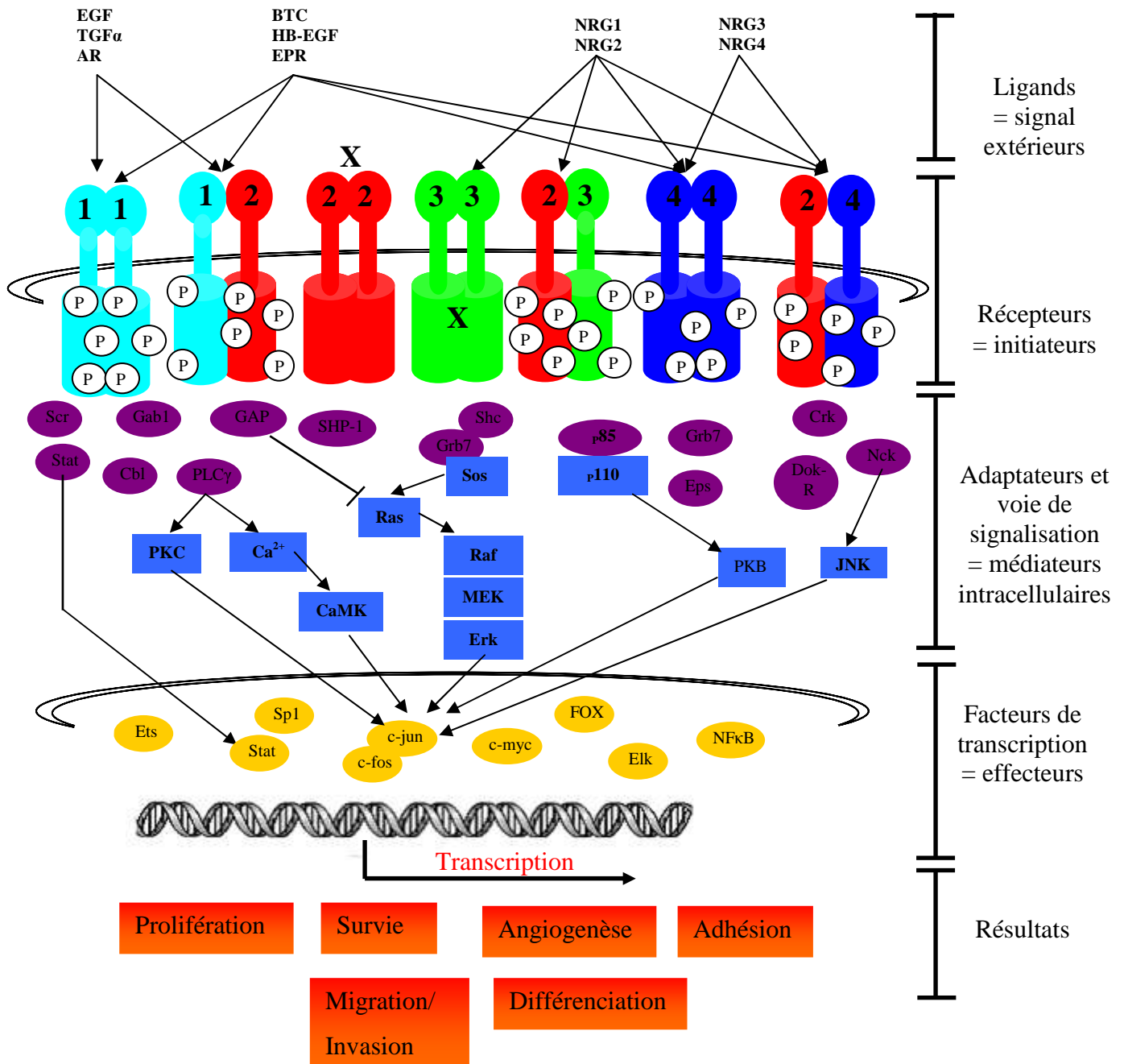
Schlessinger a récemment émis l'hypothèse de l'existence de dimères actifs même en l'absence de fixation de ligand. Cela a été proposé suite aux études menées sur l'autophosphorylation des RTK qui peut être améliorée par des inhibiteurs des protéines tyrosine phosphatase ou par la surexpression de récepteurs (Schlessinger 2000).

## 2 Les voies de signalisation des RTK

Grâce aux études biochimiques, des composants clés des voies de signalisation ont été identifiés. De plus, les mêmes protéines sont souvent identifiées comme les produits de gènes qui sont mutés et ce dans différentes maladies humaines comme le cancer, les immunodéficiences, les maladies neurologiques etc. Les RTK activent plusieurs voies de signalisation *via* des protéines possédant des modules protéiques pour transmettre le signal au noyau ou à différentes enzymes. La **figure 1.2** montre les différentes voies de transduction du signal et des réseaux de signalisation qui émergent depuis quelques années. Les principes généraux qui gouvernent le flux d'informations spatio-temporelles de la surface cellulaire au noyau et les modes de communication entre les différentes voies de signalisation sont de mieux en mieux connus.

Les principales voies de signalisation sont : Ras/MAP kinases (pour mitogen-activated protein kinases) et celle du métabolisme du phosphoinositol. Ces voies jouent un rôle important dans de nombreux processus métaboliques, le cycle cellulaire, la migration cellulaire, la survie, la prolifération, l'adhésion et la différenciation cellulaire, etc. (**voir figure 1.2**).

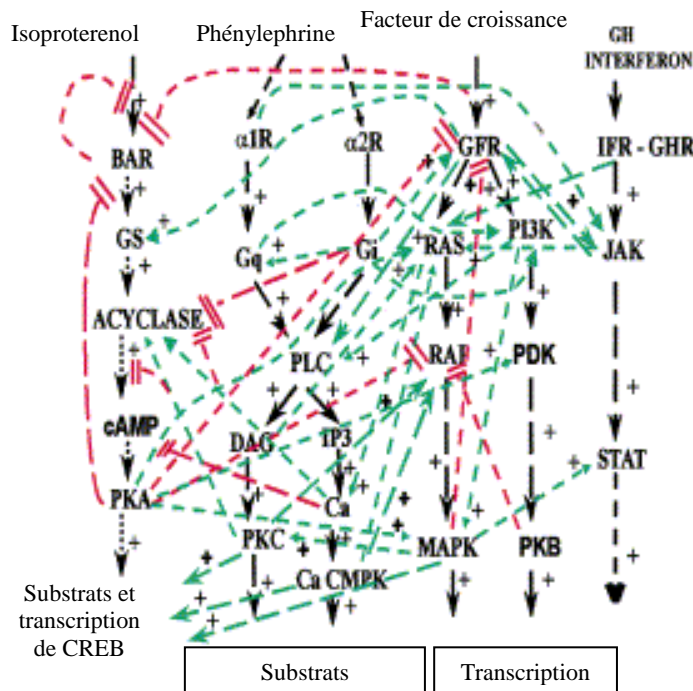
L'activation et l'autophosphorylation des RTK permettent aussi de recruter une autre classe d'enzyme, à savoir les PI3-K (pour phosphoinositide 3-kinase), tout comme la protéine kinase *via* la phospholipase C- $\gamma$  (PLC- $\gamma$ ). C'est le domaine SH2 qui se fixe aux récepteurs autophosphorylés.



**Figure 1.2 :** Représentation schématique de la redistribution au sein du réseau de signalisation des ligands de la famille EGF et de leurs récepteurs. Les ligands de la famille de l'EGF se lient avec plus ou moins de sélectivité aux récepteurs (numérotés de 1 à 4, en bleu, vert, rouge et cyan), entraînant la formation de différents dimères. ErbB2 ne possède pas de ligand connu, et ErbB3 est dépourvu d'activité kinase. Une fois les récepteurs activés et phosphorylés, diverses molécules adaptatrices (Shc, Grb2, P85,...) ou dotées d'activité enzymatique (Src, PLC $\gamma$ ) sont recrutées (en violet). Elles vont ensuite activer diverses voies de signalisation (entre autres la voie MAP kinase représentée ici par la cascade Ras-Raf-MEK-Erk, ou la voie de la PI3-kinase avec la p110 et PKB ; en bleu foncé), ce qui aboutit à l'activation de différents facteurs de transcription (en jaune). Le type et le degré d'activation des voies enclenchées, et donc le résultat final, dépendent à la fois des ligands et des dimères impliqués. D'après (Holbro and Hynes 2004).

Il est indispensable de souligner que d'autres récepteurs tels que les récepteurs couplés aux protéines G et des récepteurs à d'autres hormones peuvent activer les voies de signalisation intracellulaire activées par les RTK. En fait, une cascade de signalisation peut non seulement

être activée par différents types de récepteurs, mais elle peut également moduler d'autres cascades et cela à plusieurs niveaux. Il est donc impossible d'affirmer que l'activation d'un récepteur aboutira à une réponse biologique particulière par la seule étude des voies de signalisation. Dumont et collaborateurs ont résumé (après démonstration expérimentale) les croisements possibles entre quatre cascades de signalisation décrites dans la littérature : le résultat présenté sur la **figure 1.3** a été qualifié par ces auteurs par le terme évocateur «horror graph» (Dumont et al. 2001).



**Figure 1.3 : Réseaux établis entre 4 exemples de voies de signalisation différentes : «Everything does everything»** selon Dumont (Dumont et al. 2001).

Signalisation croisée entre 4 cascades de transduction décrites dans la littérature. La cascade linéaire est représentée en noir. Les croisements entre les différentes voies de signalisation sont représentés en couleur : les voies inhibitrices sont représentées en rouge; les voies activatrices sont représentées en vert.

Bien que presque tous les RTK soient activés par dimérisation, les mécanismes liés à la transduction du signal de l'extérieur vers le cytoplasme *via* la dimérisation des récepteurs demeurent encore mal connus.

Dans la partie qui suit, j'ai choisi de présenter les données concernant les récepteurs de la famille ErbB/HER.

## II Les récepteurs de la famille ErbB/HER

Les récepteurs de la famille ErbB/HER sont des protéines transmembranaires appartenant à la classe I des RTK (**figure 1.1**). Cette classe est composée de 4 membres à savoir : ErbB1/HER1 ou récepteur au facteur de croissance de l'épiderme (EGFR), premier RTK purifié et cloné (Ullrich et al. 1984) ; ErbB2/HER2 (ou p185<sup>c-ErbB2</sup>) également appelé Neu chez le rat (ces 2 récepteurs de 1255 et 1257 acides aminés respectivement partagent 88% d'homologie sur l'ensemble de leur séquence : alignement réalisé avec ClustalW), ErbB3/HER3 (Kraus et al. 1989; Plowman et al. 1990) et ErbB4/HER4, le plus récent (Plowman et al. 1993).

Comme les autres récepteurs RTK, les récepteurs de la famille ErbB sont exprimés dans une grande variété de tissus (Srinivasan et al. 1998). Par exemple, ces récepteurs sont fortement exprimés dans le mésenchyme épithélial et le tissu neuronal. Ils sont impliqués dans plusieurs processus biologiques tels que la prolifération, la différenciation, la migration et la survie cellulaire (**voir figure 1.2**). Ils sont aussi impliqués dans le développement des glandes mammaires pendant la puberté et la grossesse (Marmor et al. 2004). L'absence de ErbB1 est souvent liée à des anomalies au niveau cérébral, placentaire et au niveau des structures épithéliales (Miettinen et al. 1995; Sibilina and Wagner 1995) tandis que celle de ErbB3, conduit à une absence de cellules de Schwann et à de graves neuropathies (Riethmacher et al. 1997). Les récepteurs ErbB2 et ErbB4 sont impliqués aussi dans le développement cardiaque et cérébral (Gassmann et al. 1995; Lee et al. 1995).

La dérégulation de la signalisation des récepteurs ErbB entraîne de nombreuses pathologies humaines (cancer du sein, poumon,...) (Alroy and Yarden 1997) en particulier dues à l'implication de ErbB1 et ErbB2.

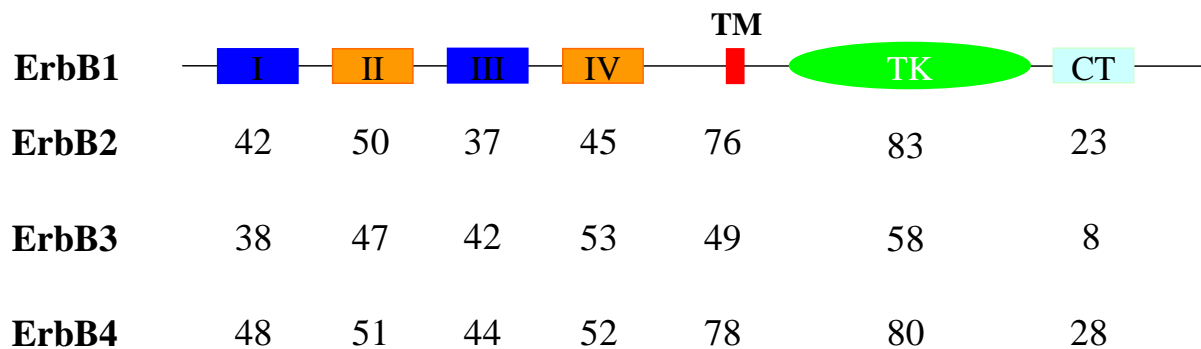
Ces récepteurs sont composés de trois parties. Un domaine extracellulaire (portant le site de fixation du ligand) relié au domaine intracellulaire (présentant une partie tyrosine kinase) par un simple segment transmembranaire (Olayioye et al. 2000). Je présente ici un état des lieux des connaissances sur la structure de ces récepteurs et leurs domaines.

Comme la plupart des RTK, c'est la fixation du ligand qui activerait ces récepteurs entraînant l'homo ou l'hétéro dimérisation. On connaît actuellement une douzaine de ligands se fixant sur les 4 récepteurs de la famille ErbB, offrant un grand nombre de combinaisons possibles, procurant ainsi une grande diversité des voies de signalisation (Gullick 2001).

La cible de mes travaux est l'association hétérodimérique des domaines TM des récepteurs ErbB.

## A Structure

Les récepteurs ErbB2, ErbB3 et ErbB4 présentent une assez forte homologie de séquence sur l'ensemble du récepteur ErbB1 de 40 à 50% selon l'alignement réalisé avec ClustalW : (<http://www.ebi.ac.uk/clustalw/index.html>). Les domaines tyrosine kinase possèdent un haut degré d'homologie (> 58%) (Peles and Yarden 1993) alors que les domaines extracellulaires sont moins conservés (**figure 1.4**). Les différences de séquences des domaines extracellulaires expliquent leur différence dans la spécificité pour la fixation du ligand (Normanno et al. 2003).

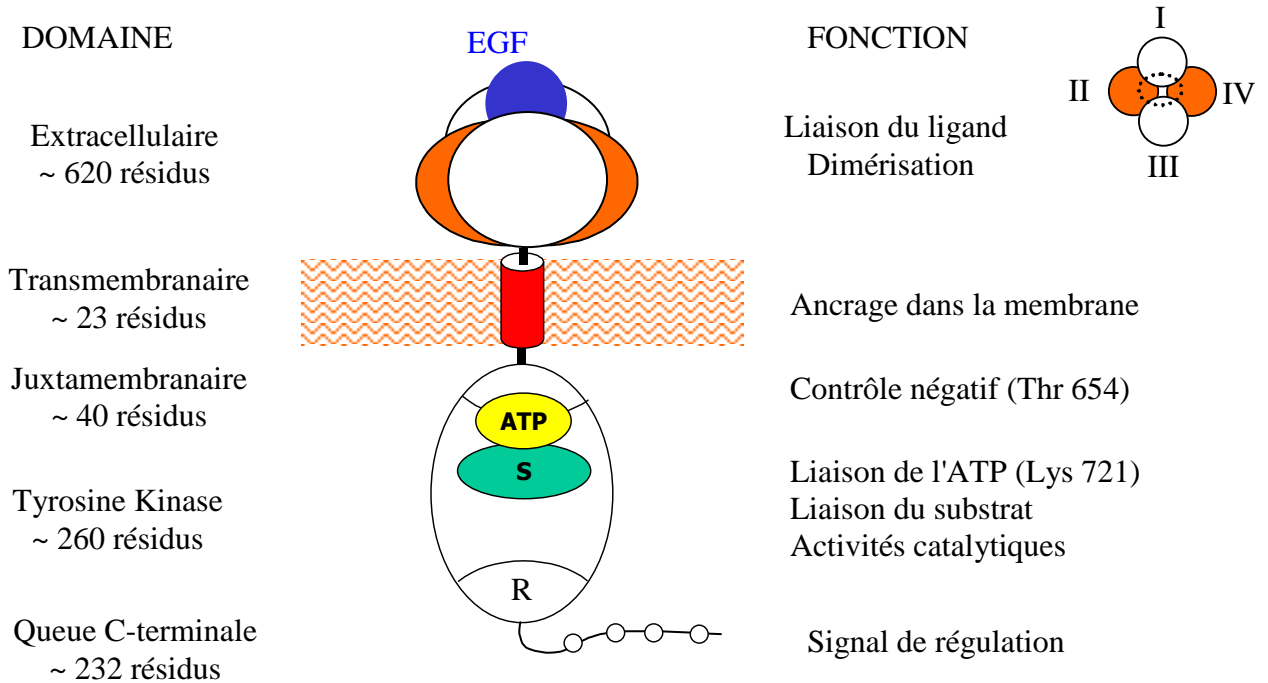


**Figure 1.4 :** Illustration schématique et homologie des récepteurs de la famille de ErbB avec le récepteur ErbB1/EGFR ou HER1. D'après (Peles and Yarden 1993).

La **figure 1.5** représente une illustration du récepteur ErbB1 ou EGFR montrant les trois domaines caractéristiques des RTK. Le domaine extracellulaire fait environ 620 résidus et est relié au domaine cytoplasmique (~ 532 résidus) par une courte hélice transmembranaire (~ 23 résidus).

Actuellement, il existe un certain nombre d'informations structurales sur les différents domaines de ces récepteurs grâce notamment à la cristallographie.





**Figure 1.5 : Topologie de EGFR.** Les 4 récepteurs ErbB sont des glycoprotéines avec une seule hélice transmembranaire (~ 23 résidus environ), reliant le domaine extracellulaire (~ 620 résidus environ) au domaine intracellulaire (~ 532 résidus environ). Les sous domaines II et IV (en orange) représentent les régions riches en cystéine. L'EGF (en bleu) se fixe sur la surface formée par les domaines I et III. Les symboles S et R représentent respectivement le site d'interaction du domaine tyrosine kinase avec le substrat et les facteurs de régulation. D'après (Ullrich and Schlessinger 1990).

Les données structurales des différents domaines de ces récepteurs sont de plus en plus nombreuses. Le **tableau 1.2** est un résumé de quelques structures obtenues par RX et répertoriées dans la PDB. Le récepteur ErbB1, seul ou en complexe est le plus étudié.

Récepteur	Code PDB	Titre	Résolution (RX) en Å	Date	Auteurs
ErbB1	2J5F	Crystal structure of EGFR kinase domain in complex with an irreversible inhibitor 13-JAB*	3.00	27-02-07	(Blair et al. 2007)
	1XKK	EGFR kinase domain complexed with a quinazoline inhibitor- GW572016*	2.40	07-12-04	(Wood et al. 2004)
	2GS6	Crystal Structure of the active EGFR kinase domain in complex with an ATP analog-peptide conjugate	2.60	20-06-06	(Zhang et al. 2006)
	2GS7	Crystal Structure of the inactive EGFR kinase domain in complex with AMP-PNP*			
	2GS2	Crystal Structure of the active EGFR kinase domain	2.80	03-04-07	(Yun et al. 2007)
	2J6M	Crystal structure of EGFR kinase domain in complex with AEE788*	3.10		
	2ITP	Crystal Structure of the active EGFR kinase domain G719S mutation in complex with AEE788*	2.74		
	2ITU	Crystal Structure of the active EGFR kinase domain L858R mutation in complex with AFN941*	2.80		
	2ITW	Crystal Structure of the active EGFR kinase domain in complex with AFN941*	2.88		
	2ITX	Crystal Structure of the active EGFR kinase domain in complex with AMP-PNP*	2.98		
	2ITY	Crystal Structure of the active EGFR kinase domain in complex with IRESSA*	3.42		
	1M14	Tyrosine Kinase Domain from Epidermal Growth Factor Receptor	2.60	04-09-02	(Stamos et al. 2002)
	1M17	Epidermal Growth Factor Receptor tyrosine kinase domain with 4-anilinoquinazoline inhibitor erlotinib			
	1MOX	Crystal Structure of Human Epidermal Growth Factor Receptor (residues 1-501) in complex with TGF-alpha	2.50	10-09-03	(Garrett et al. 2002)
	1IVO	Crystal Structure of the Complex of Human Epidermal Growth Factor and Receptor Extracellular Domains.	3.30	16-10-02	(Ogiso et al. 2002)
1YY9	Structure of the extracellular domain of the epidermal growth factor receptor in complex with the Fab fragment of cetuximab/Erbitux/IMC-C225*	2.61	26-04-05	(Li et al. 2005)	
ErbB2	1MFG	The Structure of ERBIN PDZ* domain bound to the Carboxy-terminal tail of the ErbB2 Receptor	1.25	21-01-03	(Birrane et al. 2003)
	1N8Z	Crystal structure of extracellular domain of human HER2 complexed with Herceptin Fab	2.52	18-02-03	(Cho et al. 2003)
	1S78	Insights into ErbB signaling from the structure of the ErbB2-pertuzumab complex	3.25	27-04-04	(Franklin et al. 2004)
	2A91	Crystal structure of ErbB2 domains 1-3	2.50	26-07-05	(Garrett et al. 2002)
ErbB3	1M6B	Structure of the HER3 (ERBB3) Extracellular Domain	2.60	02-08-02	(Cho and Leahy 2002)
ErbB4	2AHX	Crystal structure of ErbB4/HER4 extracellular domain	2.40	27-09-05	(Bouyain et al. 2005)

**Tableau 1.2** : Exemple de quelques structures obtenues par cristallographie des rayons X des différents domaines des récepteurs ErbB issus de la Protein Data Bank (PDB), \* voir PDB (<http://www.rcsb.org/pdb/home/home.do>).

## 1 Domaine extracellulaire

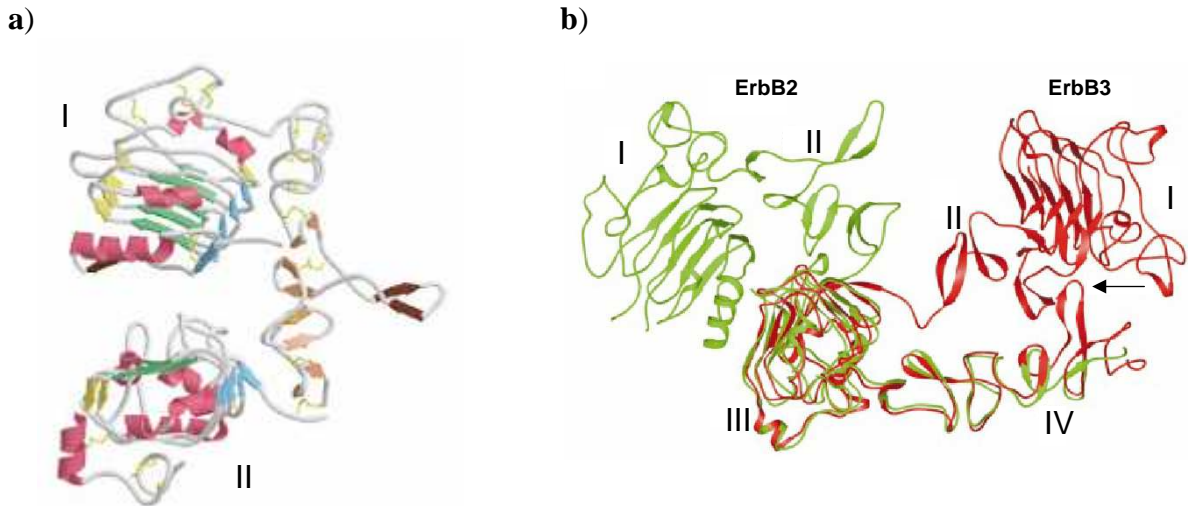
Le domaine extracellulaire est formé de 4 régions : I, II, III et IV (**figure 1.5**) (Jorissen et al. 2003; Ullrich and Schlessinger 1990). L'extrémité N-terminale de ce domaine possède un peptide signal de 20 résidus. La région I est constituée de 140 résidus. Les régions, II et IV, sont riches en cystéines et séparées par la partie III. Ces domaines II et IV ne forment pas de pont disulfure (Bazley and Gullick 2005).

Les domaines I et III présentent une homologie d'environ 40% entre les récepteurs tandis que les domaines II et IV présentent une homologie voisine de 50% (Peles and Yarden 1993).

Ces dernières années une percée importante a été effectuée dans l'étude structurale du domaine extracellulaire de ces récepteurs. Ogiso et collaborateurs ont obtenu la structure cristallographique du récepteur EGFR lié à l'EGF (code PDB : 1IVO) (Ogiso et al. 2002) tandis que d'autres auteurs l'ont cristallisé avec le TGF $\alpha$  (Ferguson et al. 2003; Garrett et al. 2002). Ils montrent l'implication des domaines I et III dans la fixation de ligand.

Les structures du domaine extracellulaire de Neu seul (code PDB : 1N8Y) et de ErbB2 complexé avec l'herceptine (code PDB : 1N8Z) sont aussi obtenues avec une résolution de 2.4 Å et 2.5 Å respectivement par Cho et collaborateurs grâce à des études cristallographiques (Cho et al. 2003). L'herceptine, molécule anti-cancéreuse dirigée contre ErbB2, se lie sur la portion C-terminale du domaine IV du récepteur ErbB2, en un site incluant la boucle étendue du domaine II dans l'état inactivé de ErbB1 et ErbB3.

La comparaison des structures des domaines extracellulaires de ErbB2, ErbB3 (Cho and Leahy 2002) et ErbB1 (**figure 1.6 a**) (Ferguson et al. 2003; Garrett et al. 2002) indique que les domaines I et II et les domaines III et IV conservent la même orientation relative. Mais l'orientation de la paire I-II est très différente de celle de la paire III-IV. Le changement conformationnel implique des mouvements de corps rigides de la paire I-II relativement à la paire III-IV autour d'un pivot situé entre les domaines II et III (Cho et al. 2003). La superposition des structures des domaines extracellulaires de ErbB2 et ErbB3 est représentée dans la **figure 1.6 b**.



**Figure 1.6 : Représentation de la structure du domaine extracellulaire a)** de ErbB1/EGFR (code PDB : 1MOX). D'après (Garrett et al. 2002) ; **b)** de ErbB2 (code PDB : 1N8Z) en vert et de ErbB3 (1M6B) en rouge après superposition des domaines III et IV. La flèche indique la relation étroite en les domaines II et IV du récepteur ErbB3. D'après (Cho et al. 2003).

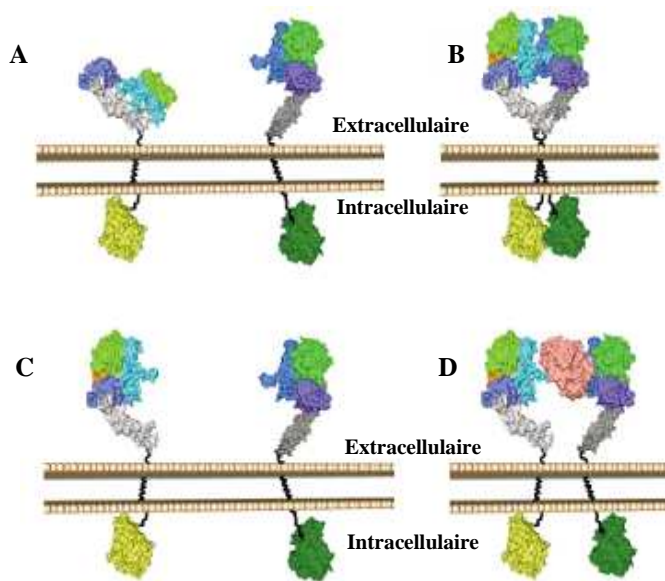
La proximité et l'orientation relative entre les domaines I et III de ErbB2 sont très similaires au réarrangement des domaines observés dans la structure de l'EGFR complexé avec le TGF- $\alpha$  ou l'EGF. Le récepteur ErbB2 est en configuration ouverte alors que les autres récepteurs en absence de ligand sont en configuration fermée. ErbB1 adopte une configuration ouverte similaire à celle adoptée par ErbB2 après fixation du ligand. La configuration fermée est caractérisée par des contacts étroits entre les domaines II et IV et la configuration ouverte par des contacts étroits entre les domaines I et III. La configuration ouverte est plus favorable à la dimérisation *via* les boucles des domaines II et IV. L'absence d'affinité de ErbB2 pour un ligand peut-être expliquée par sa structure ouverte qui ne requiert pas la fixation de ligand pour relâcher la conformation auto-inhibée (ou fermée). Dans cette configuration, ErbB2 est capable d'interagir avec les autres partenaires d'où son rôle de co-récepteur avec chaque autre récepteur en absence de liaison directe avec le ligand. L'absence de conformation auto-inhibée de ErbB2 explique son potentiel transformant quand il est surexprimé.

Franklin et collaborateurs ont résolu en 2004, la structure cristallographique du domaine extracellulaire de ErbB2 complexé avec un fragment de pertuzumab (Franklin et al. 2004). Comme le trastuzumab, le pertuzumab est un anticorps monoclonal dirigé contre le récepteur ErbB2 (Vajdos et al. 2002).

La structure de ce complexe ErbB2/pertuzumab a été obtenue avec une résolution de 3.2 Å. Le pertuzumab se fixe spécifiquement sur le récepteur ErbB2 (pas de fixation spécifique sur

Neu). Sa fixation bloque la boucle de dimérisation de ErbB2 et empêche ainsi l'hétérodimérisation avec l'EGFR ou ErbB3 comme le montre la **figure 1.7**.

La structure du domaine extracellulaire de ErbB2 couplé au pertuzumab (Franklin et al. 2004) est très similaire à la structure seule et dans le complexe avec le trastuzumab (Cho et al. 2003). Seul le domaine IV du complexe ErbB2/pertuzumab subit une rotation de 20° par rapport au reste de la protéine.



**Figure 1.7 : Hétérodimérisation de l'EGFR avec ErbB2 en présence ou non du pertuzumab.**

Les surfaces des régions extracellulaires de ErbB2 et EGFR sont représentées avec les domaines I, II, III et IV respectivement en vert, bleu, violet et gris. L'EGFR est toujours placé du côté gauche et ErbB2 à droite. L'EGF est en orange et le pertuzumab en rose (D). La membrane lipidique est sous forme de sphères et le domaine transmembranaire est représenté par une hélice noire. Les régions intracellulaires sont en jaune et vert pour l'EGFR et ErbB2 respectivement.

**A :** monomère EGFR sans molécule de ligand fixé en proximité de ErbB2. Le domaine extracellulaire de EGFR est dans une conformation fermée qui ne permet pas la dimérisation.

**B :** Dimérisation de EGFR - ErbB2 *via* leur boucle de dimérisation.

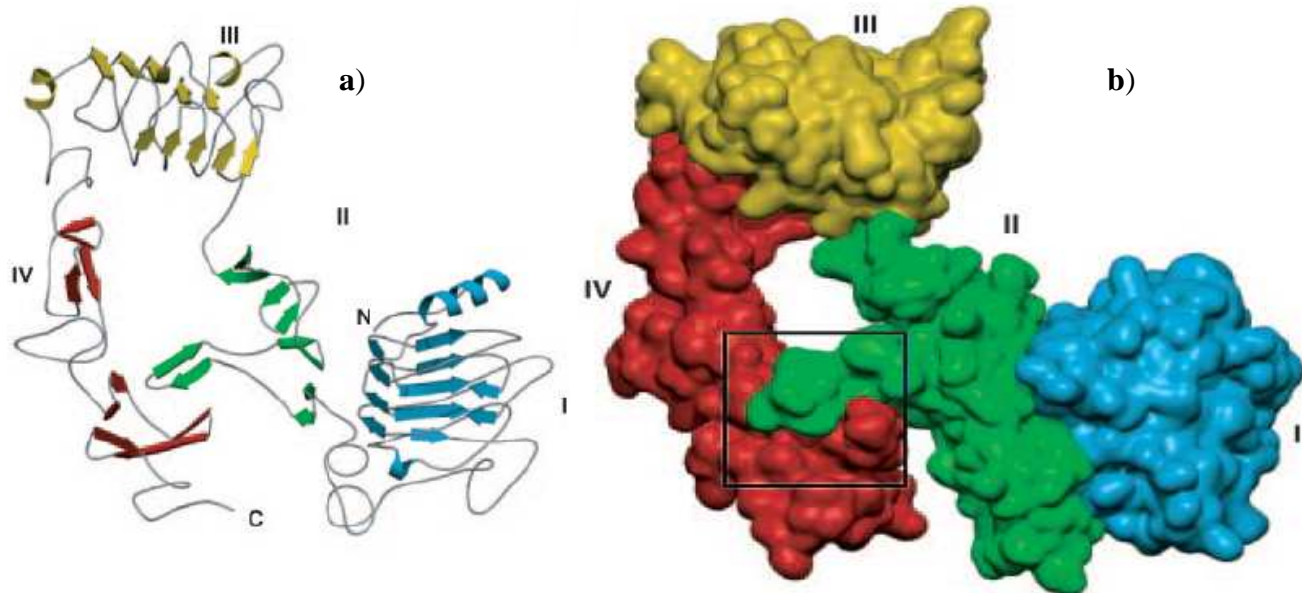
**C :** EGFR après fixation du ligand EGF (en orange) est en configuration ouverte, avec exposition des boucles de dimérisation.

**D :** la fixation du pertuzumab sur ErbB2 bloque l'hétérodimérisation en empêchant la boucle de dimérisation de l'EGFR d'interagir avec ErbB2.

D'après (Franklin et al. 2004).

La récente structure cristallographique du récepteur ErbB4 (**figure 1.8 a**) obtenue en l'absence de ligand complète les données sur ce domaine (Bouyain et al. 2005). Les auteurs ont montré que ce récepteur adopte une conformation similaire à la forme inactive de ErbB1 et ErbB3. Ils démontrent que tous les récepteurs de cette famille adoptent une conformation auto inhibée en l'absence de ligands. Contrairement à la fixation de l'EGF et du TGF sur le récepteur ErbB1, une diminution du pH n'affecte pas la fixation du ligand NRG1 sur les récepteurs ErbB3 et ErbB4 et celle du ligand BTC sur les deux récepteurs ErbB4 et ErbB1. Ces résultats montrent un important rôle du ligand dans la détermination de la fixation pH dépendante et expliqueraient les différentes réponses obtenues quand le même récepteur ErbB est stimulé par différents ligands (Bouyain et al. 2005). Dans la forme inactive, le réarrangement des domaines du récepteur ErbB4 est guidé par une boucle en épingle à cheveux de 20 Å (**figure 1.8 b**). Ce contact entraîne l'inaccessibilité d'une surface équivalente

à 674 Å<sup>2</sup> avec des paramètres comparables à ceux observés pour les récepteurs ErbB1 (Cho and Leahy 2002) et ErbB3 inactifs (Ferguson et al. 2003).



**Figure 1.8 : Structure de ErbB4.** a) Représentation en ruban de ErbB4. Les domaines I, II, III, et IV sont respectivement en bleu, vert, jaune, et rouge. N et C correspondent aux extrémités N et C. b) Représentation en surface de ErbB4. Les contacts entre les domaines II et IV sont encadrés. Code PDB : 2AHX. D'après (Bouyain et al. 2005).

Les structures cristallographiques obtenues permettent une meilleure compréhension de la fixation du ligand et de la dimérisation de ces récepteurs. Toutefois ces données sont à prendre avec précaution dans la mesure où :

- Les structures sont souvent obtenues à des températures différentes de celle du milieu biologique.
- Ce sont uniquement des structures du domaine extracellulaire. La dimérisation de ces récepteurs implique aussi les parties transmembranaire et cytoplasmique.

## 2 Domaine intracellulaire

La partie intracellulaire est composée de trois domaines : le domaine juxtamembranaire suivi du domaine tyrosine kinase (TK) et la queue autophosphorylante.

La partie juxtamembranaire, de structure encore inconnue, est composée d'environ 40 acides aminés. Ce domaine contient des sites phosphorylés par des protéines à activité tyrosine

kinase et des résidus basiques qui peuvent influencer la dimérisation des récepteurs ErbB (Aifa et al. 2005; Aifa et al. 2006).

Le domaine TK est très bien conservé entre les récepteurs ErbB1, ErbB2 et ErbB4. En effet l'alignement des séquences révèle une homologie d'environ 80% entre ces récepteurs alors qu'elle est de 58% entre ErbB1 et ErbB3 (alignement réalisé avec ClustalW : (<http://www.ebi.ac.uk/clustalw>), (**figure 1.4.**), (Peles et al. 1993). ErbB3 présente des substitutions au niveau de quatre résidus importants du domaine TK (résidus conservés pour l'ensemble des RTK) qui pourraient être liées à une activité enzymatique pour ErbB3 (Carraway and Cantley 1994). Cette homologie de séquences suggère que les domaines TK des récepteurs de la famille EGFR adoptent une structure très similaire.

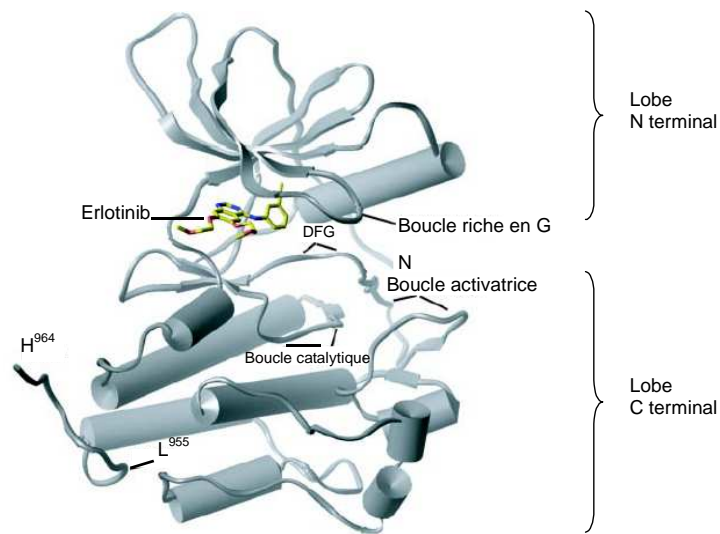
Le domaine cytoplasmique est une région importante pour la transmission du signal des récepteurs *via* les réseaux de signalisation intracellulaire, et aurait au moins un rôle stabilisateur dans la formation des dimères. Il a été montré que des mutants EGFR dépourvus de la partie extracellulaire forment des dimères avec une activité kinase, capables de s'autophosphoryler (Chantry 1995). Des études montrent que les domaines intracellulaires des récepteurs EGFR peuvent s'associer pour former des structures énergétiquement favorables en présence d'une membrane qui stabiliserait le dimère du fait de la distribution des charges électriques (Chantry 1995; Kokai et al. 1989; Murali et al. 1996).

Des études structurales ont été réalisées sur ce domaine intracellulaire (Stamos et al. 2002 ; Zhang et al. 2006).

Ainsi, en 2002, Stamos et collaborateurs ont résolu la structure du domaine TK du récepteur EGFR (code PDB : 1M14) par cristallographie aux rayons X (Stamos et al. 2002). Cette structure, incluant 40 acides aminés de la queue C-terminale, complexée ou non avec l'erlotinib a été déterminée à 2.6 Å de résolution. L'erlotinib (ou Tarceva) est un agent anti-cancéreux actuellement utilisé dans la lutte contre certains cancers mais reste encore en essai pour d'autres types de cancers (**voir annexe 1**).

Le domaine kinase est constitué de deux lobes. Le lobe N-terminal est composé de feuillets  $\beta$  et d'une hélice  $\alpha$ . Il inclut une boucle riche en glycine qui permet la fixation des nucléotides phosphates. Le lobe C-terminal adopte une organisation en hélice  $\alpha$  et comprend la boucle activatrice, la boucle catalytique et le motif DFG. Ils sont séparés par un sillon, site de

fixation de l'ATP ou de ses analogues. La structure générale du domaine kinase de l'EGFR complexé avec l'erlotinib est représentée dans la **figure 1.9**.



**Figure 1.9 : Structure du domaine tyrosine kinase de l'EGFR avec ses principales caractéristiques (code PDB : 1M14).**

L'erlotinib est situé dans le sillon séparant le lobe N-terminal du lobe C-terminal. D'après (Stamos et al. 2002).

Le lobe N-terminal a une structure tertiaire très similaire à celle observée dans les autres RTK (le RMSD calculé sur les C $\alpha$  entre la structure de l'EGFR et du FGFR est de 1.2 Å). L'absence du motif de deux résidus WE, présent dans d'autres kinases telles que le récepteur à l'insuline (InsulinR) (Hubbard 1997), le FGFR (Mohammadi et al. 1996) et la lymphocyte tyrosine kinase (Lck) (Yamaguchi and Hendrickson 1996) ne modifie pas l'activité de ce lobe. Ce motif WE permet d'importantes stabilisations structurales et apparaît comme une frontière entre le domaine kinase et le domaine juxtamembranaire. Dans l'EGFR ce motif est remplacé par les résidus Arg et Ile qui conduisent à un effet identique.

Le lobe C-terminal a une structure également similaire à celle décrite dans d'autres structures (RMSD entre le lobe C-terminal de EGFR et celui de IR et Lck est de 1.1 Å).

La boucle activatrice dans les RTK change de conformation après phosphorylation d'une tyrosine ou d'une thréonine. La structure de la boucle activatrice de EGFR avec ou sans inhibiteur diffère significativement des boucles activatrices non phosphorylées des autres RTK. Vers la fin du lobe C, le motif LVI régule la transphosphorylation des substrats tyrosines. Le site catalytique tyrosine kinase est également très conservé (RMSD entre l'EGFR, l'IR et le FGFR est de 0.1 Å). Le motif DFG, important et très conservé, occupe le site de fixation de l'ATP et adopte le même arrangement que dans le récepteur à l'insuline (Stamos et al. 2002).



Récemment, Zhang et collaborateurs (Zhang et al. 2006) ont montré, suite à des études cristallographiques et biochimiques, que la kinase de ErbB1 est activée par un mécanisme allostérique. Dans ce mécanisme, le lobe C du domaine kinase d'un des récepteurs dans le dimère forme un contact spécifique avec le lobe N du domaine kinase de l'autre récepteur pour l'activer formant ainsi un dimère asymétrique. Ce mécanisme explique pourquoi le récepteur ErbB3, dépourvu d'activité kinase intrinsèque, est capable d'activer les autres membres de la famille ErbB (Hubbard and Miller 2007).

La queue C-terminale du domaine cytoplasmique est une partie non conservée pour l'ensemble des RTK (Dougall et al. 1994). Du fait de leur faible homologie, la modélisation de la structure de cette région est limitée. Néanmoins, les modèles du domaine TK de l'EGFR obtenus par homologie et simulations de dynamique moléculaire à partir des structures cristallographiques du domaine TK du récepteur à l'insuline suggèrent un repliement de cette queue pour interagir avec le site actif du domaine kinase opposé (Groenen et al. 1997).

### 3 Le segment transmembranaire

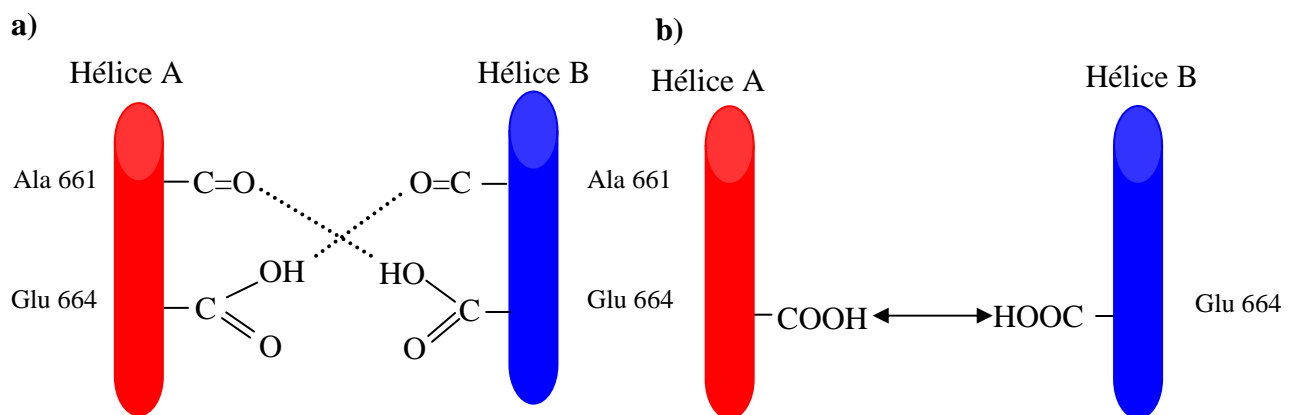
Les premiers travaux qui ont montré l'implication du domaine TM dans la dimérisation et l'activation de ces récepteurs datent d'une vingtaine d'années (Bargmann et al. 1986; Weiner et al. 1989). En plus de son rôle de lien entre les deux domaines extra- et intracellulaire, le domaine transmembranaire contribue à la dimérisation des récepteurs ErbB.

Le domaine TM des récepteurs ErbB est constitué d'environ 25 résidus hydrophobes. Ce domaine structuré en hélice  $\alpha$  a fait l'objet de plusieurs études expérimentales comme par exemple celles de (Deber and Li 1995; Miloso et al. 1995; Smith et al. 1996).

L'importance du domaine TM a été mise en évidence suite à la découverte de la forme oncogénique de Neu, résultant de la mutation ponctuelle Val664 en Glu664 dans ce domaine (Bargmann et al. 1986; Bargmann and Weinberg 1988).

Le mécanisme par lequel cette mutation active le récepteur n'est pas connu. Sternberg et Gullick ont proposé un modèle basé sur des considérations stéréochimiques supposant que cette mutation stabilise les interactions entre les hélices  $\alpha$  correspondant aux domaines TM (Sternberg and Gullick 1989).

En effet, ils ont proposé un modèle dans lequel, en considérant la conformation du domaine TM structuré en hélice  $\alpha$ , le groupe carboxyle de Glu 664 de l'hélice d'un récepteur forme une liaison hydrogène (LH) avec l'oxygène du groupe carbonyle de l'Ala 661 de l'hélice  $\alpha$  du second récepteur. Une autre liaison hydrogène symétrique est formée entre l'oxygène du groupe carbonyle de Ala 661 du domaine TM du premier récepteur avec le groupe carboxyle de Glu 664 du domaine TM du second récepteur (**figure 1.10 a**). Ainsi, cette dimérisation accrue de l'oncogène Neu mimerait une association des récepteurs induite par leur ligand et entraînerait une activation constitutive de l'activité TK (Sternberg and Gullick 1989). Un second modèle complémentaire à celui de Sternberg et Gullick a été proposé par Smith et collaborateurs qui ont étudié les séquences TM des récepteurs Neu sauvage et oncogénique par des techniques de spectroscopie infrarouge polarisée et de RMN (Smith et al. 1996). Les résultats de cette étude suggèrent l'existence de liaisons hydrogène fortes entre les chaînes latérales des résidus Glu 664 (**figure 1.10 b**).



**Figure 1.10 : Représentation schématique des interactions de liaisons hydrogène (LH) entre les hélices transmembranaires.**

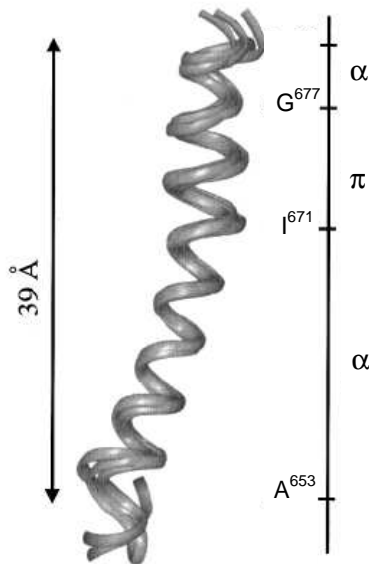
**a)** Les chaînes latérales protonnées de Glu 664 d'une hélice forment des LH avec le groupement carbonyle de Ala 661 de l'hélice opposée. D'après (Sternberg and Gullick 1989).

**b)** les chaînes latérales protonnées des résidus Glu 664 forment des LH entre elles (Smith et al. 1996).

C'est dans la partie N-terminale du domaine TM de Neu, reconnue pour induire une dimérisation des hélices TM qu'est localisé le résidu Valine 664 à l'intérieur d'une séquence consensus (Sternberg and Gullick 1990).

Goetz et collaborateurs ont étudié, la structure du domaine TM de Neu, homologue de ErbB2 (Goetz et al. 2001). Par dichroïsme circulaire il est montré que ce domaine solubilisé dans le TFE (trifluoroéthanol) ou incorporé dans des solutions micellaires de SDS (sodium dodécyl sulfate) et d'eau ou dans des bicelles composées de mélange de DMPC (dimiristoyl-

phosphatidylcholine) / DCPC (dicaproyl-phosphatidylcholine) bien hydratées, adopte une structure hélicoïdale pour environ 80% de ses acides aminés. L'étude par RMN du proton en solution dans du TFE, montre une hélice  $\alpha$  régulière présentant une déformation  $\pi$  sur deux tours (Ile 671-Gly 677) couvrant la région riche en valines (Val 673-Val 676). La **figure 1.11** représente la structure du domaine transmembranaire de Neu.



**Figure 1.11 :** Superposition des 5 structures minimisées (séquence de 35 résidus : E<sup>650</sup>- R<sup>684</sup>) du segment TM de Neu, après modélisation (code PDB : 1IIJ). Meilleure superposition des structures entre les résidus V<sup>656</sup> et K<sup>681</sup> avec un RMSD de 1 Å calculé sur les atomes du squelette peptidique. Le motif  $\alpha \pi \alpha$  est indiqué. Hélice  $\alpha$  LH : CO<sub>i</sub> et NH<sub>i+4</sub> Hélice  $\pi$  LH : CO<sub>i</sub> – NH<sub>i+5</sub>. D'après (Goetz et al. 2001).

Récemment, Houliston et collaborateurs ont mis en évidence de façon moins marquée la présence de la hernie  $\pi$  dans le segment TM de Neu. Ainsi, ils ont observé par RMN en solution dans le TFE et dans des micelles de DPC (dodécylphosphocholine) une différence structurale entre le segment TM sauvage et muté Val664/Glu (Houliston et al. 2003, 2004). Ces résultats ont été confirmés dans nos études antérieures (**voir annexe 2**) (Samna Soumana et al. 2005).

Le domaine TM des récepteurs ErbB2 ou Neu a été largement étudié au sein de l'équipe (Duneau et al. 1999; Duneau et al. 1999; Duneau et al. 1997; Duneau et al. 1996; Garnier et al. 1997; Garnier et al. 1994; Sajot et al. 1999; Sajot and Genest 2000, 2001).

Les premières études de dynamique moléculaire sur le récepteur Neu/ErbB2 ont été réalisées sans tenir compte de l'environnement membranaire. Les simulations *in vacuo* sont une approximation valable de cet environnement membranaire si l'on compare les constantes diélectriques relatives (dans le vide  $\epsilon = 1$  et dans la membrane  $\epsilon = 2-4$ ).

Garnier et collaborateurs ont montré que le peptide ErbB2 muté Val659/Glu est plus flexible que son homologue sauvage (Garnier et al. 1994). Toutefois, la structure hélicoïdale de ces

deux peptides est conservée, néanmoins il apparaît des changements conformationnels tels une transition  $\alpha \rightarrow \pi$ , au niveau du site de la mutation et en aval, ces résultats sont aussi observés dans le cas de Neu sauvage et muté (Samna Soumana et al. 2005) (**voir annexe 2**). La présence de déformation  $\pi$  a été également observée dans le cas de ErbB2 sauvage et muté Val659/Glu (Duneau et al. 1996). Elles sont observées aussi bien dans le vide que dans un environnement lipidique composé de DLPE (dilauroyl-phosphatidylethanolamine), ou de DMPC (Dimirystoyl-phosphatidylcholine) (Duneau et al. 1997; Duneau et al. 1996; Samna Soumana et al. 2005). L'origine de la déformation vient de la succession de résidus  $\beta$  branchés (comme les résidus valine) suivis d'une glycine, qui agissent comme déstabilisateur de la structure en hélice  $\alpha$ . D'autres séries de simulations ont montré que la mutation oncogénique n'empêche pas la transition locale  $\alpha \rightarrow \pi$ .

Par mécanique et dynamique moléculaires la recherche des modèles d'homodimérisation des segments TM de Neu, ErbB2 proto-oncogénique et oncogénique a été effectuée en se basant sur les données biologiques (Duneau et al. 1999; Duneau et al. 1999; Duneau et al. 1997; Duneau et al. 1996; Garnier et al. 1997; Garnier et al. 1994; Sajot et al. 1999; Sajot and Genest 2000, 2001). Il a été montré que les hélices ont une préférence pour les interactions de type gauche. Le modèle de dimère muté Val659/Glu est stabilisé par des liaisons hydrogène inter-hélices entre les chaînes latérales de Glu, en accord avec les résultats expérimentaux (Smith et al. 1996). Le rôle du motif VEG a été mis en évidence pour l'association des hélices TM. Au cours de ces simulations, différents modes d'association des hélices TM de ErbB2 ont été observés, et sont également décrits par les techniques Monte Carlo (Kim et al. 2003). Dans la plupart des simulations sur les dimères, les hélices présentent des hernies  $\pi$ .

L'étude des homodimères du domaine TM des récepteurs ErbB2/Neu dans un environnement membranaire formé de DMPC (Aller et al. 2005) ou de POPC (Palmitoyloléyl-phosphatidylcholine) (Aller et al. 2006) a fait l'objet d'une thèse effectuée par Pierre Aller.

#### 4 Le récepteur entier

Il n'existe pas de structure cristallographique complète des récepteurs de la famille ErbB. Seules des informations structurales « morcelées » sur les domaines extracellulaire,

intracellulaire et transmembranaire sont disponibles, mais permettent l'ébauche d'une vision globale de la structure de ces récepteurs.

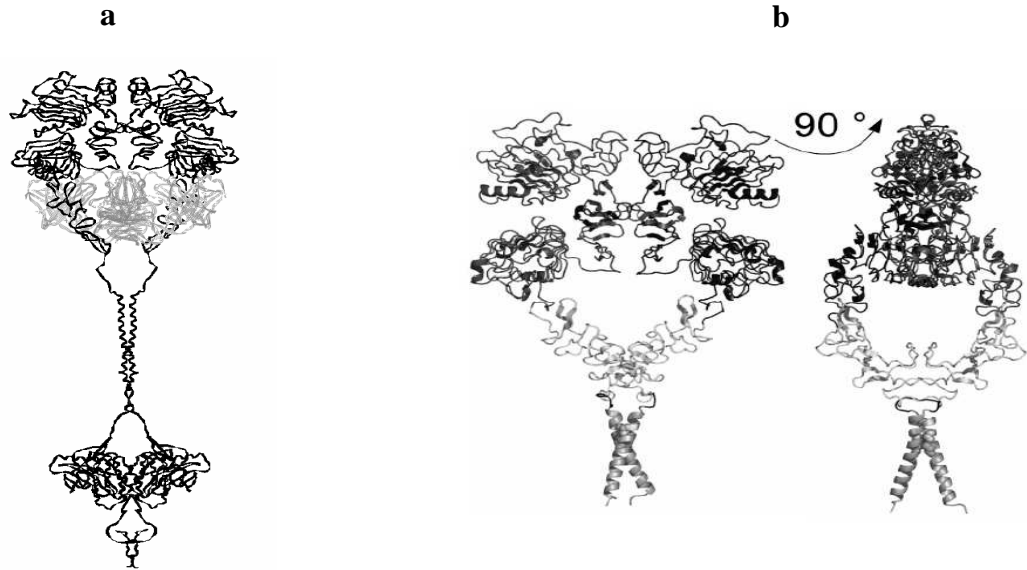
Récemment des auteurs (Bagossi et al. 2005) ont proposé un modèle de la structure complète du récepteur ErbB2 complexé à l'herceptine Fab obtenu par modélisation par homologie. Les données utilisées sont comparées à celles de Duneau et collaborateurs (Duneau et al. 2007) et résumées dans le **tableau 1.3** suivant. Les deux structures obtenues sont représentées sur la **figure 1.12**.

Les auteurs ont utilisés pratiquement les mêmes données même si pour le domaine TM, Bagossi et collaborateurs ont fait appel à plusieurs structures contrairement à Duneau et collaborateurs. De plus ces derniers n'ont pas inclut la partie intracellulaire de ces récepteurs.

La structure dimérique du domaine extracellulaire est compatible avec la structure dimérique du domaine TM. Le modèle suggère une relation étroite entre ces deux domaines (Duneau et al. 2007). De plus, ils montrent des hélices du domaine TM et une autre boucle du second domaine riche en cystéine à l'interface.

Données utilisées	(Bagossi et al. 2005)	(Duneau et al. 2007)
Domaine extracellulaire	Combinaison des structures : - 1M6B (Cho and Leahy 2002; Ogiso et al. 2002) - 1IVO (Cho and Leahy 2002; Ogiso et al. 2002)	Combinaison des structures : - 1M6B (Cho and Leahy 2002; Ogiso et al. 2002) - 1IVO (Cho and Leahy 2002; Ogiso et al. 2002)
Domaine TM	Combinaison des structures : - 1IIIJ (Goetz et al. 2001) - 1AFO (MacKenzie et al. 1997) - 1PRC (Deisenhofer et al. 1995) - 1AP9 (Pebay-Peyroula et al. 1997)	- 1AFO (MacKenzie et al. 1997)
Partie juxtamembranaire	Méthode de prédiction sur <a href="http://www.expasy.org">http://www.expasy.org</a>	
Domaine intracellulaire	- 1M14 (Stamos et al. 2002)	

**Tableau 1.3** : Comparaison des différentes données utilisées pour la modélisation de l'homodimère ErbB1 (Duneau et al. 2007) et de l'homodimère ErbB2 (Bagossi et al. 2005).



**Figure 1.12. : Modèles des récepteurs entiers ErbB1 et de ErbB2.**

**a)** Modèle du dimère ErbB2 complexé à l'herceptine Fab. Représentation en ruban du modèle de la structure presque complète du dimère ErbB2 complexé à l'herceptine Fab (en gris) (Bagossi et al. 2005).

**b)** Modèle du dimère ErbB1. Les deux domaines solubles de la structure cristallographique sont en gris foncé et le domaine riche en cystéine en gris clair. Les quatre résidus du domaine juxtamembranaire sont colorés en noir. Le domaine TM est modélisé sur la base de la structure de la GpA (1AFO) (MacKenzie et al. 1997) en utilisant comme moule le premier motif SmxxxSm (Sm pour small residue : Gly, Ala, Ser ou Thr et x, un résidu quelconque) du récepteur ErbB1. D'après (Duneau et al. 2007).

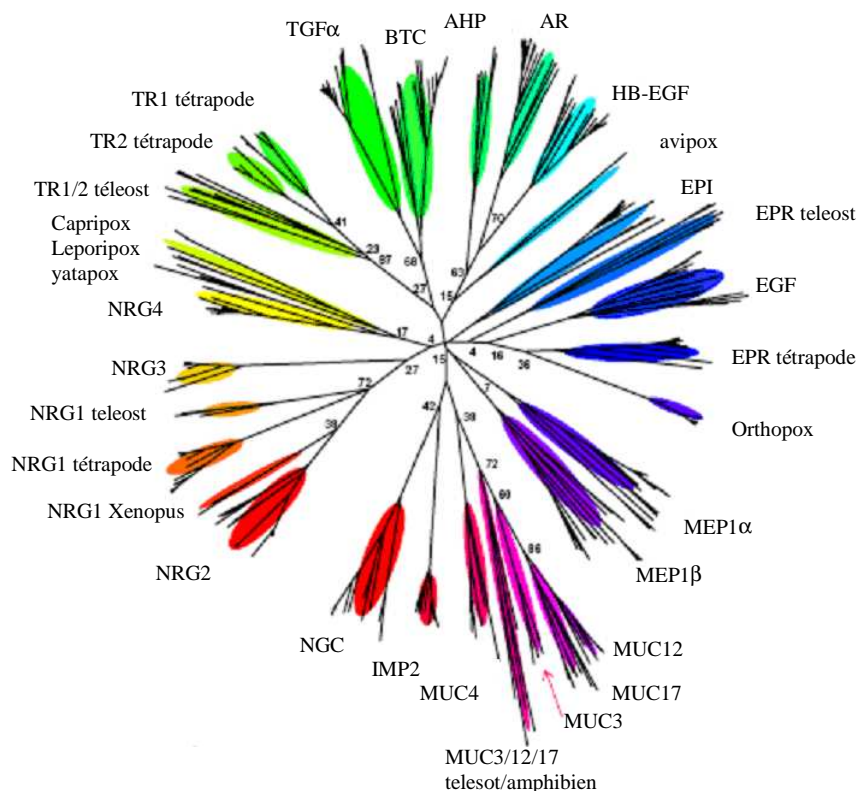
## **B Activation**

### **1 Les ligands**

La famille du facteur de croissance de l'épiderme (EGF) est constituée de plusieurs ligands. Stein et Staros (Stein and Staros 2006) ont donné une représentation en arbre phylogénétique des ligands de la famille de l'EGF chez les vertébrés. Cette étude sur les ligands des récepteurs de la famille ErbB est la plus récente. Ces auteurs ont généré un arbre sans racine montrant les différents ligands de la famille de l'EGF sous forme de branche (**figure 1.13**). Des différences existent selon la méthode utilisée pour générer l'arbre, toutefois, il y a certaines similarités nonobstant la méthode d'analyse utilisée. Généralement, l'arbre sépare les ligands du récepteur EGFR de ceux des récepteurs ErbB3 et ErbB4. Une autre caractéristique de l'arbre est l'apparition des paires apparentes des ligands suggérant des duplications de gène. Par exemple au sein de la branche des ligands du récepteur de l'EGF, ces paires sont : TR1/TR2, TGF $\alpha$ /BTC, AR/HB-EGF, et EPR/EGF.

Chez les mammifères, chaque ligand des récepteurs ErbB contient un domaine de type EGF, constitué approximativement de 60 acides aminés, nécessaire et suffisant pour l'activation (Barbacci et al. 1995), lui conférant sa spécificité de fixation (Holbro and Hynes 2004). Il est aussi marqué par l'existence d'un motif structural de type immunoglobuline, de sites de fixation de l'héparine et des sites de glycosylation (Normanno et al. 2006).

Cette étude montre que la présence de motifs facilement identifiables permet de discriminer les ligands de la famille de l'EGF des modules du type EGF. La famille de l'EGF pourrait compter plusieurs nouveaux ligands. Ces résultats suscitent aussi de nombreuses questions sur l'évolution de ErbB2 et ErbB3, sur l'activité tyrosine kinase de ErbB3, *etc.*



**Figure 1.13 : Arbre phylogénétique des modules EGF des ligands des récepteurs ErbB des vertébrés.** Cet arbre est généré en utilisant le rapprochement entre voisins avec une correction de Poisson sur des séquences protéiques avec MEGA (Molecular Evolutionary Genetics Analysis) version 3.1 (Kumar et al. 2004). Chaque zone ovale colorée correspond à un cluster de ramifications pour un ligand différent. Certains pourcentages autosuffisants pour la séparation entre deux familles de ligands sont montrés. Malgré le fait que ces pourcentages soient peu fiables pour certaines ramifications de l'arbre, des arbres générés par différentes méthodes de correction de la distance ont montré des séparations similaires des ligands du récepteur de l'EGF et des ligands de ErbB3/ErbB4. Les positions relatives entre eux sont comparables. Des arbres similaires ont été générés en utilisant le groupe de programme Phylip version 3.5 (Felsenstein 1989). D'après (Stein and Staros 2006).

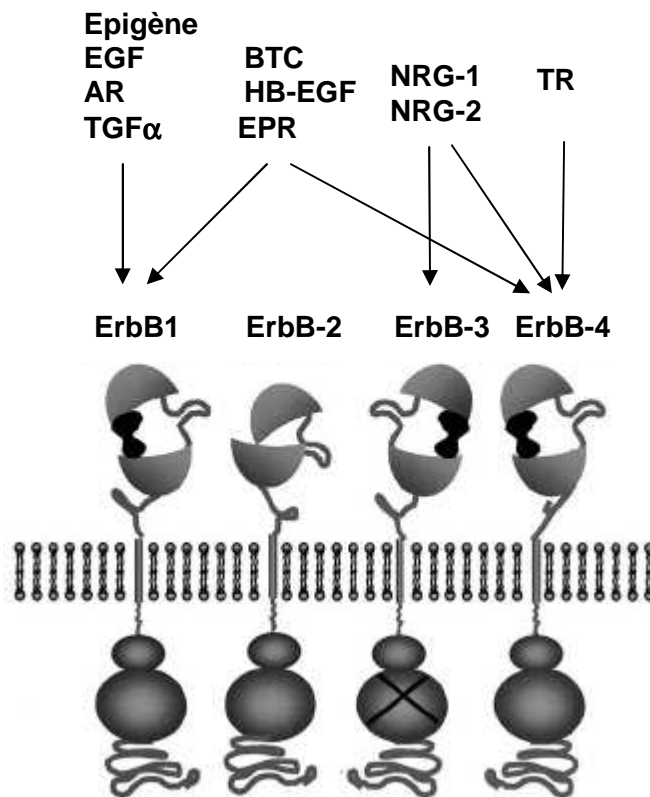
La **figure 1.14** résume les différentes interactions entre les récepteurs et les principaux ligands.

Le récepteur ErbB2 n'a pas de ligand connu à ce jour (Citri et al. 2003; Citri and Yarden 2006). L'absence de ligand capable d'activer spécifiquement le récepteur ErbB2 n'a pas permis des études directes de la dimérisation, mais sa recherche a abouti à la découverte d'une nouvelle famille de molécules : les neurégulines (NRG) ou hérégulines (HRG) qui sont des facteurs de différenciation de la protéine Neu, NDF, GGF ou ARIA (Carraway et al. 1997; Chang et al. 1997; Harari et al. 1999; Zhang et al. 1997). Ces molécules se présentent sous différents isoformes et sont les ligands des récepteurs ErbB3 et ErbB4.

Plusieurs ligands sont spécifiques de certains récepteurs. Ces ligands peuvent être scindés en trois groupes. Ainsi nous avons des ligands reconnus spécialement par ErbB1, ceux qui se



fixent sur ErbB1 et ErbB4 et un autre groupe formé de ligands se fixant sur ErbB3 et ErbB4 (NRG1 et NRG2) ou sur ErbB4 seulement (NRG3 et NRG4).



**Figure 1.14 : Récepteurs ErbB et leurs ligands.** Les récepteurs ErbB sont décrits comme des structures transmembranaires multi lobulaires. Le domaine extracellulaire comprend 2 parties riches en cystéine (domaines II et IV représentés par des boucles), impliqués dans la dimérisation induite par la fixation de ligand et 2 régions sans cystéine (domaines I et III) permettant la fixation de facteur de croissance (globule noir). Une dizaine de ligands connus peuvent se fixer sur les récepteurs de la famille ErbB. Il n'existe pas de ligand connu pour le récepteur ErbB2. Le domaine TK intracellulaire est bi-lobulaire, et est entouré par un court domaine juxta membranaire et par une longue queue carboxy-terminale contenant les sites de phosphorylation des tyrosines. Le domaine TK de ErbB3 est inactif (porte une croix). D'après (Marmor et al. 2004).

Bien que ne se liant pas directement à ErbB2, il a été montré que les ligands EGF et NRG peuvent activer ce récepteur par un mécanisme d'hétérodimérisation avec ErbB1, ErbB3 et ErbB4 (Beerli et al. 1995; Karunagaran et al. 1996).

## 2 L'activation des récepteurs ErbB

La fixation d'un ligand sur la partie extracellulaire des récepteurs ErbB induit la formation d'homo- ou d'hétéro-dimères puis l'activation du domaine TK (Olayioye et al. 2000).

L'activation des récepteurs conduit à une phosphorylation des résidus tyrosine de la queue C-terminale. Une fois phosphorylés, ces résidus sont utilisés comme des sites d'amarage pour les protéines contenant les domaines PTB et SH2. C'est ce qui initie le réseau de signalisation intracellulaire (Normanno et al. 2006).

Bien que la dimérisation induite par la fixation du ligand soit considérée depuis une vingtaine d'années comme une étape cruciale de l'activation des récepteurs ErbB, les mécanismes d'interactions étaient inconnus jusqu'en 2002 avec la résolution de la structure des domaines extracellulaires de ErbB1 (Ferguson et al. 2003; Garrett et al. 2002; Ogiso et al. 2002), ErbB2 (Cho et al. 2003; Franklin et al. 2004; Garrett et al. 2003), ErbB3 (Cho and Leahy 2002) et ErbB4 (Bouyain et al. 2005). Ces structures montrent qu'en absence de ligand, ErbB1, ErbB3 et ErbB4 sont dans une conformation inactivée au niveau intramoléculaire dans laquelle le motif de dimérisation est incapable de permettre les interactions monomère-monomère. Le récepteur ErbB1 lié au ligand, et le récepteur ErbB2 (qui n'a pas de ligand connu) sont dans une conformation favorable à la dimérisation. Dans cette conformation, le bras de dimérisation (une séquence de 10 résidus) est exposé à la surface du récepteur (Linggi and Carpenter 2006). Les contacts entre les domaines extracellulaires des deux monomères se font principalement au niveau de ce bras de dimérisation, bien que d'autres contacts existent. Des expériences de mutagenèse du bras (Garrett et al. 2002; Ogiso et al. 2002) ou de la boucle adjacente (Dawson et al. 2005) suppriment l'activation du récepteur mais permettent la fixation du ligand. En l'absence de fixation de ligand, ce bras est inactivé par des contacts intramoléculaires, empêchant les interactions intermoléculaires des monomères.

Des mutations supposées relâcher le bras inactivé, augmentent l'affinité de fixation du ligand mais ne provoquent ni la dimérisation ni l'activation des récepteurs (Ferguson et al. 2003; Mattoon et al. 2004; Walker et al. 2004). Ces résultats montrent que l'exposition de ce bras suite à la fixation du ligand est nécessaire mais pas suffisante pour entraîner la dimérisation de ces récepteurs. La dimérisation induite par la fixation du ligand est supposée entraîner l'activation de la kinase. Ainsi, les domaines cytoplasmiques se retrouvent dans une conformation permettant une transphosphorylation.

La dimérisation induite par le ligand est un processus réversible. De ce fait, certains dimères se dissocient pour former des monomères. En présence de ligand EGF, certains des monomères phosphorylés vont dimériser pour former des dimères ayant des résidus tyrosine phosphorylés (Gamett et al. 1997; Graus-Porta et al. 1997). C'est par ce « trans-mécanisme » que le récepteur ErbB2, transphosphorylé et actif après avoir dimérisé avec ErbB1 en présence de l'EGF, peut aussi phosphoryler d'autres récepteurs ErbB2 (Spivak-Kroizman et al. 1992). Ces monomères phosphorylés peuvent aussi interagir avec un monomère non phosphorylé de ErbB3 ou de ErbB4 et les activer pour former un second dimère (Gamett et al. 1997; Graus-Porta et al. 1997).

*La dimérisation et l'activation de ces récepteurs impliquent très fortement le domaine TM. Ce domaine joue aussi un rôle dans l'internalisation des récepteurs et sur l'affinité du récepteur pour son ligand. En outre, il est activement impliqué dans le processus de transduction du signal de l'extérieur vers l'intérieur de la cellule.*

*La compréhension du mode d'association des domaines TM de ces récepteurs constitue le principal objectif de cette étude. L'association de ces domaines TM se fait de deux manières : par homo ou hétérodimérisation.*

#### a\_ L'homodimérisation

Tandis que les homodimères sont la forme naturelle de ErbB1 et ErbB4, ceux de ErbB3 sont inactifs. D'après Holbro et Hynes (Holbro and Hynes 2004), la surexpression du récepteur ErbB2 déplace l'équilibre monomère  $\rightleftharpoons$  homodimère conduisant ainsi à l'activation des récepteurs sans fixation de ligand. Cette situation est notamment observée dans certains cancers (Yarden and Sliwkowski 2001).

Les études structurales des homodimères du domaine TM des récepteurs de la famille ErbB/HER (plus particulièrement de ErbB2) ont été initiées au sein de l'équipe par Norbert et Monique (Garnier et al. 1997). Elle repose sur le modèle proposé par Popot et Engelman (Popot and Engelman 1990). Dans ce modèle, deux hélices préformées vont s'associer pour former un dimère stable.

Ces premières simulations des dimères de ErbB2 ont été conduites en l'absence d'environnement membranaire (Garnier et al. 1997) puis dans un modèle de membrane formé de DMPC (Garnier et al. 2003). Récemment des simulations des homodimères du domaine TM de Neu sauvage et muté ont été effectuées dans de modèles de membrane de DMPC (Aller et al. 2005) et de POPC (Aller et al. 2006) au cours de la thèse de Pierre Aller.

Je ne reviendrai donc pas sur les homodimères.

#### b\_ L'hétérodimérisation

Sous l'influence des ligands, les 4 récepteurs de la famille ErbB peuvent s'associer pour former soit des homodimères soit des hétérodimères (Alroy and Yarden 1997). Bien que

les homodimères actifs soient les formes naturelles pour ErbB1 et ErbB4, la formation d'un hétérodimère suite à la fixation d'un ligand est souvent observée.

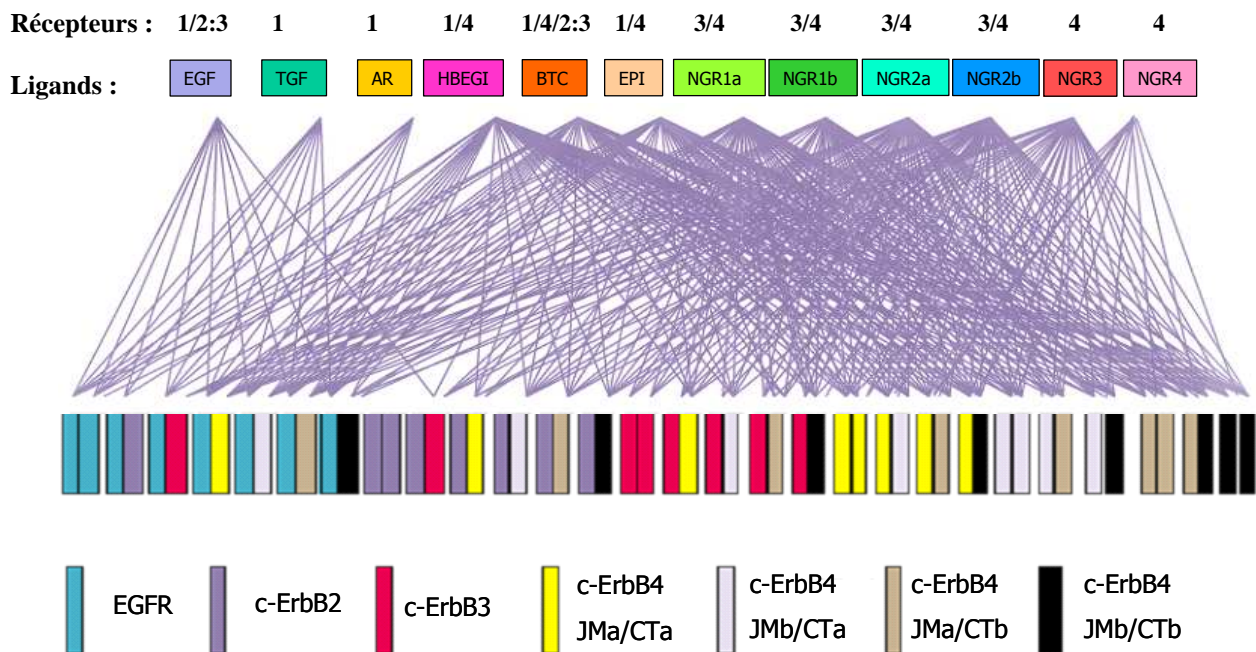
Plus particulièrement, l'hétérodimérisation stimulée par un ligand est un pré-requis pour une signalisation activée par les récepteurs ErbB2 et ErbB3 (Harari and Yarden 2000). ErbB2, une fois stimulé, est seulement recruté comme co-récepteur avec un autre membre de la famille ErbB (Klapper et al. 1999). En fait, il n'y a pas de ligand connu qui active le récepteur ErbB2. Ainsi, la fixation des ligands au récepteur ErbB2 se fait toujours de manière indirecte et fait intervenir un autre récepteur de cette famille pour former un hétérodimère (Karunagaran et al. 1996). A l'inverse, ErbB3 lie plusieurs ligands à haute affinité mais, est exempt d'activité tyrosine kinase, c'est pourquoi un co-recrutement avec un autre membre de la famille ErbB est nécessaire pour être trans-activer (Carraway and Cantley 1994). Une conséquence de ce modèle d'hétérodimérisation est qu'un récepteur particulier requiert deux autres composants de la signalisation pour initier son activation : un co-récepteur et un ligand de haute affinité. Le concept de ce modèle est essentiel à la compréhension des mécanismes liés aux stades précoces de certains cancers impliquant ErbB2.

En effet, ErbB2 n'est pas activé seul mais plutôt en concert avec au moins deux autres composants de sa voie de signalisation. Une surexpression d'ErbB2 peut augmenter sa capacité à former des hétérodimères avec un récepteur activé par son ligand. En accord avec le modèle d'hétérodimérisation, l'activité transformante des mutants Neu/ErbB2 de rongeurs dépend de la protéine ErbB coexprimée (Cohen et al. 1996).

Ainsi, les récepteurs ErbB surexprimés peuvent induire une carcinogénèse essentiellement lorsqu'un ligand induit la formation d'un hétérodimère (Harari and Yarden 2000). Alroy et Yarden (Alroy and Yarden 1997) ont décrit un certain nombre de combinaisons d'hétérodimères possibles au sein de la famille ErbB/HER, ce qui a pour effet d'augmenter la diversité des signaux intracellulaires. En 2001, Gūllick (Gullick 2001) propose une représentation schématique du réseau des couples ligands-dimères des récepteurs ErbB. En tenant compte de toutes les formes des récepteurs, 28 dimères sont potentiellement observables dans une cellule. Ce nombre est encore plus élevé en prenant en compte les ligands impliqués dans cette dimérisation (**figure 1.15**).

L'hétérodimérisation des récepteurs ErbB est un moyen d'amplification et de diversification du signal. Les hétérodimères sont plus puissants et plus mitogéniques que les homodimères (Marmor et al. 2004). La preuve de l'hétérodimérisation des récepteurs ErbB vient de la mise

en évidence par Stern et Kamps (Stern and Kamps 1988) de la phosphorylation sur les résidus tyrosine de ErbB2 suite à la fixation du ligand EGF sur le récepteur EGFR. L'expression cellulaire d'une forme mutée de ErbB2 dépourvue d'activité tyrosine kinase peut supprimer la signalisation normale de l'EGFR après stimulation par le ligand EGF (Qian et al. 1994). Grâce à l'utilisation d'agents pontants bifonctionnels, Wada et collaborateurs ont détecté la formation d'hétérodimères ErbB1-ErbB2 (Wada et al. 1990). D'autres hétérodimères impliquant les récepteurs de la famille ErbB ont aussi été décrits. Ainsi il a été mis en évidence une phosphorylation de ErbB2 grâce à une hétérodimérisation avec les récepteurs ErbB3 ou ErbB4 suite à la fixation de leur ligand, l'héréguline (Wallasch et al. 1995). Graus-Porta et collaborateurs ont aussi montré que ErbB2 était le partenaire privilégié pour s'hétérodimériser avec l'EGFR en présence de l'EGF et avec ErbB3 et ErbB4 sous l'effet des NRG (Graus-Porta et al. 1995). Ces résultats ont été obtenus grâce à l'expression intracellulaire d'anticorps monoclonaux, capables d'inhiber la translocation d'ErbB2 à la surface cellulaire. ErbB3 et ErbB4 peuvent s'associer avec l'EGFR sous l'action des NRG et des ligands de l'EGFR en l'absence de ErbB2. Dans ce cas cependant, on observe une phosphorylation réduite de ces récepteurs. Cela suggère un rôle important de ErbB2 dans l'activation complète des autres récepteurs ErbB. De plus, l'activité transformante d'ErbB2 est significativement augmentée quand ce récepteur est co-exprimé soit avec ErbB1, soit avec ErbB3. Cependant, dans les deux cas, les ligands respectifs semblent essentiels (Kokai et al. 1989; Wallasch et al. 1995). Li et collaborateurs ont montré que NRG1 induit seulement la trans-phosphorylation entre les récepteurs des hétérodimères ErbB2-ErbB3 ou ErbB2-ErbB4. La trans-phosphorylation du récepteur ErbB2 a uniquement lieu en présence de ErbB4 et non celle de ErbB3, car ce dernier est dépourvu d'activité tyrosine kinase. Il apparaît que les interactions entre les hétérodimères ne résultent pas de la cis-phosphorylation de ErbB2 (Li et al. 2007).



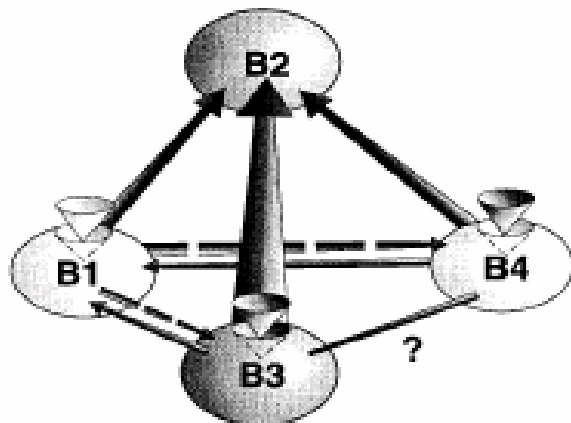
**Figure 1.15 : Représentation schématique du réseau des couples ligands récepteurs ErbB.** Les lignes en violet clair indiquent les interactions spécifiques entre les récepteurs dimériques et leurs ligands. JMa/JMb/CTa/CTb indiquent les différentes formes de ErbB4 après épissage alternatif. Les numéros sous les dimères indiquent les récepteurs interagissant ou non avec les différents ligands. D'après (Gullick 2001).

### c\_ Hiérarchie des hétérodimères

Il y a une certaine hiérarchie dans la formation des hétérodimères des récepteurs ErbB. Cela est dicté non seulement par la nature du ligand mais aussi par la composition en récepteurs de la cellule. Ainsi, le récepteur ErbB2 est décrit comme étant préférentiellement retrouvé dans la formation des hétérodimères (**figure 1.16**) (Tzahar et al. 1996). Le partenaire principal de ErbB2 est le récepteur dépourvu d'activité tyrosine kinase à savoir le récepteur ErbB3. Cette forte interaction pourrait être due à une compensation de la déficience en activité TK de ErbB3. Ce dimère ErbB2-3 posséderait un potentiel mitogénique bien supérieur à celui des autres dimères existant dans cette famille (Pinkas-Kramarski et al. 1996; Tzahar et al. 1996). Suite à l'inactivation de ErbB2 il a été constaté une baisse de la phosphorylation de tyrosine du récepteur ErbB3. De plus la perte de ErbB3 empêche la surexpression du récepteur ErbB2 : cela montre clairement que pour former un hétérodimère, le partenaire privilégié de ErbB2 est le récepteur ErbB3 (Citri et al. 2003). Cet hétérodimère est le plus important (Holbro and Hynes 2004), le plus puissant sur le plan mitogénique et aussi le plus répandu

(Citri et al. 2003). Les hétérodimères contenant ErbB2 jouent un rôle central aussi bien dans la fixation du ligand que dans la transduction du signal (Muthuswamy et al. 1999).

Pour Tzahar et collaborateurs, la hiérarchie de formation des dimères ErbB est la suivante : ErbB2-3 > ErbB2-4 ≈ ErbB2-1 > ErbB1-4 > ErbB1-3. Ces auteurs n'ont pas pu observer de dimère ErbB3-4 au cours de leurs études de culture cellulaire et de co-immunoprécipitation.



**Figure 1.16 : Hiérarchie des interactions inter - récepteurs dans la famille ErbB.** Sur la figure les récepteurs sont sous forme ovale et les ligands sous forme de cône. L'épaisseur de chaque flèche correspond à la force de l'interaction entre les récepteurs concernés. Bien que toutes les combinaisons entre les récepteurs ErbB formant des hétérodimères existent, ceux contenant ErbB2 sont prédominants. D'après (Tzahar et al. 1996).

Normanno et collaborateurs trouvent quant à eux que les hétérodimères les plus formés à la surface de la cellule sont : ErbB2-3, ErbB2-4 et ErbB1-4 (Normanno et al. 2003). Des études récentes montrent une abondance des dimères ErbB2-3 dans des cellules produisant les récepteurs ErbB1, 2 et 3. Ces analyses ont été effectuées en présence des ligands EGF ou NRG (Shankaran et al. 2006). L'abondance d'un type de dimère est fonction de son affinité de dimérisation et aussi de la localisation relative des récepteurs formant ce dimère. Il apparaît donc que l'ordre de formation des hétérodimères est proche de celle de Tzahar et collaborateurs.

En 2007, Duneau et collaborateurs ont proposé une autre hiérarchie de formation d'hétérodimères des domaines TM des récepteurs ErbB (Duneau et al. 2007). Ces résultats sont obtenus suite à des expériences de FRET menées sur les hélices TM des 4 récepteurs ErbB. Pour ces auteurs, le dimère préférentiellement formé est ErbB1-2 suivi de ErbB2-3, de ErbB1-3, ErbB2-4 et ErbB1-4. L'association des récepteurs ErbB3 et ErbB4 semble la moins favorable.

Certaines combinaisons, particulièrement celles impliquant ErbB2, sont plus efficaces dans la division cellulaire du fait de la forte affinité du ligand et de leur stabilité dans la membrane cellulaire. Cela augmente aussi la durée du signal (Gullick 2001).

La dimérisation est régulée par le niveau d'expression des différents récepteurs ErbB dans la cellule. L'hétérodimérisation permet l'amplification du signal mais aussi sa diversification (Olayioye et al., 2000). La force et la durée du signal sont influencées par l'affinité du ligand. L'activation simultanée de deux types de récepteurs pourrait améliorer la transduction du signal par rapport au cas où il y a des homodimères.

Le domaine TM joue aussi un rôle important dans la dimérisation et la stabilisation du dimère. Il est montré que les domaines TM ErbB s'autoassocient fortement en l'absence des domaines extracellulaires dans l'ordre suivant : ErbB4-TM > ErbB1-TM  $\approx$  ErbB2-TM > ErbB3-TM pour former des homodimères (Mendrola et al. 2002).

Ces interactions pourraient impliquer le motif de dimérisation de type GxxxG présent dans la séquence du domaine transmembranaire. Ce même motif est le motif dimérisant de la glycophorine A (Russ and Engelman 1999). Cette séquence est retrouvée dans le domaine TM de tous les membres de la famille des récepteurs de l'ErbB (**voir tableau 2.1 du chapitre méthodologie**). En plus de ce motif de dimérisation situé en N-terminal, ErbB2 tout comme ErbB1 et ErbB4 contiennent un second motif de dimérisation de type GxxxG dans la partie C-terminale du segment TM (Russ and Engelman 1999; Senes et al. 2004; Senes et al. 2000). Chacun de ces motifs pourrait intervenir dans la dimérisation des domaines TM de ErbB2 dans la membrane cellulaire (Gerber et al. 2004). Le motif en N-terminal serait impliqué dans la formation d'hétérodimères alors que le motif en C-terminal serait impliqué dans l'homodimérisation des récepteurs ErbB1 et ErbB2.

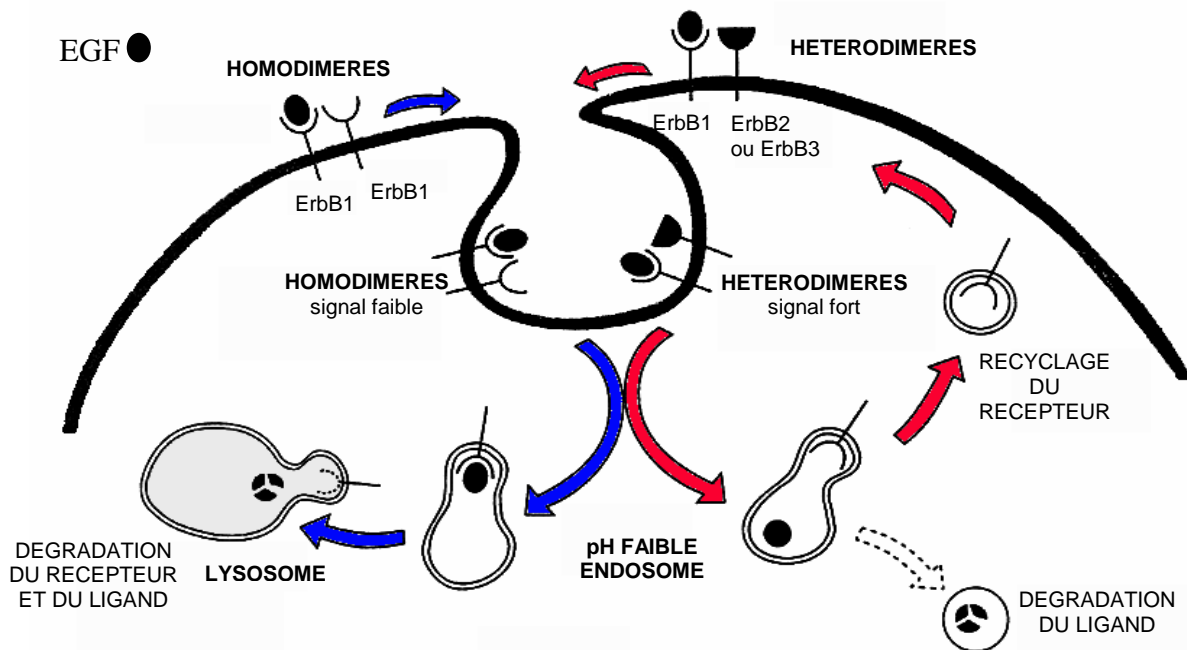
#### d\_ Dégradation des dimères

La transformation des cellules après surexpression de ErbB2 peut être également due à la nature même des ligands à motif EGF. Comme il a été décrit précédemment, ces ligands sont bivalents et se fixent avec un premier site très spécifique et très sélectif à un premier récepteur. Le second site, lui, est moins sélectif et dans tous les cas fixe ErbB2 (Tzahar et al. 1997). La présence excessive d'ErbB2 peut donc forcer la formation d'hétérodimère ErbBX-ErbB2 (X=1, 3, 4) et ainsi servir d'amplificateur du signal (Karunakaran et al. 1996; Klapper et al. 1999).

Il a été montré que les hétérodimères retardaient la dissociation du complexe ligand-récepteur (Karunakaran et al. 1996) favorisant le recyclage des récepteurs à la surface cellulaire comme



le montre la **figure 1.17**. Cela permet aussi d'augmenter la durée et l'intensité du signal (Lenferink et al. 1998).



**Figure 1.17 : Modèles de recyclage des récepteurs.** Les homodimères de ErbB1 formés sous l'action de l'EGF sont rapidement endocytés et dégradés avec le ligand au niveau lysosomal terminant ainsi la signalisation. En présence de ErbB2 ou ErbB3, les signaux induits par l'EGF sont augmentés. En effet la présence d'hétérodimères provoquera la dissociation rapide du complexe ligand-récepteur au moment de l'endocytose. Le ligand ira vers une voie de dégradation tandis que le récepteur sera recyclé à la surface de la cellule, prêt à être de nouveau activé (Lenferink et al. 1998).

Les homodimères de ErbB1 portant EGF sont dégradés dans les lysosomes alors que les hétérodimères sont régénérés à partir de l'endosome. Dans les deux voies d'endocytose (sous formes homo- et hétérodimérique) de EGF/ErbB1 il y a une dégradation intracellulaire du ligand, soit parce que les hétérodimères relâchent EGF quand le pH diminue, donc au niveau des endosomes, soit parce que le complexe ligand-récepteurs est dégradé dans le cas des homodimères au niveau des lysosomes. Les récepteurs ErbB sont recyclés à la surface cellulaire grâce aux hétérodimères alors qu'ils sont dégradés dans les lysosomes s'ils sont sous forme homodimérique.

## **CHAPITRE II : METHODOLOGIE**

## **I Introduction**

Les expériences de RMN et de cristallographie nous donnent certes des informations sur la structure des systèmes biologiques, mais l'image statique de cette structure doit être complétée par des informations sur la dynamique de ces systèmes. Les macromolécules sont des systèmes en perpétuel mouvement : ce sont donc des systèmes moléculaires dynamiques. Ces données dynamiques sont rendues accessibles grâce aux techniques de mécanique moléculaire et de simulation de dynamique. Deux niveaux sont alors à considérer. Tout d'abord le niveau microscopique, qui est représenté par le mouvement des atomes, puis le niveau macroscopique qui est représenté par les fluctuations globales de la molécule. Ces fluctuations ne sont évidemment pas visualisées par une structure tridimensionnelle statique puisqu'elle correspond, soit à un seul état piégé dans le cristal, soit à la moyenne de différents états présents en solution dans la cuve du spectromètre. Ces mouvements moléculaires sont nécessaires au bon fonctionnement des systèmes biologiques. Ils peuvent entraîner un changement important de la structure globale et peuvent être dus à la flexibilité d'une région particulière de la macromolécule. Ces deux types de mouvements sont interdépendants. Le but des simulations numériques est triple. Cette technique permet tout d'abord, au niveau microscopique d'étudier les propriétés dynamiques et structurales d'un système. Elle permet de ce fait de "mesurer" des grandeurs inaccessibles par l'expérience comme le déplacement des atomes, de caractériser les interactions et ceci pour tous les atomes. Un des avantages de cette technique est de pouvoir varier très facilement des paramètres physiques difficilement modifiables dans l'expérience, telles que la densité, la température, la pression. Enfin, les simulations permettent de tester la validité d'une approche expérimentale physique, biologique ou biochimique, de cibler les paramètres à l'origine d'un phénomène particulier.

Les premières simulations ont été effectuées par Rahman (Rahman 1964) sur l'argon liquide en 1964 et sur l'eau liquide en 1971 (Rahman 1971). Celles effectuées sur des macromolécules biologiques datent de la fin des années 1970 et ont été réalisées par McCammon et collaborateurs (McCammon et al. 1977). La première protéine étudiée est l'inhibiteur de la trypsine pancréatique bovine (BPTI), une protéine de 58 résidus durant 9.2 ps.

La formation des complexes hautement spécifiques entre les protéines et l'ADN, les protéines et leurs ligands est une nécessité pour presque tous les processus biologiques. La forme tridimensionnelle de ces complexes est souvent la clé pour une compréhension détaillée de la façon dont ils réalisent leur tâche spécifique.

Malgré les progrès technologiques réalisés pour déterminer les structures 3D des protéines isolées ou d'autres macromolécules, il existe encore de nombreuses difficultés pour accéder à la structure de très gros complexes comme les systèmes membranaires. Les simulations de dynamique moléculaire sont intéressantes dans ce domaine pour obtenir des informations structurales et d'autres paramètres comme l'élasticité, la diffusion, *etc.*

L'application de la dynamique moléculaire aux systèmes biologiques permet l'échantillonnage des régions de l'espace conformationnel pouvant être explorées. L'espace des configurations n'est pas accessible par la mécanique moléculaire. Cet échantillonnage donne à la fois des informations sur la structure et la dynamique interne des molécules.

Une autre utilisation des techniques de dynamique moléculaire est la recherche des différentes conformations d'une macromolécule. Dans ce cas, on s'intéresse plus aux mouvements macroscopiques de la molécule qu'aux mouvements microscopiques. La structure d'une molécule biologique fluctue autour d'une position moyenne. Cette notion de mouvement n'est pas accessible directement par les techniques de détermination de structure. C'est à ce niveau que la dynamique moléculaire intervient pour donner une image non statique d'une structure établie par RMN ou par cristallographie aux rayons X. Les méthodes de dynamique moléculaire classiques permettent d'explorer les configurations autour d'une structure donnée. Pour effectuer une recherche plus exhaustive, la méthode dite de recuit simulé ou dynamique moléculaire à haute température est utilisée pour franchir des barrières énergétiques suffisantes pour visualiser les différents états conformationnels. Cette étape est suivie de phases de minimisation d'énergie et de dynamique moléculaire de quelques ps.

En connaissant la position des atomes et leur vitesse à un temps  $t$ , il est possible de calculer leur position et leur vitesse à un temps  $t+\delta t$  et d'obtenir ainsi une trajectoire pour chacun d'eux. Ce calcul prend en compte les interactions entre atomes, interactions qui seront détaillées un peu plus loin. Par itérations successives, on peut accéder aux mouvements se déroulant sur une échelle de temps de l'ordre de la nanoseconde (ns). A partir de ces résultats, il est possible de recalculer des quantités macroscopiques que l'on pourra comparer aux données expérimentales disponibles.

Pour l'étude des hétérodimères de la famille des récepteurs ErbB/HER, les techniques de simulations de dynamique moléculaire ont été utilisées comme méthode de raffinement de

différents états initiaux. En simulant les mouvements atomiques il est ainsi possible de rechercher les meilleures positions des atomes répondant aux meilleures interactions. Ces simulations ont été conduites dans le vide. Les systèmes étudiés comprennent environ 1000 atomes. Grâce aux moyens de calcul disponibles au sein de l'équipe, des milliers de trajectoires d'environ 4 nanosecondes chacune ont été générées pour cette étude. Tous les calculs ont été effectués sur les plates-formes LINUX multi-processeurs de type cluster Beowulf de l'équipe.

Dans une première partie, je présenterai plus en détail le principe des techniques de simulation de dynamique moléculaire (MD). La deuxième partie portera sur le logiciel utilisé pour les calculs de MD : CHARMM et ceux utilisés pour la visualisation : VMD et MolMol. Je terminerai en décrivant la construction des dimères et les étapes de simulation de dynamique moléculaire.

## **II Principe de la dynamique moléculaire et mécanique moléculaire**

### **A Le champ de forces**

Les lois de la mécanique classique font intervenir une fonction d'énergie potentielle notée  $E_p$  qui dépend des coordonnées atomiques du système. Cette fonction  $E_p$  est décrite grâce à un ensemble de paramètres appelé champ de forces. Ces paramètres doivent être bien ajustés pour reproduire au mieux des résultats expérimentaux. A ce champ de forces on associe une expression analytique  $E_p$  de l'énergie potentielle du système. L'électron n'est pas explicitement considéré dans les calculs de dynamique moléculaire contrairement à la mécanique quantique et il est impossible de décrire des événements tels que la formation de liaisons atomiques dans une réaction chimique.

Le champ de forces est la base des techniques de simulations moléculaires. Il permet de quantifier les interactions interatomiques et intermoléculaires du système aussi précisément que possible. L'énergie potentielle, définie empiriquement, est une somme de plusieurs termes indépendants. En fait, l'énergie potentielle est universellement décomposée en une somme de termes internes (longueur de liaisons, angles de liaisons, angles dièdres et dièdres impropres), qui permettent de maintenir la géométrie covalente de la molécule et de termes

d'énergies externes (interactions de Van der Waals, électrostatiques et liaisons hydrogène) qui représentent les interactions entre atomes non liés. Les paramètres permettant d'évaluer les différents termes d'énergie sont déduits d'études détaillées sur des systèmes simples et transférés sur des molécules de plus grosse taille. On peut également introduire des termes de contraintes, de position, de géométrie, *etc.*, pour rendre compte des conditions expérimentales connues.

L'énergie potentielle totale s'exprime comme la somme des termes des énergies internes et externes. Dans le logiciel CHARMM l'équation utilisée pour le calcul de cette énergie potentielle est la suivante :

$$E_p = \underbrace{E_b + E_\theta + E_\phi + E_\omega}_{\text{Termes internes}} + \underbrace{E_{vdW} + E_{el} + E_{hb}}_{\text{Termes externes}} + \underbrace{E_{cr} + E_{cd} + E_{c\phi}}_{\text{Contraintes}} \quad (1)$$

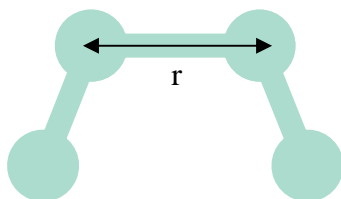
Nous décrivons dans ce paragraphe la forme générale d'une fonction d'énergie potentielle.

## 1 Interactions entre atomes liés

Les différents termes représentant les interactions entre atomes liés, c'est-à-dire séparés par au plus trois liaisons, sont décrits dans cette partie. Chacun de ces termes s'exprime sous la forme d'un potentiel harmonique et rend compte du coût énergétique lié à la déformation de géométrie des longueurs de liaisons, des angles de valence et des angles dièdres impropres et des angles de torsion.  $r_0$ ,  $\theta_0$  et  $\omega_0$  correspondent respectivement aux valeurs standard des longueurs de liaisons, des angles de valences et des angles dièdres impropres. Ces valeurs sont déduites des études cristallographiques.  $k_b$ ,  $k_\theta$  et  $k_\omega$ , les constantes de forces qui leur sont associées.  $k_\phi$  est la constante de force liée aux angles de torsion.

### a\_ Elongation des liaisons covalentes

$$E_b = \sum_{\text{liaisons}} k_b (r - r_0)^2 \quad (2)$$



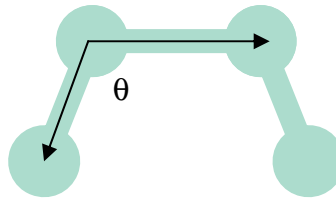
**Figure 2.1 :** Longueur  $r$  entre 2 atomes liés.

$E_b$  est minimale pour

$$r = r_0$$

b\_ Distorsion des angles de valence

$$E_{\theta} = \sum_{\text{angle-valence}} k_{\theta} (\theta - \theta_0)^2 \quad (3)$$



**Figure 2.2 :** Angle de valence.  
 $E_{\theta}$  est minimale pour  $\theta = \theta_0$ .

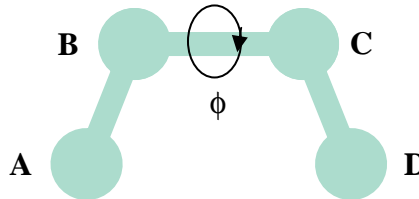
c\_ Angle de torsion

$$E_{\phi} = \sum_{\text{angle}\phi} |k_{\phi}| - k_{\phi} \cos(n\phi) \quad (4)$$

$k_{\phi}$  : constante de force,

$\phi$  : angle dièdre

n est la multiplicité de la rotation et est égal à 1, 2, 3, 4, 6

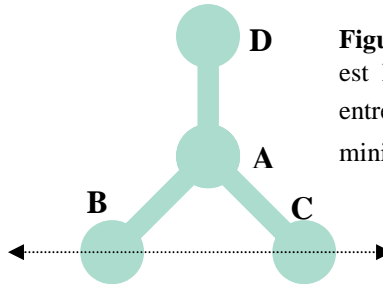


**Figure 2.3 :** L'angle dièdre  $\phi$  est l'angle entre les plans  $ABC$  et  $BCD$ .

Cette fonction sinusoidale d'énergie potentielle résulte de la torsion autour des liaisons simples. La constante de force  $k_{\phi}$  et la valeur de n sont sélectionnées dans les paramètres de CHARMM selon les atomes impliqués dans la torsion.

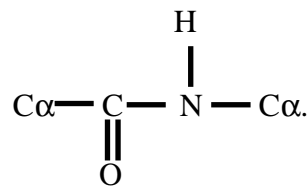
d\_ Angle dièdre impropre

$$E_{\omega} = \sum_{\text{angle-dièdre}} k_{\omega} (\omega - \omega_0)^2 \quad (5)$$



**Figure 2.4 :** L'angle dièdre impropre  $\omega$  est l'angle de torsion  $\phi_{ABCD}$  ou l'angle entre les plans  $ABC$  et  $BCD$ .  $E_{\omega}$  est minimale pour  $\omega = \omega_0$ .

Cette fonction s'applique à un groupe de 4 atomes qui ne sont pas linéairement liés. Elle permet de maintenir la planéité d'un groupe de 4 atomes comme dans le cas de la liaison peptidique



## 2 Interactions entre atomes non liés

Il s'agit ici des différents termes d'énergies représentant les interactions entre atomes séparés par plus de trois liaisons. Ces termes décrivent les interactions de van der Waals, électrostatiques et de liaisons hydrogène.

### a\_ Interactions de van der Waals

L'énergie de van der Waals est représentée par une fonction de Lennard-Jones qui peut être modulée par une fonction « switch ».

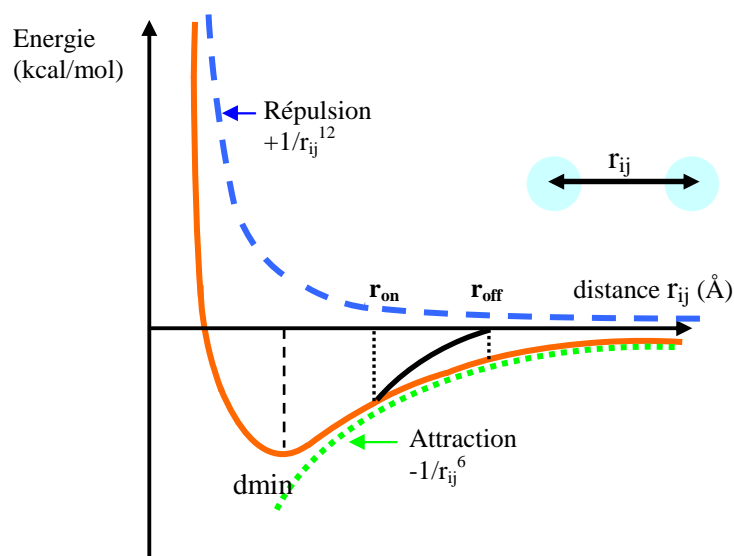
$$E_{vdW} = \sum_{\text{paires}(ij)} \left( \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) sw(r_{ij}^2, r_{on}^2, r_{off}^2) \quad (6)$$

Les paramètres  $A_{ij}$  et  $B_{ij}$  sont des constantes de van der Waals dépendants de la nature des atomes  $i$  et  $j$  et sont évalués à partir de polarisabilité atomique.

L'énergie de van der Waals entre 2 atomes non liés  $i$  et  $j$ , distants de  $r_{ij}$ , comprend un terme attractif et un terme répulsif. Le terme attractif connu historiquement sous le nom d'énergie de dispersion de London varie en  $-\frac{1}{r_{ij}^6}$  tandis que le terme répulsif qui traduit le principe de

Pauli, varie en  $\frac{1}{r_{ij}^{12}}$  et représente la répulsion des nuages électroniques à courte distance.

L'énergie de van der Waals entre 2 atomes  $i$  et  $j$  distants de  $r_{ij}$  est illustrée dans la figure 2.5.



**Figure 2.5 :** Représentation de la courbe d'énergie de van der Waals en fonction de la distance  $r$  sous la forme d'un potentiel Lennard-Jones (en rouge) : [somme des termes répulsif (en bleu) et attractif (en vert)]. Pour diminuer le temps de calcul, l'énergie est progressivement amenée à zéro (noir) dans l'intervalle  $r_{on} < r < r_{off}$ . Pour  $r > r_{off}$ , les interactions ne sont pas calculées.



Dans l'équation (6),  $sw$  désigne une fonction « switch » appliquée lorsque  $r_{ij}$  est compris entre  $r_{on}$  et  $r_{off}$ . La fonction « switch » amène progressivement l'énergie d'interaction de van der Waals à une valeur nulle dans l'intervalle  $[r_{on}, r_{off}]$ .

La fonction « switch » est définie de la manière suivante :

$$\begin{aligned}
 sw(r, r_{on}, r_{off}) &= 1 && \text{si } r \leq r_{on} \\
 sw(r, r_{on}, r_{off}) &= \frac{(r_{off} - r)^2 (r_{off} + 2r - 3r_{on})}{(r_{off} - r_{on})^3} && \text{si } r_{on} < r \leq r_{off} \\
 sw(r, r_{on}, r_{off}) &= 0 && \text{si } r > r_{off}
 \end{aligned}$$

Cette fonction évite le calcul des interactions de van der Waals à longue distance. Cette approximation est meilleure que celle de la distance de coupure (cutoff). On définit une distance maximale (cutoff) entre deux atomes au delà de laquelle les énergies d'interaction ne sont plus calculées. Le problème de cette technique est la brutalité de son exécution du tout ou rien. Il n'y a pas de continuité dans le calcul des interactions comme dans le cas de l'application de la fonction « switch ».

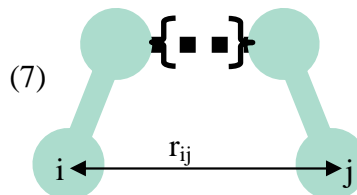
L'énergie d'interaction de van der Waals est minimale lorsque les deux atomes sont à une distance  $d_{min}$ , telle que l'équilibre de forces attractives et répulsives soit atteint.

## b\_ Interactions électrostatiques

La plupart des atomes porte une charge partielle, même si globalement la molécule est neutre, cela conduit à une distribution électronique asymétrique.

Les interactions électrostatiques sont calculées en appliquant la loi de Coulomb (7).

$$E_{el} = \sum_{\text{paires}(ij)} \frac{q_i q_j}{4\pi\epsilon_0 \epsilon_r r_{ij}} sw(r_{ij}, r_{on}, r_{off}) \quad (7)$$



**Figure 2.6 : Interactions électrostatiques.**  
 $r_{ij}$  = distance entre les 2 atomes i et j

avec  $q_i$  et  $q_j$ , les charges partielles des atomes i et j calculées par des méthodes de mécanique quantique,  $r_{ij}$ , la distance séparant les atomes i et j,  $r_{ij}$ , la distance entre les centres de géométrie de groupe d'atomes contenant les atomes i et j,  $\epsilon_0$  constante diélectrique du vide et

$\epsilon_r$  constante diélectrique relative. La valeur de la permittivité relative permet de simuler les effets du solvant autour de la molécule lorsque celui-ci n'est pas introduit explicitement. La fonction « switch » utilisée est la même que celle introduite pour le calcul des interactions de van der Waals. La seule différence est qu'elle est appliquée à des groupes d'atomes et non à des atomes seuls.

### c\_ Liaisons hydrogène

Pour décrire les interactions de liaison hydrogène (LH), certains champs de forces utilisent généralement un potentiel supplémentaire de type Lennard-Jones 10-12 associé à la loi de Coulomb. Les liaisons hydrogène impliquent 4 atomes : AA, A, H et D. AA est l'atome qui précède l'accepteur d'hydrogène A, H est l'hydrogène et D le donneur d'hydrogène. CHARMM offre la possibilité de calculer l'énergie des liaisons hydrogène en appliquant la fonction suivante :

$$E_{hb} = \sum \left( \frac{A'}{r_{AD}^i} - \frac{B'}{r_{AD}^j} \right) \cos^m(\theta_{A-H-D}) \times \cos^n(\theta_{AA-A-H}) \times sw(r_{AD}, r_{hon}, r_{hoff}) \times sw[\cos^2(\theta_{A-H-D}), \cos^2(\theta_{hon}), \cos^2(\theta_{hoff})] \quad (8)$$

où i et j sont des entiers positifs (12 et 10 respectivement pour i et j)

$$m = (0, 2, 4)$$

$$n = (0, 2)$$

A' et B' sont des constantes dépendantes du couple donneur- accepteur

$r_{AD}$  est la distance entre accepteur et donneur

$\theta_{A-H-D}$  est l'angle entre A, H et D

$\theta_{AA-A-H}$  est l'angle entre AA, A et H

La formation de liaison hydrogène dépend de la distance A-D et des angles A-H-D ( $\theta_{A-H-D}$ ) et AA-A-H ( $\theta_{AA-A-H}$ ).

L'énergie d'interaction peut être nulle lorsque :  $\theta_{A-H-D} < 90^\circ$

où  $n > 0$  et  $\theta_{AA-A-H} < 90^\circ$

Les exposants i, j, n et m sont définis dans les paramètres de CHARMM. n et m dépendent respectivement du type d'accepteur et du type de donneur d'hydrogène.

Dans les simulations, l'énergie d'interactions de liaisons hydrogène est calculée simplement comme la somme des interactions de van der Waals et électrostatiques. L'utilisation de l'équation (8) est trop coûteuse en temps de calcul. Dans les récentes versions de champ de forces CHARMM, la paramétrisation adaptée des potentiels de van der Waals et électrostatique permet de prendre en compte la contribution de LH sans avoir besoin d'ajouter de terme explicite.

### 3 Termes de contraintes

Les termes de contraintes incluent des informations structurales connues (ex : distances  $r_{ij}$  issues des expériences RMN) ou des informations supposées pour tester des hypothèses structurales. Les différents termes de contraintes spéciaux utilisés couramment dans le programme CHARMM sont :

#### a\_ Les contraintes harmoniques de positions atomiques

Lorsqu'on simule un système moléculaire, il peut être utile de fixer la position de quelques atomes ou d'une partie du système. Lorsque l'on effectue des simulations de dynamique moléculaire il n'est pas judicieux d'immobiliser certains atomes puisque cela conduirait à réduire la flexibilité du système telles que les transitions et les mouvements qui interviennent normalement dans le système. Mais on peut contraindre des atomes spécifiques à rester autour d'une position déterminée. On applique un potentiel de contrainte harmonique qui autorise néanmoins une certaine mobilité et donc une certaine flexibilité.

$$E_{cr} = \sum K_i (\vec{r}_i - \vec{r}_{i0})^2 \quad (9)$$

avec  $K_i$  constante dépendante de la masse de l'atome  $i$ ,  $\vec{r}_i$  nouvelle position de l'atome  $i$  et  $\vec{r}_{i0}$ , sa position initiale.

Les contraintes harmoniques appliquées à la position des atomes permettent de maintenir une conformation particulière pour un ensemble donné d'atomes en évitant de grands déplacements mais en permettant tout de même la relaxation de la structure, par exemple, on peut maintenir le squelette de la protéine et relaxer les chaînes latérales.

b\_ Les contraintes harmoniques de distances interatomiques

C'est un potentiel harmonique de la forme

$$E_{cd} = \sum_{ij} K_{ij} (r_{ij} - r_{ij0})^2 \quad (10) \quad \text{entre les atomes } i \text{ et } j.$$

avec  $K_{ij}$  constante dépendante de la masse de l'atome  $i$ ,  $r_{ij}$  nouvelle distance entre les atomes  $i$  et  $j$  et  $r_{ij0}$ , la distance initiale.

c\_ Les contraintes d'angle de valence et d'angles dièdres  $\phi$

$$E_{c\phi} = \sum K_i (\phi_i - \phi_{i0})^2 \quad (11)$$

avec  $\phi_i$  et  $\phi_{i0}$  angles dièdres calculé et initial respectivement.

Les contraintes appliquées aux angles dièdres permettent le maintien d'une conformation locale.

d\_ Les contraintes de liaison rigide (SHAKE)

L'utilisation de l'algorithme SHAKE uniquement au cours de la dynamique moléculaire permet de réduire les variations des longueurs de la liaison dans la limite d'un seuil de tolérance choisi. Il permet d'augmenter l'intervalle de temps ( $\delta t = 2fs$ ) en éliminant les oscillations les plus rapides associées aux liaisons chimiques notamment de type atome lourd-atome d'hydrogène (X—H) ou entre 2 atomes lourds (X—X).

## B Algorithmes de dynamique moléculaire

Les techniques de simulation de dynamique moléculaire permettent de décrire l'évolution au cours du temps de la position des atomes (leur trajectoire) à une température donnée. La dynamique moléculaire standard est basée sur la résolution numérique de l'équation fondamentale de Newton générant ainsi les trajectoires des atomes.

Une macromolécule est définie par la position dans l'espace de chacun de ses atomes. Chaque atome est considéré comme une masse ponctuelle. A un instant  $t+\delta t$ , un atome  $i$  donné se trouvera à une nouvelle position  $r_{t+\delta t}$  du fait des différentes forces qui s'exercent sur lui, liées à la présence des autres atomes environnants. Le but de la dynamique moléculaire est de calculer la position de l'atome à cet instant  $t+\delta t$ . Les équations de la mécanique classique permettent d'établir les relations entre ces deux instants  $t$  et  $t+\delta t$ .

Les forces  $\vec{F}_i$  dérivent du potentiel  $E_p$  : 
$$\vec{F}_i = -\vec{\nabla} E_p = -\frac{\delta E_p}{\delta \vec{r}_i}. \quad (12)$$

L'équation de Newton relie la force  $\vec{F}_i$  à l'accélération  $\vec{a}_i$  pour un atome  $i$  de masse  $m_i$  :

$$\vec{F}_i = m_i \vec{a}_i = m_i \frac{d^2 \vec{r}_i}{dt^2} \quad (13)$$

Au début d'une simulation, une vitesse initiale est attribuée aux atomes pour se conformer à l'énergie cinétique totale du système. Cette énergie dépend elle-même de la température de simulation désirée. Elle est habituellement obtenue en élevant lentement la température d'un système moléculaire placé initialement au zéro absolu et en amenant progressivement le système vers l'équilibre.

La résolution analytique de l'équation du mouvement (12) est complexe. Les positions atomiques sont calculées à des intervalles de temps discrets grâce à des algorithmes d'intégration tels que celui de Verlet (Verlet 1967) ou de Leap-Frog (« saut de grenouille ») (Hockney 1981). Du fait de la relation entre les forces et l'énergie potentielle en fonction de la position à un instant  $t$ , il est possible de connaître les positions et les vitesses des différents atomes en appliquant un développement limité en série de Taylor. La position des atomes et leur vitesse initiale doivent être spécifiées pour initier l'algorithme de Verlet à l'instant  $t_0$ . La position des atomes est fournie par l'expérience (structure RMN ou cristallographique) ou déduite des expériences *in silico* (homologie, construction de modèles). Les vitesses sont choisies dans une distribution de Maxwell correspondant à la température désirée.

## 1 Algorithme de Verlet

L'emploi de cet algorithme suppose que la force  $\vec{F}_i(t)$  soit constante pendant l'intervalle de temps  $\delta t$ . Il est donc nécessaire de considérer un intervalle de temps de l'ordre de la femtoseconde, (choisi par rapport à la fréquence du mouvement de la liaison C-H), pour que  $\vec{F}_i(t)$  varie très peu. Le développement limité des positions à l'instant  $t + \delta t$  et  $t - \delta t$  est le suivant :

$$\vec{r}_i(t + \delta t) = \vec{r}_i(t) + \frac{d\vec{r}_i(t)}{dt} \delta t + \vec{a}_i(t) \frac{\delta t^2}{2!} + \vec{b}_i(t) \frac{\delta t^3}{3!} \quad (14)$$

$$\vec{r}_i(t - \delta t) = \vec{r}_i(t) - \frac{d\vec{r}_i(t)}{dt} \delta t + \vec{a}_i(t) \frac{\delta t^2}{2!} - \vec{b}_i(t) \frac{\delta t^3}{3!} \quad (15)$$

L'addition des équations (13) et (14) donne :

$$\vec{r}_i(t + \delta t) = 2\vec{r}_i(t) - \vec{r}_i(t - \delta t) + \vec{a}_i(t) \delta t^2 \quad (16)$$

Dans l'équation (16) tous les termes sont connus ou peuvent être calculés à partir de l'expression du champ de forces. La position à l'instant  $t + \delta t$  est calculée en fonction des positions à l'instant  $t$  et  $t - \delta t$ . L'accélération  $\vec{a}_i(t)$  proportionnelle à la force  $\vec{F}_i(t)$  est obtenue à partir de l'énergie potentielle  $E_p$  et de l'équation (11), elle ne dépend que de la configuration à l'instant  $t$ .

$$\vec{a}_i = \frac{\vec{F}_i}{m_i} = - \frac{dE_p}{d\vec{r}_i} \frac{1}{m_i} \quad (17)$$

Les vitesses n'apparaissent pas dans cette équation, et il n'est pas toujours nécessaire de les calculer au cours de la simulation, mais il est possible de les obtenir de la manière suivante :

$$\vec{v}_i(t) = \frac{\vec{r}_i(t + \delta t) - \vec{r}_i(t - \delta t)}{2\delta t} \quad (18)$$

L'inconvénient de la méthode de Verlet vient du fait que dans l'expression de  $\vec{r}(t)$ , un petit terme de l'ordre de  $\delta t^2$  est additionné à la différence entre deux termes d'ordre 0 en  $\delta t$ . Cette situation peut induire des erreurs de précision. Un autre algorithme dérivé dit algorithme de Leap-Frog (ou « saut de grenouille ») (Hockney 1981) a été développé dans les années 1980.

## 2 Algorithme de Leap-Frog

Il repose sur le principe suivant : les vitesses sont calculées pour des intervalles de temps demi-entiers et les positions sont obtenues pour des intervalles de temps entiers. Les vitesses sont définies au temps  $t+\delta t/2$  et  $t-\delta t/2$  :

$$\vec{v}_i(t + \delta t / 2) = \frac{\vec{r}_i(t + \delta t) - \vec{r}_i(t)}{\delta t} \quad (19)$$

$$\vec{v}_i(t - \delta t / 2) = \frac{\vec{r}_i(t) - \vec{r}_i(t - \delta t)}{\delta t} \quad (20)$$

A partir des équations (19) et (20) l'expression de la position de l'atome  $i$  à l'instant  $t+\delta t$  et  $t-\delta t$  est la suivante :

$$\vec{r}_i(t + \delta t) = \vec{r}_i(t) + \vec{v}_i(t + \delta t / 2)\delta t \quad (21)$$

$$\vec{r}_i(t - \delta t) = \vec{r}_i(t) - \vec{v}_i(t - \delta t / 2)\delta t \quad (22)$$

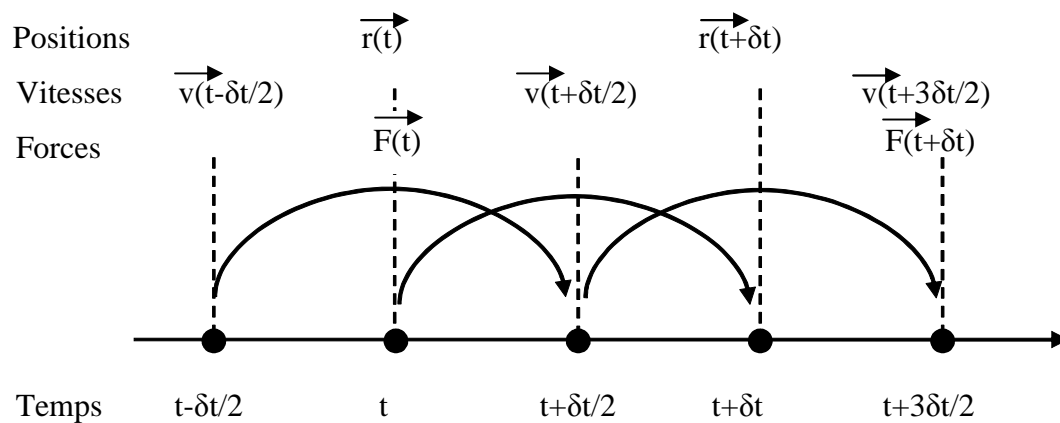
En additionnant les équations (21) et (22) on obtient :

$$\vec{r}_i(t + \delta t) = 2\vec{r}_i(t) - \vec{r}_i(t - \delta t) + \vec{v}_i(t + \delta t / 2)\delta t - \vec{v}_i(t - \delta t / 2)\delta t \quad (23)$$

En utilisant l'équation (15), on obtient l'expression suivante :

$$\vec{v}_i(t + \delta t / 2) = \vec{v}_i(t - \delta t / 2) + \frac{\vec{F}_i(t)}{m_i} \delta t \quad (24)$$

Cet algorithme est identique à celui de Verlet en ce qui concerne la génération des trajectoires, les valeurs des vitesses aux temps demi-entiers n'apparaissent que comme des intermédiaires de calcul.



**Figure 2.6 :** Schéma de principe de l'algorithme du Leap-Frog.

L'intervalle de temps  $\delta t$  doit être choisi petit par rapport à la période du mouvement de plus haute fréquence (c'est à dire celui de la liaison C-H, soit 10 fs). En général une valeur 10 fois inférieure pour le pas d'intégration soit  $\delta t = 1$  fs est utilisée. L'algorithme SHAKE, permet de s'affranchir des vibrations de liaisons rigides et de doubler le temps d'intégration soit  $\delta t = 2$  fs. Cela permet de générer des trajectoires sur des durées plus longues.

### C Méthodes de minimisation d'énergie

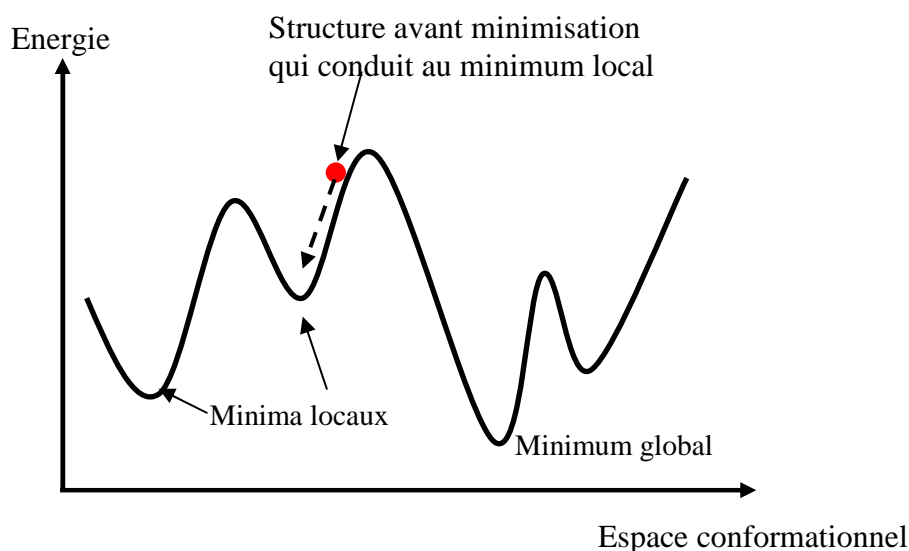
La minimisation d'énergie est l'une des premières étapes d'un protocole de MD et est utilisée aussi en alternance avec les étapes de montée en température et d'équilibration. Cette énergie est décrite dans l'espace cartésien par une fonction à  $3N$  variables ( $N$  étant le nombre d'atomes du système). La minimisation permet d'éliminer les mauvais contacts interatomiques et d'obtenir la conformation la plus stable d'une macromolécule. Cette conformation correspond généralement au minimum local le plus proche de la structure de départ.

La **figure 2.7** représente le schéma de principe de la minimisation d'énergie en fonction d'un seul paramètre. La courbe devient une surface si deux paramètres sont utilisés. Si on tient compte de plusieurs paramètres, on parle d'hyper surface d'énergie potentielle.

Il existe plusieurs méthodes de minimisation d'énergie. Dans le paragraphe qui suit nous allons donner quelques aspects des deux méthodes de minimisation utilisées au cours de cette



thèse à savoir la méthode de la plus grande pente (*Steepest Descent* ou *SD*) et la méthode ABNR pour *Adopted Basis Newton-Raphson*.



**Figure 2.7 :** Courbe d'énergie d'un espace conformationnel à une dimension

- La méthode *Steepest Descent*.

Les méthodes utilisant la dérivée première de l'énergie (ou méthodes du premier ordre) sont rapides et efficaces quand la pente est forte. L'une des méthodes couramment employée est la SD (Wiberg 1965). Cette méthode modifie la position des atomes selon la dérivée première de l'énergie (ou gradient). Elle suit donc les forces les plus importantes. On a donc :

$$\vec{S}(k) = -\frac{\vec{g}(k)}{|\vec{g}(k)|}$$

avec  $\vec{S}(k)$  direction du déplacement en fonction du gradient  $\vec{g}(k)$ , au point  $k$ .

La SD utilise une direction de recherche correspondant à l'opposé de la valeur du gradient. Elle est très efficace lorsqu'on se trouve loin du minimum et permet de supprimer les mauvais contacts stériques et d'accéder rapidement au minimum local le plus proche de la conformation de départ. Toutefois, elle présente un comportement oscillatoire autour de ce minimum local et ne converge que très lentement. Elle n'est donc pas efficace pour le raffinement autour du puits de potentiel, particulièrement pour les macromolécules.

Pour contourner cet obstacle on fait appel aux méthodes dites de dérivée seconde.

- La méthode *ABNR pour Adopted Basis Newton-Raphson* .

Autour du minimum, où le potentiel est supposé approximativement quadratique, l'algorithme de Newton-Raphson converge rapidement. Loin du minimum, cet algorithme n'est pas efficace. De plus, cette méthode nécessite beaucoup d'espace disque pour stocker les données et aussi un temps de calcul plus long. La méthode de la dérivée seconde la plus utilisée et particulièrement bien adaptée pour les gros systèmes tels que les protéines est l'ABNR.

Cette méthode utilise une base limitée au sous espace dans lequel le système a fait le plus de progrès antérieurement.

Les nouvelles positions sont déterminées dans la méthode Newton-Raphson. L'ABNR est une combinaison des meilleurs aspects des méthodes de la dérivée première en terme de rapidité et de stockage et la méthode Newton-Raphson en terme d'introduction de la plus importante information de la dérivée seconde.

Les premières étapes de cette méthode proviennent donc de la dérivée première et la SD est la méthode la plus couramment utilisée.

Le choix de l'algorithme de minimisation est dicté par un certain nombre de facteurs dont le stockage des données, le temps de calcul, la disponibilité des données analytiques, la puissance de la méthode, etc. Pour toutes ces raisons, dans les premières étapes de minimisation on utilise des méthodes robustes comme la SD avant de faire usage de l'ABNR.

### **III Les logiciels utilisés**

#### **A Le logiciel CHARMM**

Le logiciel utilisé pour les calculs de dynamique moléculaire est la version parallélisée de CHARMM (Chemistry at HARvard Macromolecular Mechanics) version 26 (Brooks 1983), avec le champ de forces PARAM 22 (ou 27) (MacKerell 1992, 1998), qui inclut une bibliothèque de phospholipides (Schlenkrich et al. 1996) et le potentiel des molécules d'eau TIP3P (Jorgensen et al. 1983).

CHARMM est un logiciel permettant de modéliser des systèmes macromoléculaires en utilisant les fonctions d'énergie empiriques définies précédemment. Ce programme permet de lire et construire des modèles de structures, de minimiser leur énergie, d'effectuer des simulations de dynamique moléculaire.

Ce logiciel permet aussi d'analyser les structures générées et de déterminer les propriétés dynamiques et énergétiques des systèmes étudiés.

## **B Les logiciels de visualisation**

### **1 VMD**

Le logiciel VMD ou Visual Molecular Dynamics (Humphrey et al. 1996) est un programme de visualisation graphique compatible avec CHARMM. Il permet une représentation en 3D ainsi qu'une animation des macromolécules biologiques. VMD a été développé par le groupe « Theoretical and Computational Biophysics Group » à l'université de l'Illinois. (<http://www.ks.uiuc.edu/Research/vmd/>).

### **2 MolMol**

Tout comme VMD, MOLMOL pour MOLEcule analysis and MOLEcule display (Koradi et al. 1996) est aussi un logiciel de visualisation graphique, d'analyse et de manipulation des structures 3D des macromolécules biologiques : (<http://hugin.ethz.ch/wuthrich/software/molmol/>)

J'ai utilisé ces deux logiciels au cours de cette thèse pour la visualisation et la création des images issues des fichiers de coordonnées et des trajectoires générées lors des simulations.

## **IV Protocole de recherche des modes d'association**

Le protocole adopté tout au long de cette étude est composé de 5 phases.

- 1) Construction du système à étudier
- 2) Chauffage du système
- 3) Equilibration
- 4) Production
- 5) Analyses des résultats

Cette partie est décrite dans le chapitre III sous forme d'article publié dans *European Biophysics Journal*.

## A Construction du système à étudier

Chaque type de dimère, homodimère pour la GpA et hétérodimère pour les récepteurs ErbB, est construit suivant le même protocole à partir des séquences des domaines TM ci-dessous

	Motif I	Motif II	Numérotation protéine
GpA	I T L I I F <b>G V M A G</b> V I G T I L L I S Y G I		73-95
ErbB1	P S I A T <b>G M V G A</b> L L L L L V V <b>A L G I G</b> L F M		644-667
ErbB2	L T S I V <b>S A V V G</b> I L L V V V L <b>G V V F G</b> I L I		651-675
ErbB3	L T M A L <b>T V I A G</b> L V V I F M M L G G T F L Y W		642-666
ErbB4	P L I A A <b>G V I G G</b> L F I L V I V <b>G L T F A</b> V Y V		651-675

**Figure 2.8** : Séquence du domaine TM des récepteurs ErbB (Coussens et al. 1985; Kraus et al. 1989; Plowman et al. 1993; Ullrich et al. 1984; Yamamoto et al. 1986) et du domaine TM de la GpA (Tomita et al. 1978) montrant les motifs de dimérisation du type GxxxG ou plus largement les sites d'interactions du type SmxxxSm (résidus soulignés) de chaque récepteur.

La **figure 2.9** résume les principales étapes de la procédure.

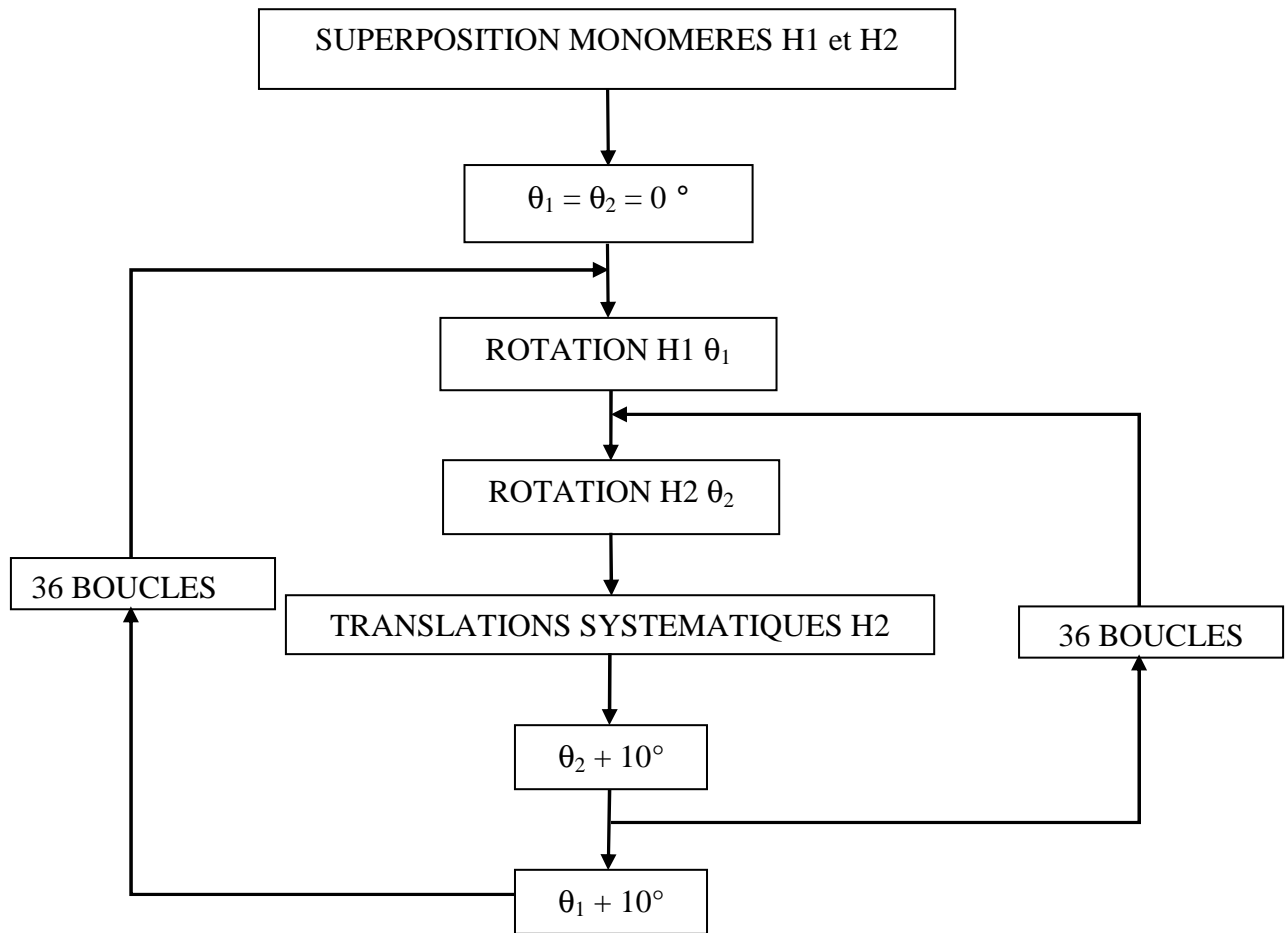
J'ai commencé par construire les monomères sous forme d'hélice  $\alpha$  canonique (3,6 résidus/tour d'hélice, angles  $\varphi$  et  $\psi$  respectivement égaux à  $-57^\circ$  et  $-47^\circ$ ) aussi bien pour la GpA que pour les récepteurs ErbB. A partir des quatre récepteurs transmembranaires ErbB, nous avons six hétérodimères possibles ; à savoir ErbB<sup>1-2, 1-3, 1-4, 2-3, 2-4 et 3-4</sup>.

Aux extrémités N- et C-terminales de chaque monomère, nous avons ajouté un groupe méthyle acétamide ( $\text{CH}_3\text{CO}$ ) et un groupe méthyle amide ( $\text{NHCH}_3$ ) respectivement. Les deux hélices construites sont superposées de telle sorte que l'axe principal de chacune d'elles coïncide avec l'axe Z du référentiel.

Pour obtenir les structures initiales d'un dimère, les deux hélices H1 et H2 subissent une rotation complète autour de leur axe hélicoïdal (d'un angle  $\theta_1$  pour H1,  $\theta_2$  pour H2). Pour chaque valeur de  $\theta_1$  une rotation complète de l'hélice H2 est effectuée pour balayer toutes les faces de la seconde hélice. Le pas de rotation est de  $10^\circ$  et 36 rotations sont effectuées (soit  $36 \times 10^\circ = 360^\circ$ ). 1296 structures initiales sont ainsi générées pour chaque dimère ( $36 \times 36$ ).

Pour chaque couple  $(\theta_1, \theta_2)$ , l'hélice H2 est translatée le long de l'axe X d'une distance de 20 Å (distance axe hélicoïdal – axe hélicoïdal). Cette distance est progressivement réduite par pas de 0.25 Å jusqu'à trouver la plus courte distance inter-hélice qui n'engendre pas de

contacts stériques entre les chaînes latérales situées à l'interface. A cette étape, des courtes minimisations SD et ABNR sont effectuées pour éliminer les contacts stériques éventuels. Cette structure minimisée constitue le point de départ des simulations MD.



**Figure 2.9** Organigramme résumant les principales étapes de la construction des dimères des hélices transmembranaires. A la fin de la construction, chaque configuration du dimère est minimisée et les coordonnées sont sauvegardées.

## B Simulation MD dans le vide

### 1 Chauffe

L'étape dite de montée en température est une période pendant laquelle la température du système est progressivement élevée à 300 K et dure 18 ps. Elle se déroule par palier de 50 K durant 3 ps. Au cours de cette phase, des vitesses initiales sont aléatoirement attribuées à chacun des atomes. Ces vitesses sont tirées à partir d'une distribution de Maxwell centrée sur

la température choisie. A chaque nouveau pallier, de nouvelles vitesses sont affectées à chaque atome du système.

Dans le cas de nos simulations dans le vide, nous avons appliqué des contraintes de position et de liaison hydrogène pour maintenir les deux hélices en structures  $\alpha$ .

## 2 Equilibration - Production

La période dite d'équilibration dépend de la taille du système. Dans notre cas les systèmes dimériques simulés dans le vide ne dépassent pas 1000 atomes. Cette période transitoire, d'une durée de 1 ns, permet une redistribution des vitesses relativement isotrope sur l'ensemble des atomes en utilisant un couplage faible à un bain thermique. Au cours de cette phase, les constantes de force appliquées aux contraintes sont progressivement réduites jusqu'à être annulées pour laisser les systèmes évoluer librement.

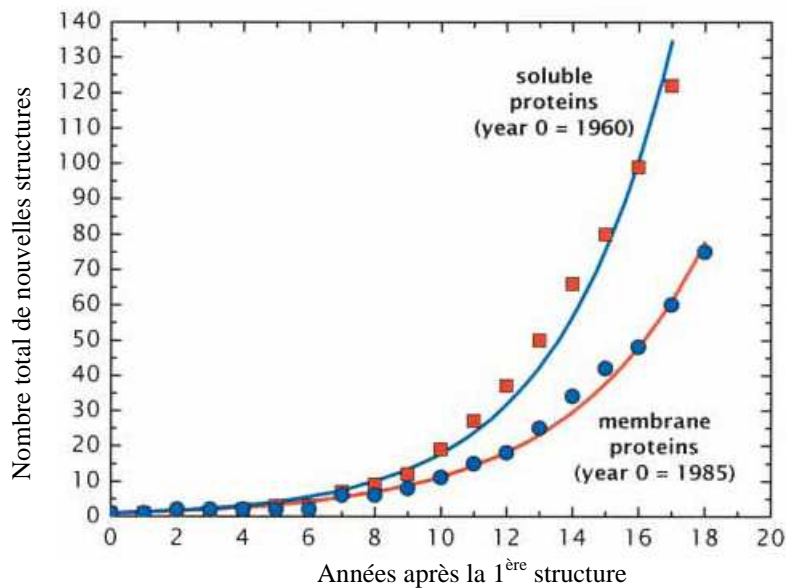
Cette étape est suivie d'une période de production d'une durée de 3 ns. Durant cette étape, aucune contrainte n'est appliquée, les caractéristiques énergétiques du système sont contrôlées et les coordonnées sont stockées pour produire les trajectoires. Les analyses sont effectuées au cours de la dernière ns de production.

## 3 Analyses

Les structures des dimères sont sélectionnées en appliquant des critères géométriques. Ces critères sont basés sur l'assemblage des hélices membranaires observé dans les structures RX des protéines membranaires. Le nombre de protéines membranaires dont la structure a été résolue par RX est faible comparé à celui des protéines solubles (**graphe 2.10**) (White 2004). Plusieurs études ont été menées pour comprendre le mode d'assemblage des hélices TM, leur orientation dans la membrane, leur longueur, leur inclinaison et la recherche de motifs structuraux (Bowie 1997).

Ces études ont permis de définir les critères géométriques utiles à la sélection des structures des hétérodimères ErbB.

Les caractéristiques des hélices TM sont décrites dans les paragraphes suivants.



**Figure 2.10 :** Evolution du nombre de structures des protéines solubles et des protéines membranaires dont la structure est résolue. D'après (White 2004).

a\_ Longueur des hélices

La longueur des hélices TM varie fortement de 14 à 39 résidus, avec une moyenne de 26.4 résidus (Bowie 1997). La taille généralement observée des hélices est de plus de 20 résidus, qui est un nombre optimal pour traverser la membrane (Engelman et al. 1986). Beaucoup d'hélices TM s'étendent en dehors de la membrane et contribuent à la structure des domaines solubles. Ainsi, nous avons pris des segments TM de 23 et 25 résidus respectivement pour les monomères de la GpA et les hélices ErbB. Ces séquences correspondent à la partie hydrophobe des domaines TM de ces récepteurs.

b\_ Arrangement des hélices

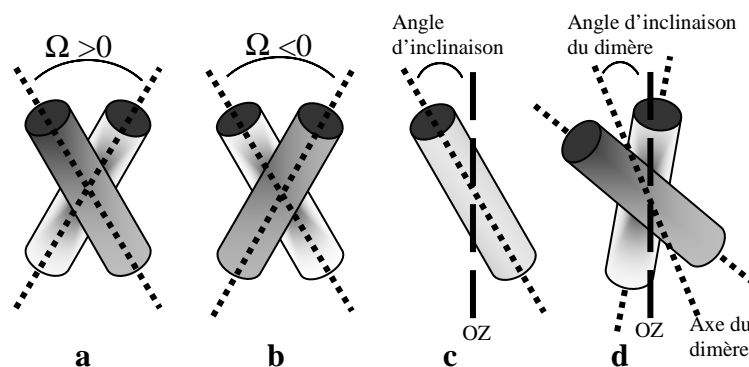
Dans un dimère, les hélices peuvent avoir deux orientations : parallèle ou anti-parallèle. L'orientation du N-terminal vers le C-terminal est la même pour les hélices parallèles et opposée pour les hélices anti-parallèles. Sur les 88 interactions entre hélices TM (Bowie 1997, 1997), 63% des hélices sont dans une orientation anti-parallèle.

Pour éviter d'avoir des hélices trop décalées l'une par rapport à l'autre le long de leur axe, nous avons appliqué un filtre. Ainsi, le déplacement suivant l'axe Z entre les deux monomères

ne doit pas être supérieur à 3.5 Å aussi bien du côté N-terminal que du côté C-terminal. De plus la différence entre les distances croisées ne doit pas elle aussi dépasser 3.5 Å (**voir légende du tableau 2.1**).

#### c\_ Angles d'inclinaison et de croisement

La moyenne des angles d'inclinaison des hélices par rapport à la normale à la bicouche lipidique est de 21° (**figure 2.11**), mais avec une préférence pour les petits angles compris entre 10 et 20°.

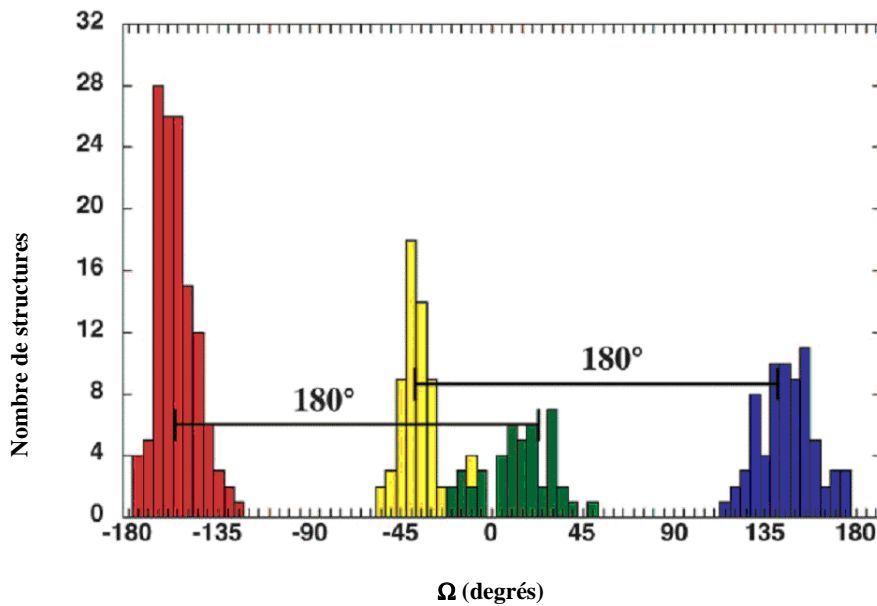


**Figure 2.11 : Orientation des hélices TM.** a) Par définition si l'hélice en premier plan du dimère doit effectuer une rotation dans le sens des aiguilles d'une montre pour s'aligner avec l'hélice en arrière plan, l'angle de croisement  $\Omega$  est positif et caractérise une association gauche. b) Si l'hélice en premier plan doit effectuer une rotation dans le sens inverse des aiguilles d'une montre pour s'aligner avec l'hélice en arrière plan,  $\Omega$  est négatif et caractérise une association droite. Dans cette définition la direction des hélices est ignorée. c) l'angle d'inclinaison est l'angle entre l'axe de l'hélice et l'axe OZ. d) l'angle d'inclinaison du dimère est l'angle formé par l'axe du dimère, (moyenne des deux axes de chacune des hélices), et l'axe Z. D'après (Chothia et al. 1981).

Les angles de croisement entre les hélices TM sont distribués entre  $-56^\circ$  et  $+67^\circ$ . La valeur la plus fréquemment rencontrée est de  $+20^\circ$  pour les protéines membranaires, alors que dans les protéines solubles celle-ci est de  $-35^\circ$ . La différence de  $\Omega$  entre les protéines solubles et les protéines TM vient de la restriction de l'angle d'inclinaison des hélices lorsqu'elles sont insérées dans la membrane.

L'étude de Bowie (Bowie 1997, 1997) a été élargie par une analyse plus récente sur 11 protéines TM comprenant 125 interactions hélice-hélice (Senes et al. 2001) et par celle de (Walters and DeGrado 2006). La **figure 2.12** représente la distribution de l'angle de croisement pour des associations des hélices parallèles et anti-parallèles.

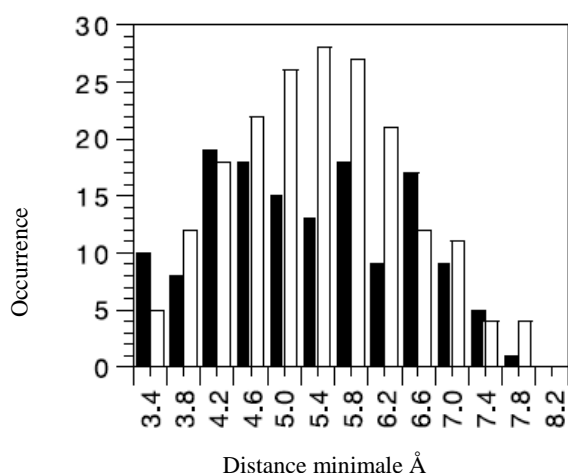




**Figure 2.12 : Distribution de l'angle de croisement ( $\Omega$ ) entre les hélices TM observés dans les protéines membranaires.** Ce graphe met en évidence 4 blocs rouge, jaune, vert et bleu qui représentent respectivement les structures antiparallèles gauches, parallèles droites, antiparallèles droites et parallèles gauches, D'après (Walters and DeGrado 2006).

d\_ Distance inter-hélices

La distance la plus faible mesurée entre les axes des hélices en interaction est qualifiée de distance de plus petite approche (Walther et al. 1996). Dans les membranes, la distance de plus petite approche est respectivement de 9.1 Å et 10.5 Å pour les paires anti-parallèles et les paires parallèles. La distance moyenne est donc de 9.6 Å, très proche de celle mesurée sur 2145 interactions hélice-hélice dans des protéines solubles (Bowie 1997). Les distances inter-hélices moyennes entre les atomes C $\alpha$  des résidus qui se font face sont plus courtes dans les protéines membranaires (5.5 Å) que dans les protéines solubles (6.0 Å) (Gimpelev et al. 2004).



**Figure 2.13 : Distribution des distances inter-hélices minimales entre les squelettes peptidiques des protéines membranaires (barres noires) et solubles (barres blanches).**

La distance minimale moyenne entre les axes hélicoïdaux est de 5.10 Å et 5.22 Å respectivement pour les protéines membranaires et solubles. D'après (Eilers et al. 2002).

#### e\_ Choix des critères géométriques

Les différents critères géométriques utilisés pour la sélection des structures des hétérodimères sont donnés dans le **tableau 2.1**.

	$L_H$ (Å)		$\Delta dc$ (Å)	$\Delta Z_{N\text{-ter}}$	$\Delta Z_{C\text{-ter}}$ (Å)	$\Omega$ (°)	
Dimer	min	max				min	max
HER <sub>TM</sub>	32.0	38.0	$\leq 3.5$		$\leq 3.5$	-65°	+45°
GpA <sub>TM</sub>	29.0	35.0	$\leq 3.5$		$\leq 3.5$	-65°	+45°

**Tableau 2.1: Paramètres définissant la qualité de la structure du dimère.**

Les valeurs min et max de la longueur de l'hélice ( $L_H$ ) sont définies à partir de la longueur d'une hélice  $\alpha$  canonique ( $L_{H\alpha} = 35.25$  Å pour les hélices HER<sub>TM</sub> et 32.5 Å pour l'hélice GpA<sub>TM</sub>).  $\Omega$  est l'angle de croisement entre les deux axes principaux des deux monomères. Les valeurs min et max proviennent de l'étude de Bowie (Bowie 1997) et des analyses statistiques des protéines membranaires.  $\Delta dc$  correspond à la valeur absolue de la différence des distances croisées entre les atomes  $C_\alpha$  terminaux :  $C_\alpha^{H1}N\text{ter}-C_\alpha^{H2}C\text{ter}$  et  $C_\alpha^{H2}N\text{ter}-C_\alpha^{H1}C\text{ter}$ .  $\Delta Z(N\text{ter}) = |Z(C_\alpha^{H1} N\text{ter}) - Z(C_\alpha^{H2} N\text{ter})|$  et  $\Delta Z(C\text{ter}) = |Z(C_\alpha^{H1} C\text{ter}) - Z(C_\alpha^{H2} C\text{ter})|$  correspondent respectivement à la valeur absolue de la différence des positions Z des atomes  $C_\alpha$  des résidus terminaux des hélices H1 et H2.

Il n'y avait pas de limitation de l'angle de croisement contrairement aux méthodes courantes (Adamian and Liang 2006; Adams et al. 1995; Akula and Pattabiraman 2005; Beevers and Kukol 2006; Bu et al. 2007; Kim et al. 2003). Avec ce protocole, nous avons généré un

nombre raisonnable de structure (1296 structures initiales) comparativement à Kukol et Beevers qui ont généré 72 structures de départ puis les ont copiés 4 fois donnant ainsi 288 structures. Toutefois ces auteurs ont effectué des simulations de dynamique moléculaire dans un modèle de membrane. D'autres auteurs ont généré un très grand nombre de structures dimériques de la GpA (97200) et la sélection est basée sur des critères énergétiques (Adamian and Liang 2006; Adams et al. 1995; Akula and Pattabiraman 2005; Beevers and Kukol 2006; Bu et al. 2007; Kim et al. 2003) sans faire de simulation MD.

La méthode que j'ai développée n'utilise pas de contrainte de symétrie, ni d'angle de croisement des hélices. La sélection des meilleures structures est basée uniquement sur des critères géométriques qui rendent compte implicitement de l'existence de la membrane. Elle est indépendante de la séquence des hélices TM. Cette méthode présente l'avantage de pouvoir s'appliquer à des hétérodimères. Elle pourrait s'appliquer à l'étude des oligomères. Le nombre de configurations initiales sera beaucoup plus grand et nécessitera un temps de calcul plus long.

## **BIBLIOGRAPHIE CHAPITRES I et II**

- Adamian L, Liang J (2006) Prediction of transmembrane helix orientation in polytopic membrane proteins. *BMC Struct Biol* 6:13
- Adams PD, Arkin IT, Engelman DM, Brunger AT (1995) Computational searching and mutagenesis suggest a structure for the pentameric transmembrane domain of phospholamban. *Nat Struct Biol* 2:154-62
- Aifa S, Aydin J, Nordvall G, Lundstrom I, Svensson SP, Hermanson O (2005) A basic peptide within the juxtamembrane region is required for EGF receptor dimerization. *Exp Cell Res* 302:108-14
- Aifa S, Miled N, Frikha F, Aniba MR, Svensson SP, Rebai A (2006) Electrostatic interactions of peptides flanking the tyrosine kinase domain in the epidermal growth factor receptor provides a model for intracellular dimerization and autophosphorylation. *Proteins* 62:1036-43
- Akula N, Pattabiraman N (2005) A systematic search method for the identification of tightly packed transmembrane parallel alpha-helices. *J Biomol Struct Dyn* 22:625-34
- Aller P, Garnier N, Genest M (2006) Transmembrane Helix Packing of ErbB/Neu Receptor in Membrane Environment: A Molecular Dynamics Study. *J Biomol Struct Dyn* 24:209-28
- Aller P, Voiry L, Garnier N, Genest M (2005) Molecular dynamics (MD) investigations of preformed structures of the transmembrane domain of the oncogenic Neu receptor dimer in a DMPC bilayer. *Biopolymers* 77:184-197
- Alroy I, Yarden Y (1997) The ErbB signaling network in embryogenesis and oncogenesis: signal diversification through combinatorial ligand-receptor interactions. *FEBS Lett* 410:83-6
- Bagossi P, Horvath G, Vereb G, Szollosi J, Tozser J (2005) Molecular modeling of nearly full-length ErbB2 receptor. *Biophys. J.* 88:1354-63
- Barbacci EG, Guarino BC, Stroh JG, Singleton DH, Rosnack KJ, Moyer JD, Andrews GC (1995) The structural basis for the specificity of epidermal growth factor and heregulin binding. *J Biol Chem* 270:9585-9
- Bargmann CI, Hung MC, Weinberg RA (1986) Multiple independent activations of the neu oncogene by a point mutation altering the transmembrane domain of p185. *Cell* 45:649-657

- Bargmann CI, Weinberg RA (1988) Oncogenic activation of the neu-encoded receptor protein by point mutation and deletion. *EMBO J* 7:2043-2052
- Bazley LA, Gullick WJ (2005) The epidermal growth factor receptor family. *Endocr Relat Cancer* 12 Suppl 1:S17-27
- Beerli RR, Graus-Porta D, Woods-Cook K, Chen X, Yarden Y, Hynes NE (1995) Neu differentiation factor activation of ErbB-3 and ErbB-4 is cell specific and displays a differential requirement for ErbB-2. *Mol Cell Biol* 15:6496-505
- Beevers AJ, Kukol A (2006) Systematic molecular dynamics searching in a lipid bilayer: Application to the glycoporphin A and oncogenic ErbB-2 transmembrane domains. *J Mol Graph Model* 25:226-233
- Birrane G, Chung J, Ladias JA (2003) Novel mode of ligand recognition by the Erbin PDZ domain. *J Biol Chem* 278:1399-402
- Blair JA, Rauh D, Kung C, Yun CH, Fan QW, Rode H, Zhang C, Eck MJ, Weiss WA, Shokat KM (2007) Structure-guided development of affinity probes for tyrosine kinases using chemical genetics. *Nat Chem Biol* 3:229-38
- Blume-Jensen P, Hunter T (2001) Oncogenic kinase signalling. *Nature* 411:355-65
- Bouyain S, Longo PA, Li S, Ferguson KM, Leahy DJ (2005) The extracellular region of ErbB4 adopts a tethered conformation in the absence of ligand. *Proc Natl Acad Sci U S A* 102:15024-9
- Bowie JU (1997) Helix packing angle preferences. *Nat. Struc. Biol.* 4:915-917
- Bowie JU (1997) Helix packing in membrane proteins. *J. Mol. Biol.* 272:780-789
- Brooks BR, Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., Karplus, M. (1983) CHARMM: A program for macromolecular energy minimization, and dynamics calculations. *J. Comp. Chem.* 4:187-217
- Bu L, Im W, Brooks CL, 3rd (2007) Membrane assembly of simple helix homo-oligomers studied via molecular dynamics simulations. *Biophys J* 92:854-63
- Carraway KL, Cantley LC (1994) A neu acquaintance for erbB3 and erbB4: a role for receptor heterodimerization in growth signaling. *Cell* 78:5-8
- Carraway KL, Weber JL, Unger MJ, Ledesma J, Yu N, Gassmann M, Lai C (1997) Neuregulin-2, a new ligand of ErbB3/ErbB4-receptor tyrosine kinases. *Nature* 387:512-6
- Chang H, Riese DJ, Gilbert W, Stern DF, McMahan UJ (1997) Ligands for ErbB-family receptors encoded by a neuregulin-like gene. *Nature* 387:509-12

- Chantry A (1995) The kinase domain and membrane localization determine intracellular interactions between epidermal growth factor receptors. *J Biol Chem* 270:3068-73
- Cheatham B, Shoelson SE, Yamada K, Goncalves E, Kahn CR (1993) Substitution of the erbB-2 oncoprotein transmembrane domain activates the insulin receptor and modulates the action of insulin and insulin- receptor substrate 1 [published erratum appears in *Proc. Natl. Acad. Sci. U S A* 1993 Oct 15;90(20):9739]. *Proc. Natl. Acad. Sci. U S A* 90:7336-40
- Cho HS, Leahy DJ (2002) Structure of the extracellular region of HER3 reveals an interdomain tether. *Science* 297:1330-3
- Cho HS, Mason K, Ramyar KX, Stanley AM, Gabelli SB, Denney DW, Jr., Leahy DJ (2003) Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. *Nature* 421:756-60
- Chothia C, Levitt M, Richardson D (1981) Helix to helix packing in proteins. *J. Mol. Biol.* 145:215-250
- Citri A, Skaria KB, Yarden Y (2003) The deaf and the dumb: the biology of ErbB-2 and ErbB-3. *Exp Cell Res* 284:54-65
- Citri A, Yarden Y (2006) EGF-ERBB signalling: towards the systems level. *Nat Rev Mol Cell Biol* 7:505-16
- Cohen BD, Kiener PA, Green JM, Foy L, Fell HP, Zhang K (1996) The relationship between human epidermal growth-like factor receptor expression and cellular transformation in NIH3T3 cells. *J Biol Chem* 271:30897-903
- Coussens L, Yang-Feng TL, Liao YC, Chen E, Gray A, McGrath J, Seeburg PH, Libermann TA, Schlessinger J, Francke U, et al. (1985) Tyrosine kinase receptor with extensive homology to EGF receptor shares chromosomal location with neu oncogene. *Science* 230:1132-9
- Dawson JP, Berger MB, Lin CC, Schlessinger J, Lemmon MA, Ferguson KM (2005) Epidermal growth factor receptor dimerization and activation require ligand-induced conformational changes in the dimer interface. *Mol Cell Biol* 25:7734-42
- Deber CM, Li SC (1995) Peptides in membranes: helicity and hydrophobicity. *Biopolymers* 37:295-318
- Deisenhofer J, Epp O, Sinning I, Michel H (1995) Crystallographic refinement at 2.3 Å resolution and refined model of the photosynthetic reaction centre from *Rhodospseudomonas viridis*. *J. Mol. Biol.* 246:429-57

- Dougall WC, Qian X, Peterson NC, Miller MJ, Samanta A, Greene MI (1994) The neurooncogene: signal transduction pathways, transformation mechanisms and evolving therapies. *Oncogene* 9:2109-2123
- Dumont JE, Pecasse F, Maenhaut C (2001) Crosstalk and specificity in signalling. Are we crosstalking ourselves into general confusion? *Cell Signal* 13:457-63
- Duneau JP, Crouzy S, Chapron Y, Genest M (1999) Dynamics of the transmembrane domain of the ErbB-2 receptor. *Theor. Chem. Acc.* 101:87-91
- Duneau JP, Crouzy S, Garnier N, Chapron Y, Genest M (1999) Molecular dynamics simulations of the ErbB-2 transmembrane domain within an explicit membrane environment: comparison with vacuum simulations. *Biophys. Chem.* 76:35-53
- Duneau JP, Garnier N, Genest M (1997) Insight into signal transduction: structural alterations in transmembrane helices probed by multi-1 ns molecular dynamics simulations. *J. Biomol. Struct. Dyn.* 15:555-72
- Duneau JP, Genest D, Genest M (1996) Detailed description of an alpha helix-->pi bulge transition detected by molecular dynamics simulations of the p185c-erbB2 V659G transmembrane domain. *J. Biomol. Struct. Dyn.* 13:753-69
- Duneau JP, Vegh AP, Sturgis JN (2007) A dimerization hierarchy in the transmembrane domains of the HER receptor family. *Biochemistry* 46:2010-9
- Eilers M, Patel AB, Liu W, Smith SO (2002) Comparison of helix interactions in membrane and soluble alpha-bundle proteins. *Biophys J* 82:2720-36
- Engelman DM, Steitz TA, Goldman A (1986) Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. *Annu Rev Biophys Biophys Chem* 15:321-53
- Felsenstein J (1989) Phylip - Phylogeny Inference Package (Version 3.2). *Cladistics* 5:164-166
- Ferguson KM, Berger MB, Mendrola JM, Cho HS, Leahy DJ, Lemmon MA (2003) EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. *Mol. Cell* 11:507-17
- Franklin MC, Carey KD, Vajdos FF, Leahy DJ, de Vos AM, Sliwkowski MX (2004) Insights into ErbB signaling from the structure of the ErbB2-pertuzumab complex. *Cancer Cell* 5:317-28
- Gamett DC, Pearson G, Cerione RA, Friedberg I (1997) Secondary dimerization between members of the epidermal growth factor receptor family. *J Biol Chem* 272:12052-6

- Garnier N, Crouzy S, Genest M (2003) Molecular Dynamics Simulations of the Transmembrane Domain of the Oncogenic ErbB2 Receptor Dimer in a DMPC Bilayer. *J. Biomol. Struct. Dyn.* 21:179-200
- Garnier N, Genest D, Duneau JP, Genest M (1997) Molecular modeling of c-erbB2 receptor dimerization: coiled-coil structure of wild and oncogenic transmembrane domains. Stabilization by interhelical hydrogen bonds in the oncogenic form. *Biopolymers* 42:157-168
- Garnier N, Genest D, Hebert E, Genest M (1994) Influence of a mutation in the transmembrane domain of the p185c-erbB2 oncogene-encoded protein studied by molecular dynamics simulations. *J. Biomol. Struct. Dyn.* 11:983-1002
- Garrett TP, McKern NM, Lou M, Elleman TC, Adams TE, Lovrecz GO, Kofler M, Jorissen RN, Nice EC, Burgess AW, Ward CW (2003) The crystal structure of a truncated ErbB2 ectodomain reveals an active conformation, poised to interact with other ErbB receptors. *Mol. Cell* 11:495-505
- Garrett TP, McKern NM, Lou M, Elleman TC, Adams TE, Lovrecz GO, Zhu HJ, Walker F, Frenkel MJ, Hoyne PA, Jorissen RN, Nice EC, Burgess AW, Ward CW (2002) Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor alpha. *Cell* 110:763-73
- Gassmann M, Casagrande F, Orioli D, Simon H, Lai C, Klein R, Lemke G (1995) Aberrant neural and cardiac development in mice lacking the ErbB4 neuregulin receptor. *Nature* 378:390-4
- Gerber D, Sal-Man N, Shai Y (2004) Two motifs within a transmembrane domain, one for homodimerization and the other for heterodimerization. *J Biol Chem* 279:21177-82
- Gimpelev M, Forrest LR, Murray D, Honig B (2004) Helical packing patterns in membrane and soluble proteins. *Biophys J* 87:4075-86
- Goetz M, Carlotti C, Bontems F, Dufourc E (2001) Evidence for an  $\alpha$  helix- $\pi$  bulge helicity modulation for the neu/erbB-2 membrane spanning segment. A  $^1\text{H}$  NMR and Circular dichroism study. *Biochemistry* 40:6534-6540
- Goncalves E, Yamada K, Thatte HS, Backer JM, Golan DE, Kahn CR, Shoelson SE (1993) Optimizing transmembrane domain helicity accelerates insulin receptor internalization and lateral mobility. *Proc Natl Acad Sci U S A* 90:5762-6
- Graus-Porta D, Beerli RR, Daly JM, Hynes NE (1997) ErbB-2, the preferred heterodimerization partner of all ErbB receptors, is a mediator of lateral signaling. *Embo J* 16:1647-55



- Graus-Porta D, Beerli RR, Hynes NE (1995) Single-chain antibody-mediated intracellular retention of ErbB-2 impairs Neu differentiation factor and epidermal growth factor signaling. *Mol Cell Biol* 15:1182-91
- Groenen LC, Walker F, Burgess AW, Treutlein HR (1997) A model for the activation of the epidermal growth factor receptor kinase involvement of an asymmetric dimer? *Biochemistry* 36:3826-36
- Gullick WJ (2001) The Type 1 growth factor receptors and their ligands considered as a complex system. *Endocr Relat Cancer* 8:75-82
- Harari D, Tzahar E, Romano J, Shelly M, Pierce JH, Andrews GC, Yarden Y (1999) Neuregulin-4: a novel growth factor that acts through the ErbB-4 receptor tyrosine kinase. *Oncogene* 18:2681-9
- Harari D, Yarden Y (2000) Molecular mechanisms underlying ErbB2/HER2 action in breast cancer. *Oncogene* 19:6102-14
- Heldin CH (1995) Dimerization of cell surface receptors in signal transduction. *Cell* 80:213-223
- Hockney RW, Eastwood, J. W. (1981) *Computer simulation using particles*. McGraw-Hill Book Company
- Holbro T, Hynes NE (2004) ErbB receptors: directing key signaling networks throughout life. *Annu Rev Pharmacol Toxicol* 44:195-217
- Houliston RS, Hodges RS, Sharom FJ, Davis JH (2003) Comparison of proto-oncogenic and mutant forms of the transmembrane region of the Neu receptor in TFE. *FEBS Lett* 535:39-43
- Houliston RS, Hodges RS, Sharom FJ, Davis JH (2004) Characterization of the proto-oncogenic and mutant forms of the transmembrane region of Neu in micelles. *J Biol Chem* 279:24073-80
- Hubbard SR (1997) Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. *Embo J* 16:5572-81
- Hubbard SR (2002) Protein tyrosine kinases: autoregulation and small-molecule inhibition. *Current Opinion in Structural Biology* 12:735-741
- Hubbard SR, Miller WT (2007) Receptor tyrosine kinases: mechanisms of activation and signaling. *Curr Opin Cell Biol* 19:117-23
- Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. *J Mol Graph* 14:33-8, 27-8

- Hunter T (1995) Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. *Cell* 80:225-36
- Hunter T (1998) The Croonian Lecture 1997. The phosphorylation of proteins on tyrosine: its role in cell growth and disease. *Philos Trans R Soc Lond B Biol Sci* 353:583-605
- Hunter T, Sefton BM (1980) Transforming gene product of Rous sarcoma virus phosphorylates tyrosine. *Proc Natl Acad Sci U S A* 77:1311-5
- Jiang G, Hunter T (1999) Receptor signaling: when dimerization is not enough. *Curr Biol* 9:R568-71
- Jorgensen WL, Impey RW, Chandrasekhar J, Madura JD, Klein ML (1983) Comparison of simple potentials functions for simulating liquid water. *J. Chem. Phys.* 79:926-935
- Jorissen RN, Walker F, Pouliot N, Garrett TP, Ward CW, Burgess AW (2003) Epidermal growth factor receptor: mechanisms of activation and signalling. *Exp Cell Res* 284:31-53
- Karunakaran D, Tzahar E, Beerli RR, Chen X, Graus-Porta D, Ratzkin BJ, Seger R, Hynes NE, Yarden Y (1996) ErbB-2 is a common auxiliary subunit of NDF and EGF receptors: implications for breast cancer. *Embo J* 15:254-64
- Kim S, Chamberlain AK, Bowie JU (2003) A simple method for modeling transmembrane helix oligomers. *J. Mol. Biol.* 329:831-840
- Klapper LN, Glathe S, Vaisman N, Hynes NE, Andrews GC, Sela M, Yarden Y (1999) The ErbB-2/HER2 oncoprotein of human carcinomas may function solely as a shared coreceptor for multiple stroma-derived growth factors [In Process Citation]. *Proc Natl Acad Sci U S A* 96:4995-5000
- Kokai Y, Myers JN, Wada T, Brown VI, LeVea CM, Davis JG, Dobashi K, Greene MI (1989) Synergistic interaction of p185c-neu and the EGF receptor leads to transformation of rodent fibroblasts. *Cell* 58:287-92
- Koradi R, Billeter M, Wuthrich K (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph* 14:51-5, 29-32
- Kraus MH, Issing W, Miki T, Popescu NC, Aaronson SA (1989) Isolation and characterization of ERBB3, a third member of the ERBB/epidermal growth factor receptor family: evidence for overexpression in a subset of human mammary tumors. *Proc Natl Acad Sci U S A* 86:9193-7
- Krebs EG, Fischer EH (1956) The phosphorylase b to a converting enzyme of rabbit skeletal muscle. *Biochim Biophys Acta* 20:150-7

- Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform* 5:150-63
- Lee KF, Simon H, Chen H, Bates B, Hung MC, Hauser C (1995) Requirement for neuregulin receptor erbB2 in neural and cardiac development. *Nature* 378:394-8
- Lemmon MA, Schlessinger J (1994) Regulation of signal transduction and signal diversity by receptor oligomerization. *Trends Biochem Sci* 19:459-63
- Lenferink AE, Pinkas-Kramarski R, van de Poll ML, van Vugt MJ, Klapper LN, Tzahar E, Waterman H, Sela M, van Zoelen EJ, Yarden Y (1998) Differential endocytic routing of homo- and hetero-dimeric ErbB tyrosine kinases confers signaling superiority to receptor heterodimers. *Embo J* 17:3385-97
- Li S, Schmitz KR, Jeffrey PD, Wiltzius JJ, Kussie P, Ferguson KM (2005) Structural basis for inhibition of the epidermal growth factor receptor by cetuximab. *Cancer Cell* 7:301-11
- Li Y, Mangasarian K, Mansukhani A, Basilico C (1997) Activation of FGF receptors by mutations in the transmembrane domain. *Oncogene* 14:1397-406
- Li Z, Mei Y, Liu X, Zhou M (2007) Neuregulin-1 only induces trans-phosphorylation between ErbB receptor heterodimer partners. *Cell Signal* 19:466-71
- Linggi B, Carpenter G (2006) ErbB-4 s80 intracellular domain abrogates ETO2-dependent transcriptional repression. *J Biol Chem* 281:25373-80
- Lofts FJ, Hurst HC, Sternberg MJ, Gullick WJ (1993) Specific short transmembrane sequences can inhibit transformation by the mutant neu growth factor receptor in vitro and in vivo. *Oncogene* 8:2813-2820
- Longo N, Shuster RC, Griffin LD, Langley SD, Elsas LJ (1992) Activation of insulin receptor signaling by a single amino acid substitution in the transmembrane domain. *J. Biol. Chem.* 267:12416-12419
- MacKenzie KR, Prestegard JH, Engelman DM (1997) A transmembrane helix dimer: structure and implications. *Science* 276:131-3
- MacKerell JAD, Bashford, D., Bellot, M., Dunbrack, R.L., Evansek, J.D., Field, M.J., Fisher, S., Gao, J., Guo, H., Ha, S., Joseph McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F.T.K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D.T., Prodhom, B., Reiher, W.E., Roux, B., Schlenkrich, B., Smith, J., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Karplus, M. (1992) Self-consistent parametrization of biomolecules for molecular modelling and condensed phase simulations. *Biophys. J.* 61:A143
- MacKerell JAD, Bashford, D., Bellot, M., Dunbrack, R.L., Evansek, J.D., Field, M.J., Fisher, S., Gao, J., Guo, H., Ha, S., Joseph McCarthy, D., Kuchnir, L., Kuczera, K., Lau,

- F.T.K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D.T., Prodohm, B., Reiher, W.E., Roux, B., Schlenkrich, B., Smith, J., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B.* 102:3585-3616
- Marmor MD, Skaria KB, Yarden Y (2004) Signal transduction and oncogenesis by ErbB/HER receptors. *Int J Radiat Oncol Biol Phys* 58:903-13
- Mattoon D, Klein P, Lemmon MA, Lax I, Schlessinger J (2004) The tethered configuration of the EGF receptor extracellular domain exerts only a limited control of receptor function. *Proc Natl Acad Sci U S A* 101:923-8
- McCammion JA, Gelin BR, Karplus M (1977) Dynamics of folded proteins. *Nature* 267:585-90
- Mendrola JM, Berger MB, King MC, Lemmon MA (2002) The single transmembrane domains of ErbB receptors self-associate in cell membranes. *J. Biol. Chem.* 277:4704-12
- Meyers GA, Orlow SJ, Munro IR, Przylepa KA, Jabs EW (1995) Fibroblast growth factor receptor 3 (FGFR3) transmembrane mutation in Crouzon syndrome with acanthosis nigricans. *Nat Genet* 11:462-4
- Miettinen PJ, Berger JE, Meneses J, Phung Y, Pedersen RA, Werb Z, Derynck R (1995) Epithelial immaturity and multiorgan failure in mice lacking epidermal growth factor receptor. *Nature* 376:337-41
- Miloso M, Mazzotti M, Vass WC, Beguinot L (1995) SHC and GRB-2 are constitutively by an epidermal growth factor receptor with a point mutation in the transmembrane domain. *J. Biol. Chem.* 270:19557-19562
- Moe GR, Bollag GE, Koshland DE (1989) Transmembrane signaling by a chimera of the *Escherichia coli* aspartate receptor and the human insulin receptor. *Proc. Natl. Acad. Sci. U S A* 86:5683-7
- Mohammadi M, Dikic I, Sorokin A, Burgess WH, Jaye M, Schlessinger J (1996) Identification of six novel autophosphorylation sites on fibroblast growth factor receptor 1 and elucidation of their importance in receptor activation and signal transduction. *Mol Cell Biol* 16:977-89
- Murali R, Brennan PJ, Kieber-Emmons T, Greene MI (1996) Structural analysis of p185c-neu and epidermal growth factor receptor tyrosine kinases: oligomerization of kinase domains. *Proc. Natl. Acad. Sci. U.S.A.* 93:6252-6257

- Muthuswamy SK, Gilman M, Brugge JS (1999) Controlled dimerization of ErbB receptors provides evidence for differential signaling by homo- and heterodimers. *Mol Cell Biol* 19:6845-57
- Noble ME, Endicott JA, Johnson LN (2004) Protein kinase inhibitors: insights into drug design from structure. *Science* 303:1800-5
- Normanno N, Bianco C, De Luca A, Maiello MR, Salomon DS (2003) Target-based agents against ErbB receptors and their ligands: a novel approach to cancer treatment. *Endocr Relat Cancer* 10:1-21
- Normanno N, De Luca A, Bianco C, Strizzi L, Mancino M, Maiello MR, Carotenuto A, De Feo G, Caponigro F, Salomon DS (2006) Epidermal growth factor receptor (EGFR) signaling in cancer. *Gene* 366:2-16
- Normanno N, Maiello MR, De Luca A (2003) Epidermal growth factor receptor tyrosine kinase inhibitors (EGFR-TKIs): simple drugs with a complex mechanism of action? *J Cell Physiol* 194:13-9
- Ogiso H, Ishitani R, Nureki O, Fukai S, Yamanaka M, Kim JH, Saito K, Sakamoto A, Inoue M, Shirouzu M, Yokoyama S (2002) Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. *Cell* 110:775-87
- Olayioye MA, Neve RM, Lane HA, Hynes NE (2000) The ErbB signaling network: receptor heterodimerization in development and cancer. *EMBO J* 19:3159-3167
- Parsons JT, Parsons SJ (1997) Src family protein tyrosine kinases: cooperating with growth factor and adhesion signaling pathways. *Curr Opin Cell Biol* 9:187-92
- Pawson T (1995) Protein modules and signalling networks. *Nature* 373:573-80
- Pebay-Peyroula E, Rummel G, Rosenbusch JP, Landau EM (1997) X-ray structure of bacteriorhodopsin at 2.5 angstroms from microcrystals grown in lipidic cubic phases. *Science* 277:1676-81
- Peles E, Ben-Levy R, Tzahar E, Liu N, Wen D, Yarden Y (1993) Cell-type specific interaction of Neu differentiation factor (NDF/herregulin) with Neu/HER-2 suggests complex ligand-receptor relationships. *EMBO J*. 12:961-971
- Peles E, Yarden Y (1993) Neu and its ligands: from an oncogene to neural factors. *Bioessays* 15:815-824
- Petti LM, Irusta PM, DiMaio D (1998) Oncogenic activation of the PDGF beta receptor by the transmembrane domain of p185neu\*. *Oncogene* 16:843-51
- Pinkas-Kramarski R, Soussan L, Waterman H, Levkowitz G, Alroy I, Klapper L, Lavi S, Seger R, Ratzkin BJ, Sela M, Yarden Y (1996) Diversification of Neu differentiation

- factor and epidermal growth factor signaling by combinatorial receptor interactions. *EMBO J* 15:2452-67
- Plowman GD, Culouscou JM, Whitney GS, Green JM, Carlton GW, Foy L, Neubauer MG, Shoyab M (1993) Ligand-specific activation of HER4/p180erbB4, a fourth member of the epidermal growth factor receptor family. *Proc. Natl. Acad. Sci. U S A* 90:1746-50
- Plowman GD, Whitney GS, Neubauer MG, Green JM, McDonald VL, Todaro GJ, Shoyab M (1990) Molecular cloning and expression of an additional epidermal growth factor receptor-related gene. *Proc Natl Acad Sci U S A* 87:4905-9
- Popot JL, Engelman DM (1990) Membrane protein folding and oligomerization: the two-stage model. *Biochemistry* 29:4031-4037
- Qian X, LeVea CM, Freeman JK, Dougall WC, Greene MI (1994) Heterodimerization of epidermal growth factor receptor and wild-type or kinase-deficient Neu: a mechanism of interreceptor kinase activation and transphosphorylation. *Proc. Natl. Acad. Sci. U.S.A.* 91:1500-1504
- Rahman A (1964) Correlations in the Motion of Atoms in Liquid Argon. *Phys. Rev.* 136:A405 - A411
- Rahman A, Stillinger, F.H. (1971) Molecular Dynamics Study of Liquid Water. *J. Chem. Phys.* 55:3336-3359
- Riethmacher D, Sonnenberg-Riethmacher E, Brinkmann V, Yamaai T, Lewin GR, Birchmeier C (1997) Severe neuropathies in mice with targeted mutations in the ErbB3 receptor. *Nature* 389:725-30
- Rousseau F, Bonaventure J, Legeai-Mallet L, Pelet A, Rozet JM, Maroteaux P, Le Merrer, Munnich A (1994) Mutations in the gene encoding fibroblast growth factor receptor-3 in achondroplasia. *Nature* 371:252-4
- Russ WP, Engelman DM (1999) TOXCAT: a measure of transmembrane helix association in a biological membrane. *Proc. Natl. Acad. Sci. U S A* 96:863-868
- Sajot N, Garnier N, Genest M (1999) Dimer models for ErbB-2/neu transmembrane domains from molecular dynamics simulations. *Theor. Chem. Acc.* 101:67-72
- Sajot N, Genest M (2000) Structure prediction of the dimeric neu/ErbB-2 transmembrane domain from multi-nanosecond molecular dynamics simulations. *Eur. Biophys. J.* 28:648-662
- Sajot N, Genest M (2001) Dimer interface of transmembrane domains for neu/erbB-2 receptor dimerization and transforming activation: a model revealed by molecular dynamics simulations. *J. Biomol. Struct. Dyn.* 19:15-31

- Samna Soumana O, Aller P, Garnier N, Genest M (2005) Transmembrane peptides from tyrosine kinase receptor. Mutation-related behavior in a lipid bilayer investigated by molecular dynamics simulations. *J. Biomol. Struct. Dyn.* 23:91-100
- Samna Soumana O, Garnier N, Genest M (2007) Molecular dynamics simulation approach for the prediction of transmembrane helix-helix heterodimers assembly. *Eur Biophys J*
- Schlenkrich M, Brickmann J, MacKerell ADJ, Karplus M (1996) An empirical potential energy function for phospholipids: criteria for parameter optimization and applications. In: Merz KJ, Roux B (eds) *Biological Membranes, a molecular perspective from computation and experiment*. Birkhäuser, Boston, pp 31-81
- Schlessinger J (1988) Signal transduction by allosteric receptor oligomerization. *Trends Biochem. Sci.* 13:443-7
- Schlessinger J (2000) Cell signaling by receptor tyrosine kinases. *Cell* 103:211-225
- Segatto O, King CR, Pierce JH, Di Fiore PP, Aaronson SA (1988) Different structural alterations upregulate in vitro tyrosine kinase activity and transforming potency of the erbB-2 gene. *Mol Cell Biol* 8:5570-4
- Senes A, Engel DE, DeGrado WF (2004) Folding of helical membrane proteins: the role of polar, GxxxG-like and proline motifs. *Curr Opin Struct Biol* 14:465-79
- Senes A, Gerstein M, Engelman DM (2000) Statistical analysis of amino acid patterns in transmembrane helices: the GXXXG motif occurs frequently and in association with beta-branched residues at neighboring positions. *J. Mol. Biol.* 296:921-936
- Senes A, Ubarretxena-Belandia I, Engelman DM (2001) The Calpha ---H.O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions. *Proc Natl Acad Sci U S A* 98:9056-61
- Shankaran H, Wiley HS, Resat H (2006) Modeling the effects of HER/ErbB1-3 coexpression on receptor dimerization and biological response. *Biophys J* 90:3993-4009
- Shiang R, Thompson LM, Zhu YZ, Church DM, Fielder TJ, Bocian M, Winokur ST, Wasmuth JJ (1994) Mutations in the transmembrane domain of FGFR3 cause the most common genetic form of dwarfism, achondroplasia. *Cell* 78:335-342
- Sibilia M, Wagner EF (1995) Strain-dependent epithelial defects in mice lacking the EGF receptor. *Science* 269:234-8
- Smith JM, Mayer BJ (2002) Abl: mechanisms of regulation and activation. *Front Biosci* 7:d31-42
- Smith SO, Aschheim K, Groesbeck M (1996) Magic angle spinning NMR spectroscopy of membrane proteins. *Quater Rev Biophys* 3:252-8

- Smith SO, Smith CS, Bormann BJ (1996) Strong hydrogen bonding interactions involving a buried glutamic acid in the transmembrane sequence of the neu/erbB-2 receptor. *Nat. Struct. Biol.* 3:252-258
- Spivak-Kroizman T, Rotin D, Pinchasi D, Ullrich A, Schlessinger J, Lax I (1992) Heterodimerization of c-erbB2 with different epidermal growth factor receptor mutants elicits stimulatory or inhibitory responses. *J Biol Chem* 267:8056-63
- Srinivasan R, Poulosom R, Hurst HC, Gullick WJ (1998) Expression of the c-erbB-4/HER4 protein and mRNA in normal human fetal and adult tissues and in a survey of nine solid tumour types. *J Pathol* 185:236-45
- Stamos J, Sliwkowski MX, Eigenbrot C (2002) Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor. *J. Biol. Chem.* 277:46265-72
- Stein RA, Staros JV (2006) Insights into the evolution of the ErbB receptor family and their ligands from sequence analysis. *BMC Evol Biol* 6:79
- Stern DF, Kamps MP (1988) EGF-stimulated tyrosine phosphorylation of p185neu: a potential model for receptor interactions. *Embo J* 7:995-1001
- Sternberg MJ, Gullick WJ (1989) Neu receptor dimerization. *Nature* 339:587
- Sternberg MJ, Gullick WJ (1990) A sequence motif in the transmembrane region of growth factor receptors with tyrosine kinase activity mediates dimerization. *Protein Eng.* 3:245-248
- Takahashi K, Yonezawa K, Nishimoto I (1995) Insulin-like growth factor I receptor activated by a transmembrane mutation. *J Biol Chem* 270:19041-5
- Taylor SS, Radzio-Andzelm E, Hunter T (1995) How do protein kinases discriminate between serine/threonine and tyrosine? Structural insights from the insulin receptor protein-tyrosine kinase. *Faseb J* 9:1255-66
- Tomita M, Furthmayr H, Marchesi VT (1978) Primary structure of human erythrocyte glycophorin A. Isolation and characterization of peptides and complete amino acid sequence. *Biochemistry* 17:4756-70
- Tzahar E, Pinkas-Kramarski R, Moyer JD, Klapper LN, Alroy I, Levkowitz G, Shelly M, Henis S, Eisenstein M, Ratzkin BJ, Sela M, Andrews GC, Yarden Y (1997) Bivalence of EGF-like ligands drives the ErbB signaling network. *EMBO J.* 16:4938-4950
- Tzahar E, Waterman H, Chen X, Levkowitz G, Karunakaran D, Lavi S, Ratzkin BJ, Yarden Y (1996) A hierarchical network of interreceptor interactions determines signal



- transduction by Neu differentiation factor/neuregulin and epidermal growth factor. *Mol Cell Biol* 16:5276-87
- Ullrich A, Coussens L, Hayflick JS, Dull TJ, Gray A, Tam AW, Lee J, Yarden Y, Libermann TA, Schlessinger J, et al (1984) Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. *Nature* 309:418-25
- Ullrich A, Schlessinger J (1990) Signal transduction by receptors with tyrosine kinase activity. *Cell* 61:203-12
- Vajdos FF, Adams CW, Breece TN, Presta LG, de Vos AM, Sidhu SS (2002) Comprehensive functional maps of the antigen-binding site of an anti-ErbB2 antibody obtained with shotgun scanning mutagenesis. *J Mol Biol* 320:415-28
- Van Obberghen E (1994) Signalling through the insulin receptor and the insulin-like growth factor-I receptor. *Diabetologia* 37 Suppl 2:S125-34
- Verlet L (1967) "Computer experiment" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Physical review* 159:98
- Wada T, Qian XL, Greene MI (1990) Intermolecular association of the p185neu protein and EGF receptor modulates EGF receptor function. *Cell* 61:1339-47
- Walker F, Orchard SG, Jorissen RN, Hall NE, Zhang HH, Hoyne PA, Adams TE, Johns TG, Ward C, Garrett TP, Zhu HJ, Nerrie M, Scott AM, Nice EC, Burgess AW (2004) CR1/CR2 interactions modulate the functions of the cell surface epidermal growth factor receptor. *J Biol Chem* 279:22387-98
- Wallasch C, Weiss FU, Niederfellner G, Jallal B, Issing W, Ullrich A (1995) Heregulin-dependent regulation of HER2/neu oncogenic signaling by heterodimerization with HER3. *EMBO J* 14:4267-75
- Walters RF, DeGrado WF (2006) Helix-packing motifs in membrane proteins. *Proc Natl Acad Sci U S A* 103:13658-63
- Walther D, Eisenhaber F, Argos P (1996) Principles of helix-helix packing in proteins: the helical lattice superposition model. *J Mol Biol* 255:536-553
- Wang JY, Koshland DE, Jr. (1978) Evidence for protein kinase activities in the prokaryote *Salmonella typhimurium*. *J Biol Chem* 253:7605-8
- Webster MK, Donoghue DJ (1996) Constitutive activation of fibroblast growth factor receptor 3 by the transmembrane domain point mutation found in achondroplasia. *EMBO J.* 15:520-527

- Weiner DB, Kokai Y, Wada T, Cohen JA, Williams WV, Greene MI (1989) Linkage of tyrosine kinase activity with transforming ability of the p185neu oncoprotein. *Oncogene* 4:1175-1183
- Weiner DB, Liu J, Cohen JA, Williams WV, Greene MI (1989) A point mutation in the neu oncogene mimics ligand induction of receptor aggregation. *Nature* 339:230-231
- White SH (2004) The progress of membrane protein structure determination. *Protein Sci* 13:1948-9
- Wiberg KB (1965) A Scheme for Strain Energy Minimization. Application to the Cycloalkanes. *J. AM. Chem. Soc.* 87:1070-1078
- Wides RJ, Zak NB, Shilo BZ (1990) Enhancement of tyrosine kinase activity of the Drosophila epidermal growth factor receptor homolog by alterations of the transmembrane domain. *Eur. J. Biochem.* 189:637-645
- Wood ER, Truesdale AT, McDonald OB, Yuan D, Hassell A, Dickerson SH, Ellis B, Pennisi C, Horne E, Lackey K, Alligood KJ, Rusnak DW, Gilmer TM, Shewchuk L (2004) A unique structure for epidermal growth factor receptor bound to GW572016 (Lapatinib): relationships among protein conformation, inhibitor off-rate, and receptor activity in tumor cells. *Cancer Res* 64:6652-9
- Xiong Q, Chan JL, Zong CS, Wang LH (1996) Two chimeric receptors of epidermal growth factor receptor and c-Ros that differ in their transmembrane domains have opposite effects on cell growth. *Mol Cell Biol* 16:1509-18
- Yamada K, Carpentier JL, Cheatham B, Goncalves E, Shoelson SE, Kahn CR (1995) Role of the transmembrane domain and flanking amino acids in internalization and down-regulation of the insulin receptor. *J Biol Chem* 270:3115-22
- Yamada K, Goncalves E, Carpentier JL, Kahn CR, Shoelson SE (1995) Transmembrane domain inversion blocks ER release and insulin receptor signaling. *Biochemistry* 34:946-54
- Yamada K, Goncalves E, Kahn CR, Shoelson SE (1992) Substitution of the insulin receptor transmembrane domain with the c- neu/erbB2 transmembrane domain constitutively activates the insulin receptor kinase in vitro. *J. Biol. Chem.* 267:12452-12461
- Yamaguchi H, Hendrickson WA (1996) Structural basis for activation of human lymphocyte kinase Lck upon tyrosine phosphorylation. *Nature* 384:484-9
- Yamamoto T, Ikawa S, Akiyama T, Semba K, Nomura N, Miyajima N, Saito T, Toyoshima K (1986) Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth factor receptor. *Nature* 319:230-234

- Yan H, Schlessinger J, Chao MV (1991) Chimeric NGF-EGF receptors define domains responsible for neuronal differentiation. *Science* 252:561-3
- Yarden Y, Sliwkowski MX (2001) Untangling the ErbB signalling network. *Nat Rev Mol Cell Biol* 2:127-37
- Yun CH, Boggon TJ, Li Y, Woo MS, Greulich H, Meyerson M, Eck MJ (2007) Structures of lung cancer-derived EGFR mutants and inhibitor complexes: mechanism of activation and insights into differential inhibitor sensitivity. *Cancer Cell* 11:217-27
- Zhang D, Sliwkowski MX, Mark M, Frantz G, Akita R, Sun Y, Hillan K, Crowley C, Brush J, Godowski PJ (1997) Neuregulin-3 (NRG3): a novel neural tissue-enriched protein that binds and activates ErbB4. *Proc. Natl. Acad. Sci. U S A* 94:9562-7
- Zhang X, Gureasko J, Shen K, Cole PA, Kuriyan J (2006) An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. *Cell* 125:1137-49
- Zong CS, Wang LH (1994) Modulatory effect of the transmembrane domain of the protein-tyrosine kinase encoded by oncogene *ros*: biological function and substrate interaction. *Proc. Natl. Acad. Sci. U.S.A.* 91:10982-10986

## **CHAPITRE III : RESULTATS**

## Introduction

La transduction du signal est un événement de première importance, puisqu'elle stimule la prolifération et la différenciation des cellules mais sa dérégulation entraîne les cellules sur la voie de la cancérogenèse. La famille des récepteurs ErbB est comme nous l'avons vu précédemment très souvent impliquée dans cette voie et de nombreux travaux sont effectués depuis des dizaines d'années pour trouver des solutions thérapeutiques. La plupart de ces études ont pour but la désactivation des récepteurs.

Les fragments peptidiques chimiques correspondant à l'ensemble du domaine extracellulaire ou à de courtes séquences du domaine transmembranaire sont capables d'interférer avec la dimérisation du récepteur et d'inhiber la croissance des cellules tumorales aussi bien au niveau de l'état monomérique que dimérique (Lofts et al. 1993; O'Rourke et al. 1997). Cette dernière voie nécessite de connaître la structure tridimensionnelle du récepteur ou celle des portions cibles. Il est évident que la connaissance de la structure complète des récepteurs ErbB/HER constituera une avancée très importante pour la compréhension de leur fonctionnement.

Notre objectif est de faire progresser les connaissances sur la structure de ces récepteurs en utilisant des techniques théoriques de prédiction de structure. Comme nous l'avons déjà évoqué, les techniques théoriques de prédiction combinées aux techniques de simulation de dynamique moléculaire s'appliquent généralement pour cibler les zones de flexibilité des macromolécules mais peut également être utilisée pour modéliser une structure 3D. Cependant, elle a également ses limites et il est difficile de déterminer ou de simuler la structure complète d'un récepteur du type ErbB dans son intégralité sans connaître la structure des domaines. Cette opération n'est pas faisable et le domaine le plus simple à modéliser est le segment transmembranaire (formé d'une seule hélice  $\alpha$ ) et son implication dans les phénomènes d'activation constitutive du récepteur en fait un candidat très intéressant. De plus, le succès de la prédiction de la structure de la partie transmembranaire de la Glycophorine A (Adams et al. 1996; MacKenzie et al. 1997; Smith et al. 2001) et de la structure du domaine TM de Neu par RMN (Goetz et al. 2001; Houliston et al. 2004) est encourageant. Notre ambition est de proposer des modèles d'association hétérodimérique des segments transmembranaires des récepteurs ErbB/HER en utilisant essentiellement comme principale

donnée, la composition en acides aminés de ces domaines. Ils nous permettent de prédire des modes d'association de ces hélices transmembranaires.

Je présenterai les résultats obtenus au cours de ma thèse en décrivant les études réalisées pour prédire des modes d'association hétérodimerique des hélices transmembranaires sans tenir compte des conditions environnementales de la membrane. Pour cela, je commencerai par décrire la méthodologie de prédiction d'association des hélices transmembranaires et je décrirai ensuite ceux issus d'une première étude réalisée sur les hétérodimères du domaine transmembranaire des récepteurs de la famille ErbB/HER.

Ce travail fait l'objet de deux publications dont l'une est publiée dans *European Biophysics Journal* (voir 1<sup>ère</sup> PARTIE). La deuxième publication (voir 2<sup>e</sup> PARTIE) vient d'être soumise à *Biopolymers*.

**1<sup>ère</sup> PARTIE: METHODOLOGIE DE PREDICTION DE L'ASSOCIATION DES  
HELICES TRANSMEMBRANAIRES. VALIDATION DE LA RECHERCHE DE  
STRUCTURES.**

## Introduction

Différentes méthodes ont été développées pour prédire l'assemblage des hélices TM. Toutes ces méthodes conduisent à une bonne prédiction de structures et leur performance se mesure par leur pouvoir de reproduction des résultats expérimentaux. La méthode que j'ai mise au point combine la mécanique et la dynamique moléculaire.

Dans le cas du domaine TM des récepteurs ErbB, on se heurte à une difficulté majeure puisque la structure des homodimères et des hétérodimères de ces récepteurs n'est pas connue. La structure hélicoïdale du domaine transmembranaire unique de Neu a été démontrée par voie expérimentale (Goetz et al. 2001; Gullick et al. 1992; Houliston et al. 2003, 2004; Smith et al. 1996). La flexibilité intrinsèque de l'hélice TM des récepteurs Neu et ErbB2 a été démontrée.

Dans l'équipe, une méthode de recherche globale associée aux simulations MD a été développée pour définir des modèles probables d'association des domaines TM de ErbB2 (Aller et al. 2006; Aller et al. 2005; Garnier et al. 2003; Garnier et al. 1997; Sajot et al. 1999). Son application, élargie à l'étude des associations du domaine TM de Neu, a permis de montrer que les domaines TM s'associent en formant un super enroulement gauche (Aller et al. 2006; Aller et al. 2005; Garnier et al. 2003; Garnier et al. 1997; Sajot et al. 1999).

Même si les modèles théoriques sont en très bon accord avec l'ensemble des données expérimentales, un point essentiel est de valider la méthode sur un système parfaitement connu. C'est la raison pour laquelle le dimère TM de la glycophorine A (GpA) est toujours pris comme base de validation.

La prédiction de la structure de ce domaine date des années 1994 (Langosch and Heringa 1998; Lemmon and Engelman 1994; Lemmon et al. 1994; Lupas 1996, 1997). Aujourd'hui encore ce système sert de modèle pour la validation de nouvelles méthodes de prédiction. Le dernier exemple est celui de Del'Ocro (Dell'Orco et al. 2007) qui a développé une méthode d'amarrage à corps rigides associé aux simulations MD.

La GpA est une protéine membranaire de 31 kDa présente à la surface des érythrocytes humains. C'est une sialoglycoprotéine majeure qui porte les déterminants antigéniques M et N des groupes sanguins. Elle peut également, dans certain cas, servir de ligand pour des virus, bactéries et parasites (Blanchard 1990; Cartron and Rahuel 1992).



L'enroulement droit caractéristique des domaines TM de la GpA est dicté par le motif spécifique de 7 résidus L75 I76 -- G79 V80 -- G83 V84 -- T87, disposé symétriquement à l'interface des deux hélices (Brosig and Langosch 1998; Lemmon et al. 1992; Lemmon et al. 1994). Ce domaine conserve une structure en hélice  $\alpha$  avec un angle de croisement des deux hélices voisin de  $-40^\circ$ .

Pour l'étude de la structure du domaine TM des hétérodimères ErbB, nous nous sommes basés sur la méthode initialement développée en 1997 (Aller et al. 2006; Aller et al. 2005; Garnier et al. 2003; Garnier et al. 1997; Sajot et al. 1999) et appliquée pour les études des homodimères (Aller et al. 2006; Aller et al. 2005; Garnier et al. 2003; Garnier et al. 1997; Sajot et al. 1999). La méthode a été adaptée pour traiter le cas d'association non symétrique, et les simulations MD utilisées pour le raffinement des structures initiales ont été poursuivies sur des échelles de temps relativement longues. Une méthode de sélection des structures représentatives a été développée en s'appuyant sur les connaissances acquises sur l'assemblage des hélices TM.

*Article EBJ*

# Molecular dynamics simulation approach for the prediction of transmembrane helix–helix heterodimers assembly

Oumarou Samna Soumana · Norbert Garnier ·  
Monique Genest

Received: 16 February 2007 / Revised: 11 May 2007 / Accepted: 15 May 2007  
© EBSA 2007

**Abstract** Computational methods are useful to identify favorable structures of transmembrane (TM) helix oligomers when experimental data are not available or when they cannot help to interpret helix–helix association. We report here a global search method using molecular dynamics (MD) simulations to predict the structures of transmembrane homo and heterodimers. The present approach is based only on sequence information without any experimental data and is first applied to glycoporphin A to validate the protocol and to the HER2-HER3 heterodimer receptor. The method successfully reproduces the experimental structures of the TM domain of glycoporphin A (GpA<sub>TM</sub>) with a root mean square deviation of 1.5 Å. The search protocol identifies three energetically stable models of the TM domain of HER2-HER3 receptor with favorable helix–helix arrangement, including right-handed and left-handed coiled-coils. The predicted TM structures exhibit the GxxxG-like motif at the dimer interface which is presumed to drive receptor oligomerization. We demonstrate that native structures of TM domain can be predicted without quantitative experimental data. This search protocol could help to predict structures of the TM domain of HER heterodimer family.

**Keywords** Molecular dynamics global search · Helix–helix association · Glycophorin A · ErbB/HER receptor

## Introduction

Membrane proteins encoded by nearly ~20–30% of the genome (Krogh et al. 2001; White 2004) are involved in complex signaling and transport functions. Because of their crucial role in many cellular processes they constitute the most biologically important family of proteins that serve as targets for drug discovery. However, their great biological importance contrasts with the very small number of structures resolved at high-resolution (Jayasinghe et al. 2001; White et al. 2001; Berman et al. 2002; White 2004). Solving membrane protein structures is necessary for a complete understanding of cellular processes (Fleishman and Ben-Tal 2006a, b). To overcome the difficulties to express membrane proteins and to obtain experimental structure using X-ray crystallography (White 2004) and nuclear magnetic resonance (NMR) (Opella and Marassi 2004) various computational methods for predicting the topology and the packing of transmembrane (TM) helices have been developed. Helical proteins are the dominant class of TM proteins and the identification of sequence segments embedded in membranes represents the fundamental step to progress into the membrane protein folding problem (Krogh et al. 2001; Chen et al. 2002; Kim et al. 2003; Kokubo and Okamoto 2004; Akula and Pattabiraman 2005; Adamian and Liang 2006; Fleishman et al. 2006b). Another step is to predict how TM helices associate. Various computational approaches have been developed all based on the two-stage model of membrane protein folding (Popot and Engelman 1990). The two simple rules are that helices are individually

---

Presented at the joint biannual meeting of the SFB-GEIMM-GRIP, Anglet France, 14–19 October 2006.

---

O. Samna Soumana (✉) · N. Garnier · M. Genest  
Centre de Biophysique Moléculaire, UPR 4301, CNRS,  
University of Orléans, rue Charles Sadron,  
45071 Orléans Cedex 02, France  
e-mail: soumana@cnrs-orleans.fr

O. Samna Soumana  
INSERM, rue Charles Sadron,  
45071 Orléans Cedex 02, France

inserted into the plasma membrane, and next, helices associate to shape the tertiary structure of the protein.

Specific interactions take place between hydrophobic segments within the bilayer and the major contribution to helix–helix interactions in membrane comes from van der Waals packing (Popot and Engelman 1990; White and Wimley 1999). Therefore, modeling approaches including molecular dynamics (MD) simulations and energy minimizations, combined with experimental information clearly identify structures of helix bundles (Herzyk and Hubbard 1995; Fleishman et al. 2004; Beuming and Weinstein 2005; Fleishman and Ben-Tal 2006a, b).

Computational methods can be used also to identify the most favorable interactions between TM helices when biochemical and biophysical data have no existence or when they cannot help to interpret helix–helix association. The best example is that of the TM dimer of glycoporphin A (GpA) for which global search MD studies (Treutlein et al. 1992; Adams et al. 1995, 1996) had first identified the structural model of the dimer, latter solved by NMR experiments in micelles (MacKenzie et al. 1997). Exhaustive studies have been undertaken to explore the determinants of stability in GpA dimerization highlighting the role of small residue sequences called the GxxxG motif, in mediating inter helix contacts (Russ and Engelman 2000). GxxxG-like motifs in which Gly residue is substituted by a small residue like Ala, Ser and Thr are frequently found in mediating TM helix association in a variety of protein families, like signal transduction proteins, channels, transporters and enzymes (Eilers et al. 2000; Dawson et al. 2002; Senes et al. 2004).

This motif is present in the TM domains of all the members of the HER family of growth factor receptor tyrosine kinase (ErbB1/HER or EGFR; ErbB2/HER2, ErbB3/HER3, ErbB4/HER4). Activation of HER receptors involves ligand-induced oligomerization, which in turn leads to receptor trans-phosphorylation and activation within dimers or higher order oligomers (Schlessinger 2000). Subsequent studies have shown that interactions between TM domains do contribute significantly to homo and hetero dimerization processes (Tanner and Kyte 1999; Mendrola et al. 2002). The single TM domain of HER members exhibits a conserved GxxxG-like motif at the N-terminal and the C-terminal ends. The HER<sup>3</sup><sub>TM</sub> domain exhibits this motif only at the N-terminus. Mutational analyses suggest that TM dimerization may involve one of the two motifs for a separate function (Fleishman et al. 2002; Mendrola et al. 2002). As an example, in the case of the HER<sup>1</sup><sub>TM</sub> and HER<sup>2</sup><sub>TM</sub> one motif at the C-terminus serves for homo dimerization and the other at the N-terminus for heterodimerization (Gerber et al. 2004).

Multiple short MD simulations in vacuum have been used to identify possible structures of the TM domain

association of the HER2/Neu receptors demonstrating the importance of the GxxxG-like motif in the formation of homodimers in the wild and the oncogenic forms (Garnier et al. 1997; Duneau et al. 1999a, b; Sajot and Genest 2000).

Besides the fact that the dimerization motif is involved in the association of the TM domain of the HER receptors, structural data on TM heterodimers are still unknown.

In this paper we describe a computational method that allows the prediction of structural models of HER heterodimers. For that, we have revisited our MD search protocol used to model helix–helix packing of symmetric homodimers (Garnier et al. 1997; Duneau et al. 1999a, b). In the case of heterodimers, symmetry constraints cannot be used and the overall energy surface has to be explored. The conformational sampling is described by the rotational pitch angle of each helix allowing the complete exploration of helix–helix interfaces.

Structural data of heterodimers of TM helices are not yet available (Caballero-Herrera and Nilsson 2003) and to test our search method it is first applied to the GpA<sub>TM</sub> dimer and second to the HER2-HER3 TM domain. HER2 is the preferred dimerization partner of all HER receptors and preferentially associate to HER3 (Tzahar et al. 1996; Graus-Porta et al. 1997). HER2-HER3 heterodimer shows the strongest potency in terms of stimulating cell proliferation (Pinkas-Kramarski et al. 1996; Rubin and Yarden 2001; Citri et al. 2003).

Our method successfully describes the right-handed structure of the GpA<sub>TM</sub> dimer with a root mean square deviation (RMSD) of 1.5 Å from the experimental structure (MacKenzie et al. 1997).

The excellent agreement obtained for GpA<sub>TM</sub> dimer leads us to explore TM domain structures of HER<sup>2-3</sup>. Experimental information relies on the presence of the GxxxG-like motif at the dimer interface. Our prediction shows that three MD models, with helices in right-handed and left-handed interactions conform to the presence of the GxxxG-like motif in the dimer interface.

## Methods

### TM domain systems

Two helical dimers are studied. The first one is the dimer of the GpA<sub>TM</sub> domain (MacKenzie et al. 1997) which is used as template to validate the method. Both helices comprise 23 residues from residue Ile73 to Ile95 constituting the hydrophobic sequence of the TM domain (Tomita et al. 1978). The second one results from the association of the HER2 and HER3 receptors referred to as the HER<sup>2-3</sup><sub>TM</sub> domain in the following. Each helix comprises the 25 residues of the hydrophobic sequence of the TM domains of

**Table 1** Sequence of the TM domains

Protein	Sequence	Residues
GpA <sub>TM</sub>	ITLIIFGVMAGVIGTILLISYGI	73-95
HER2 <sub>TM</sub>	LTSIVSAVVGILLVVVLGVVFGILI	651-675
HER3 <sub>TM</sub>	LTMALTVIAGLVVIFMMLGGTFLYW	642-666

The sequences of HER<sup>2</sup><sub>TM</sub> (Coussens et al. 1985; Yamamoto et al. 1986; Ehsani et al. 1993) and HER<sup>3</sup><sub>TM</sub> (Kraus et al. 1989) are listed, alongside the GpA<sub>TM</sub> sequence. Regions corresponding to the potential GxxxG dimerization motifs in each TM domain are shaded gray

the receptors (Coussens et al. 1985; Yamamoto et al. 1986; Kraus et al. 1989; Ehsani et al. 1993). The sequences of the two systems are given in Table 1. For both systems, each helix has the amino terminus acetylated and the carboxyl terminus methyl-aminated to mimic the peptide bond.

### Global search method

All calculations were undertaken by using the program CHARMM Version 26 (Brooks et al. 1983) and the all hydrogen PARAM22 force field (MacKerell et al. 1992). All calculations were carried out in vacuum and the dielectric constant was set to 1.0. This value is a good approximation of the membrane environment (Adams et al. 1995; Duneau et al. 1999a, b) and is justified for our purpose to define a set of possible structures of the TM dimers. Electrostatic interactions were not truncated and an atom-based switching function was used for the van der Waals interactions between 10 and 12 Å. The nonbonded list was generated with a cutoff of 13 Å. All calculations were performed on a multi-processor platform Linux Beowulf clusters with dual 3.06 GHz Intel Xeon processors.

### Conformational sampling

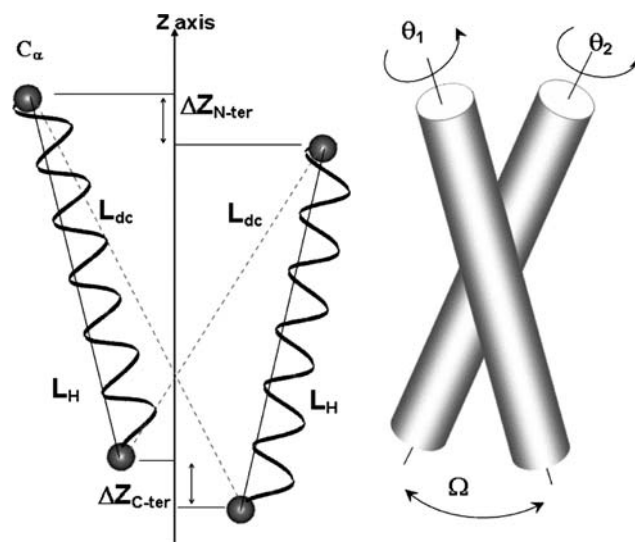
A parallel dimer was first generated with the two helices (H1 and H2) built with the standard torsion angles ( $\varphi = -57^\circ$ ,  $\psi = -47^\circ$ ) and the helix axes aligned along the Z-axis. This initial construction corresponding to two superimposed helices was used to generate multiple configurations of the dimer by rotating each helix about its own axis.  $\theta_1$  and  $\theta_2$  are the rotational angles of H1 and H2, respectively. For a complete search over the two-body rotational interaction space,  $\theta_1$  was increased by  $10^\circ$  steps from  $0^\circ$  to  $360^\circ$  and, for each  $\theta_1$  value,  $\theta_2$  was varied by  $10^\circ$  steps from  $0^\circ$  to  $360^\circ$ . A total of 1,296 ( $36 \times 36$ ) initial configurations was generated.

For each of these configurations, helix H2 was translated along the X-axis from its initial position (0 Å) by a distance  $D$  of 20 Å between the two helix axes sufficiently large to avoid atomic contacts.  $D$  was gradually decreased by increment of 0.25 Å to reach the distance Dh for which the

two helices are in favorable interactions. Dh is obtained as follows: at each translation step, the system was energy minimized using first the steepest descent (SD) method and then the Adopted Basis Newton–Raphston (ABNR) minimization algorithm for 50–100 steps. During the minimization process the two helices have been maintained in a parallel orientation by fixing the backbone atoms for the optimization of the side chain orientation. For the 1,296 minimized structures Dh ranges from 9.0 to 11 Å. These minimized structures have been used as starting structures for the simulations. This minimization process has been applied to the GpA<sub>TM</sub> dimer and to the HER<sup>2-3</sup><sub>TM</sub> dimer.

### Molecular dynamics protocol

Molecular dynamics simulations were carried out at constant temperature  $T = 300$  K. The SHAKE algorithm (Ryckaert et al. 1977) was applied to fix all the bonds involving hydrogen atoms. The integration time step was 1 fs. The MD protocol begins by a heating period (18 ps) during which the system was gradually thermalized from 0 to 300 K by increment of 50 K. New atomic random velocities were assigned from a Maxwell distribution at each temperature step. The  $\alpha$  helical structure was maintained by restraining the  $O_i-NH_{i+4}$  distance of the backbone hydrogen bonds (Hbonds) at 2.8 Å. The vertical shift of one helix relatively to the second helix along the Z-axis was



**Fig. 1** Parameters defining structure quality,  $L_H$ : distance between the  $C_\alpha$  atoms of the first residue  $C_\alpha^{H1Nter}$  and the last residue  $C_\alpha^{H1Cter}$ ;  $L_{dc}$ : crossed distance  $C_\alpha^{H1Nter}-C_\alpha^{H2Cter}$  and  $C_\alpha^{H2Nter}-C_\alpha^{H1Cter}$ ;  $\Delta dc$ : difference between the crossed distances.  $\Delta Z(Nter) = |Z(C_\alpha^{H1Nter}) - Z(C_\alpha^{H2Nter})|$ : difference in Z positions of the  $C_\alpha$  atoms of the first residue of H1 and H2;  $\Delta Z(Cter) = |Z(C_\alpha^{H1Cter}) - Z(C_\alpha^{H2Cter})|$ : difference in the Z positions of the  $C_\alpha$  atoms of the last residue of H1 and H2.  $\Omega$  ( $^\circ$ ): crossing angle.  $\theta_1$  and  $\theta_2$  are the rotational angles of H1 and H2, respectively, about their own axis

monitored by applying harmonic crossed distance restraints  $L_{dc}$  (see Fig. 1 for definition) at 35 Å and 36 Å with a force constant of 250 kcal mol<sup>-1</sup> Å<sup>-2</sup> for the GpA<sub>TM</sub> dimer and the HER<sup>2-3</sup><sub>TM</sub> dimer, respectively. The values are based on those measured on the experimental model of the GpA<sub>TM</sub> dimer (MacKenzie et al. 1997) and our previous MD models of HER<sup>2</sup><sub>TM</sub> homodimers (Sajot and Genest 2000).

Equilibration was continued for 1 ns applying the same  $\alpha$  Hbonds and crossed distances restraints with a force constant of 250 kcal mol<sup>-1</sup> Å<sup>-2</sup> for the first 300 ps and reduced to 150 kcal mol<sup>-1</sup> Å<sup>-2</sup> for the following 300 ps. The last 400 ps were continued without any constraint allowing the two helices to move freely. During the heating and the equilibration periods helix–helix separation  $D$  was not restrained.

The production period was conducted during 3 ns without any restraint. Atomic coordinates and energy terms were saved every ps. The last ns was used for averaging and further analysis.

## Analysis of helix–helix assembly

### Geometric criteria

Several geometrical parameters were used to evaluate the quality of the simulated structures. These parameters, shown in Fig. 1, are: (1) the helix length  $L_H$  defined as the distance between the C $_{\alpha}$  atoms of the first residue at the N-terminus (C $_{\alpha}$ <sup>H</sup>Nter) and of the last residue at the C-terminus (C $_{\alpha}$ <sup>H</sup>Cter). The  $L_H$  value informs directly on helix elongation or helix contraction resulting from  $3_{10}$  or  $\pi$  helix distortions. Distortions of the helical backbone are currently observed during MD simulations. We have considered, after several tests that a  $\alpha$  helix is structurally correct when  $L_H$  does not deviate up to 3 Å from 32 Å for the GpA helix and 35 Å for the HER helix. For a canonical  $\alpha$  helix,  $L_H$  is about 35.25 Å for a 25 residues helix (HER<sub>TM</sub>) and 32.5 Å for a 23 residues helix (GpA<sub>TM</sub>). These values are close to those measured along the Z-axis. One may notice that  $L_H$  can also be obtained from the calculation of the helix length along the Z-axis. However, the helix axis is not correctly defined in this case because the  $\alpha$  helix structure is not strictly maintained after MD simulations. (2) The difference between the two crossed distances ( $L_{dc}$ ). This parameter noted  $\Delta_{dc}$  allows the control of both the vertical shift of the helices and the length of the two helices. (3) The helix register  $\Delta Z$  defines the vertical position of the helices relatively to each other. It is calculated as the difference between the Z positions of the C $_{\alpha}$  atoms of the first and the last residues of H1 and H2 along the dimer axis. For bitopic membrane proteins a significant shift of the TM helices along the membrane normal would be impossible due to the limitations imposed by the lipid bilayer (Killian and von Heijne 2000; Killian 2003). (4) The crossing angle

$\Omega$ , which measures the angle between the two helical axes whereby positive and negative values indicate a left-handed and a right-handed coiled-coil. A distribution of the crossing angle values ranging from  $-65^\circ$  to  $45^\circ$  is in accordance with the various studies on membrane protein assembly (Bowie 1997; Chothia et al. 1981; Bywater et al. 2001; Kallblad and Dean 2004). Bowie (1997) examined a small database of helical membrane proteins and showed that a crossing angle of  $+20^\circ$  was the most common, in contrast to the globular proteins where packing angles around  $-35^\circ$  are the most prevalent. According to Chothia et al. (1981) inter-helical crossing angles range between  $-80^\circ$  and  $+40^\circ$ . Values of the geometric parameters used to define the dimer structures quality are given in Table 2.

## Structure selection

### Geometry-based filters

The quality of a dimer structure was determined from the analysis of the 250 configurations extracted from the last ns MD trajectory (one configuration extracted every 4 ps). The structure selection follows several steps:

(1) For each of the 250 configurations, the 6 parameters defined previously were calculated. A score value of 1 is given when one parameter is respected otherwise the score value is 0. A configuration is assumed structurally correct for a total score of 6 or close to correct for a total score of 5. Otherwise, the configuration is rejected. (2) A structure is retained if 80% of the analyzed configurations ( $N$ ) are structurally correct ( $N \geq 200$ ). Otherwise, the structure is rejected. Steps (1) and (2) applied to the 1,296 structures drastically reduce the number of structurally correct dimers. (3) Based on the distribution of the crossing angle, the remaining structures are clustered into three distinct families of coiled-coils corresponding to right-handed dimers ( $\Omega < -10^\circ$ ), parallel dimers ( $-10^\circ \leq \Omega \leq 10^\circ$ ) and left-handed dimers ( $\Omega > 10^\circ$ ). (4) For each family, several clusters including structures with a C $\alpha$  RMSD value  $\leq 2$  Å are well identified. The average structure for each cluster was then energy minimized in order to remove bad contacts

**Table 2** Parameters defining dimer structure quality

	$L_H$ (Å)		$\Delta_{dc}$ (Å)	$\Delta Z_{Nter}$ $\Delta Z_{Cter}$ (Å)	$\Omega$ (°)	
	Min	Max			Min	Max
HER <sup>2-3</sup> <sub>TM</sub>	32.0	38.0	$\leq 3.5$	$\leq 3.5$	$-65^\circ$	$+45^\circ$
GpA <sub>TM</sub>	29.0	35.0	$\leq 3.5$	$\leq 3.5$	$-65^\circ$	$+45^\circ$

Minimum and maximum values of  $L_H$  are defined from the length of a  $\alpha$  canonical helix ( $L_{Hz} = 35.25$  Å for the HER<sub>TM</sub> helix and 32.5 Å for the GpA<sub>TM</sub> helix).  $\Omega$  (°) is the crossing angle. Minimum and maximum values derived from Bowie (1997) and statistical analysis of membrane proteins. All these parameters are shown in Fig. 1

and was used as the starting point for further MD refinement. The same exact MD protocol described previously was applied.

In addition, it has been applied to the two experimental models of the GpA<sub>TM</sub> dimer, (model 16 and the average NMR structure) for comparison (see section “[Results and discussion](#)”).

#### Interface surface area calculation

The interface surface area (SA) was calculated as the difference between the solvent accessible SA of individual helices and the solvent accessible SA calculated for the dimer. Surface areas are calculated with CHARMM and values are averaged over the last ns of the simulation.

Residues in the TM helices were divided in three groups: interacting residues have over 80% of their SA buried in the helical dimer, partially buried residues have buried SAs between 50 and 80% and exposed residues have more than 50% of their SA accessible (Gimpelev et al. 2004).

#### Binding free energy calculation

We have calculated the gas-phase contribution to the binding free energy,  $\Delta G^{\circ}_{\text{bind}}$  defined as: (Zoete et al. 2005)

$$\text{Vacuo (Helix1 + Helix2)} \xrightarrow{\Delta G^{\circ}_{\text{bind}}} \text{Dimer}$$

$\Delta G^{\circ}_{\text{bind}}$  is the sum of three terms:  $\Delta E_{\text{intra}}$ , the difference in the internal energy of the two monomers H1 and H2 in the dimer  $D$  and the isolated monomers,  $E_{\text{vdW}}$ , and  $E_{\text{elec}}$ , the van der Waals and electrostatic inter-helical interactions energies.

$$\begin{aligned} \Delta G^{\circ}_{\text{bind}} &= \Delta E_{\text{intra}} + E_{\text{vdW}} + E_{\text{elec}} \\ \Delta E_{\text{intra}} &= E_{\text{intra}}^D - E_{\text{intra}}^{\text{H1}} - E_{\text{intra}}^{\text{H2}} \\ E_{\text{intra}}^X &= E_{\text{intra,bond}}^X + E_{\text{intra,vdW}}^X + E_{\text{intra,elec}}^X \end{aligned}$$

## Results and discussion

### GpA<sub>TM</sub> dimers

The 1,296 sampled structures formed three families representative of the right-handed, left-handed and parallel dimers with 73, 70 and 16 structures respectively after geometric filters.

Clustering similar structures in each family ( $C\alpha$  RMSD  $\leq 2$  Å), followed by averaging and MD refinement reduce their population. Finally, four MD models representative of the right-handed helical dimers were obtained, nine are representative of the left-handed dimers and one exhibits parallel helices.

In order to check whether the present protocol allows predicting the experimental structure,  $C\alpha$  RMSDs of all the final MD models relative to two NMR models were calculated. One is the minimized structure of model 16, referred to as 1AFO<sub>16</sub> which exhibits the lowest  $C\alpha$  RMSD value relatively to the 14 final GpA<sub>TM</sub> MD models. The second is the minimized average NMR structure which will be noted 1AFO<sub>mean</sub>. The 20 NMR models are very similar with crossed  $C\alpha$  RMSD values ranging from 0.3 to 1.76 Å. RMSDs of these 20 models from 1AFO<sub>mean</sub> range from 0.34 to 0.94 Å. 1AFO<sub>16</sub> exhibits one of the largest value (0.90 Å). The lowest  $C\alpha$  RMSDs from 1AFO<sub>16</sub> were obtained for the right-handed models ranging from 1.5 to 4.7 Å.

For the left-handed dimers and the parallel dimer RMSDs are greater than 5.9 Å. All the right-handed structures, referred to as GpA<sub>R</sub>, show relatively short  $C\alpha$ – $C\alpha$  distances between the two helices ( $\leq 4.4$  Å) which are correlated with short distances between the two helical axes ( $\leq 7.4$  Å).

Table 3 gives the various characteristics of the right-handed models and of two left-handed models, noted here GpA<sub>L</sub>, selected on the basis of the strongest helix–helix interactions (GpA<sub>L1</sub>) and the weakest helix–helix interactions (GpA<sub>L2</sub>).

Three of the GpA<sub>R</sub> models exhibit a crossing angle around  $-45^{\circ}$ , very close to that calculated for 1AFO<sub>16</sub> and 1AFO<sub>mean</sub>. Calculated helix–helix angles are also in the range of values derived from MD simulation in membrane ( $-39^{\circ}$  to  $-46^{\circ}$ ) (Petrache et al. 2000; Cuthbertson et al. 2006), or in micelles ( $-44^{\circ}$ ) (Braun et al. 2004). As shown in Table 3, the magnitude of inter-helical interactions is correlated with the magnitude of accessible SA of the dimer interface. The same is observed for binding free energy. The two left-handed structures exhibit short  $C\alpha$ – $C\alpha$  residue distances in the same range of value as observed for right-handed dimers. At least, one Gly residue—Gly79 or Gly83—is involved at the dimer interface. Left-handed structures largely differ from the experimental structure 1AFO<sub>16</sub> attested by RMSD values of 5.9 and 6.5 Å. The only parallel dimer resulting from the last filter does not exhibit interesting properties and thus is not discussed here.

It is interesting to remark that the total energy of the different MD dimers around  $-110 \pm 11$  kcal mol<sup>-1</sup> does not allow to differentiate the right-handed and the left-handed structures. The left-handed model GpA<sub>L1</sub> exhibits the best stability, the strongest helix–helix interactions and, the greatest contacting surface at the interface, but is the furthest from 1AFO<sub>16</sub>. The right-handed model GpA<sub>R1</sub> which is the less stable and shows the weakest helix–helix interactions and the smallest contacting surface is very close to the experimental structure. This is clearly shown by the very low values of RMSD calculated over the residues of the dimerization motif.

**Table 3** GpA<sub>TM</sub> simulated dimer structures

GpA <sub>TM</sub> models	$\Omega$ (°)	$D$ (Å)	$d_{C\alpha-C\alpha}$ (Å)	R <sub>i</sub> -R <sub>j</sub>	SA (Å <sup>2</sup> )	$E_{\text{tot}}$ dimer	$E_{\text{inter}}$	$\Delta G^{\circ}_{\text{bind}}$	C $\alpha$ RMSD (Å)	
									a	b
GpA <sub>R1</sub>	-44	6.2	4.1	G79-G79	513	-110	-50	-104	1.5(1.8)	0.6(0.7)
GpA <sub>R2</sub>	-43	7.0	4.2	G83-M81	585	-118	-57	-132	3.4(3.3)	2.9(2.9)
GpA <sub>R3</sub>	-46	6.8	4.0	G86-G86	580	-121	-60	-128	4.7(4.7)	4.8(4.9)
GpA <sub>R4</sub>	-35	7.4	4.4	I77-G79	598	-118	-59	-129	3.5(3.2)	2.9(2.8)
GpA <sub>L1</sub>	16	8.0	4.5	G79-F78	672	-121	-82	-152	6.5(6.0)	4.3(4.2)
GpA <sub>L2</sub>	25	8.9	4.9	G83-V84	528	-100	-43	-79	5.9(5.4)	4.0(3.9)
1AFO <sub>16</sub>	-50	6.2	3.5	V80-V80	432	-99				
1AFO <sub>mean</sub>	-40	6.9	3.6	V80-V80	424	-100				

The best representative structures of the GpA<sub>TM</sub> include right-handed coiled-coils (GpA<sub>R</sub>) and left-handed coiled-coils (GpA<sub>L</sub>). Geometric properties and energy terms are averaged over the last ns of the 4 ns MD simulation. Crossing angle  $\Omega$  (°) is calculated following a method by Chothia et al. (1981) included in CHARMM. A positive value characterizes a left-handed coiled-coil and a negative value characterizes a right-handed coiled-coil.  $D$  (Å): inter-helical distance between the two helices;  $d_{C\alpha-C\alpha}$  (Å): minimum distances between C $\alpha$  atoms of opposed residues; SA (Å<sup>2</sup>): contacting surface area at the dimer interface. The two residues in closest proximity are given as R<sub>i</sub>-R<sub>j</sub> respectively for GpA<sub>H1</sub> and GpA<sub>H2</sub>. The total energy of the dimer ( $E_{\text{tot}}$  dimer) is the sum of the intramolecular energy of the two helices and helix-helix interactions ( $E_{\text{inter}}$ ).  $\Delta G^{\circ}_{\text{bind}}$  is the mean value of the binding free energy in the gas-phase (see text for definition). All energy terms are in kcal/mol. C $\alpha$  RMSD values from the minimized structures 1AFO<sub>16</sub> and 1AFO<sub>mean</sub> (italic) (model 16 and average NMR structure): (a) over all C $\alpha$  atoms and (b) over the C $\alpha$  atoms of the G<sub>xxx</sub>G motif. [RMSD values over the C $\alpha$  atoms of the seven residues L<sub>ixx</sub>G<sub>vxx</sub>G<sub>vxx</sub>T are very similar ( $\pm 0.15$ ).] The various geometric and energetic properties are given for the minimized structure

### GpA<sub>TM</sub> dimer interface

Interfacial residues involved in helix-helix association were identified from the C $\alpha$ -C $\alpha$  distance maps and the calculations of the buried SAs.

In Fig. 2, the C $\alpha$ -C $\alpha$  contact map of GpA<sub>R1</sub> is compared to that of 1AFO<sub>16</sub> and 1AFO<sub>mean</sub>. One may note the strong similarity of distance patterns and particularly that of the two NMR models although model 16 is the more distant from the mean structure. The representation of models clearly confirm their strong similarity.

For GpA<sub>R1</sub>, the three residues, Gly79, Gly83 and Thr87, have been identified as completely buried. The contact map highlights the well-known interacting residue pairs Gly79-Gly79, Gly83-Gly83, Gly79-Val80 and Gly83-Val80 pertaining to the dimerization motif. In addition to these contacting pairs, others are revealed between Gly83-Val84, Gly79-Ile76, Leu75-Ile76 and the symmetry-related Leu75-Leu75, Ile76-Ile76, and Thr87-Thr87.

For a better comparison with our MD models we have carried on a 4 ns MD simulation of 1AFO<sub>16</sub> and of 1AFO<sub>mean</sub> under the same conditions as described in MD protocol. Mean values of different quantities have been calculated. The C $\alpha$ -C $\alpha$  distances less than 6 Å between the two helices are reported in Table 4 for the two experimental models and are compared to those of the six others MD models. It is shown that for both NMR structures, the shortest distances imply the same residues of the G<sub>xxx</sub>G motif. One may note that the V80-V80 distance evolves from the shortest to the largest during simulation process.

We demonstrate that the proximities between the residues of the G<sub>79</sub>VMAG<sub>83</sub> motif calculated for GpA<sub>R1</sub> are almost identical to those of the experimental models. For the others MD models these proximities are not respected.

Our findings are very similar to those obtained recently by (Bu et al. 2007) describing the GpA<sub>TM</sub> structure using implicit membrane generalized Born model.

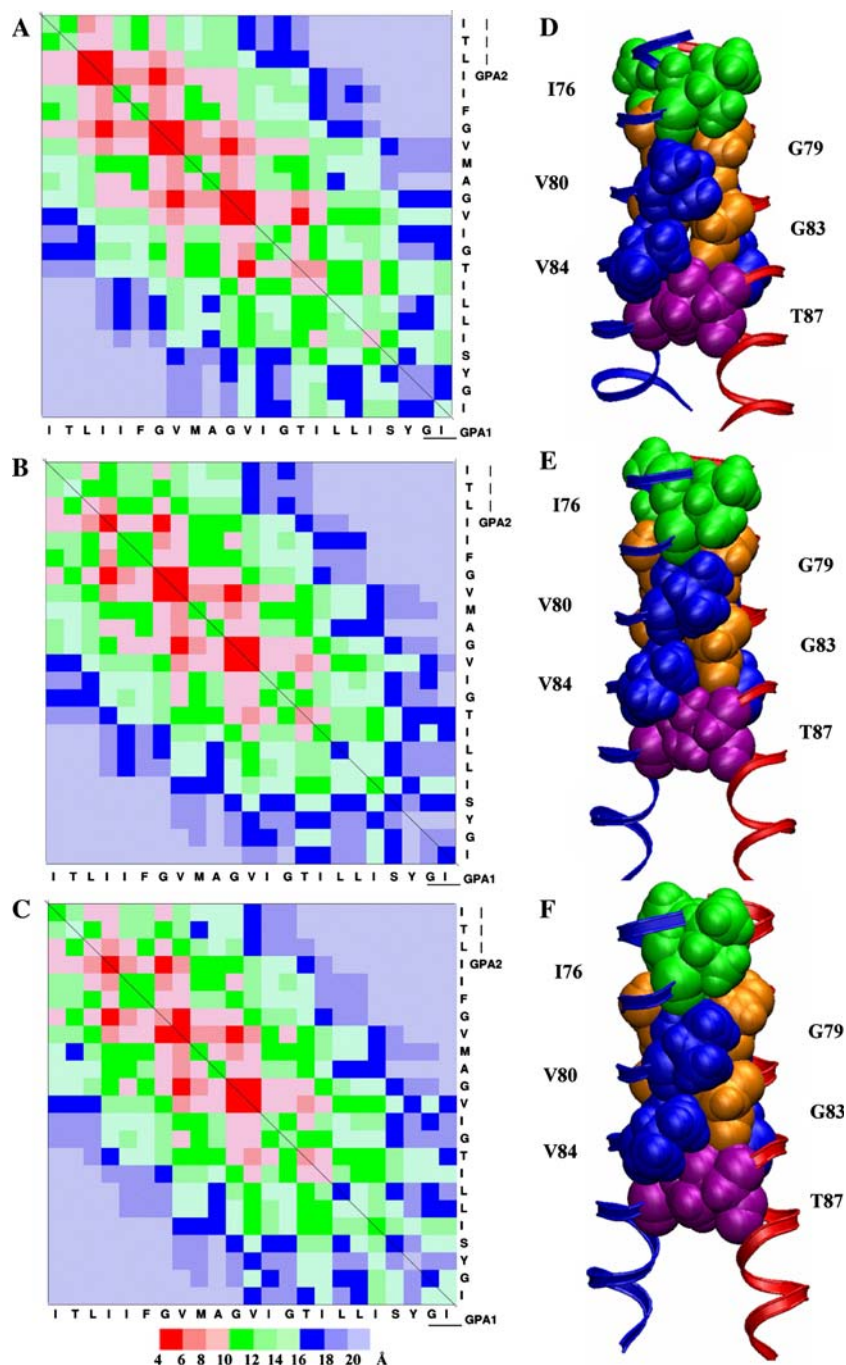
The same residue pairs were identified in GpA structures simulated in dimyristoyl-phosphatidylcholine (DMPC) bilayers and dodecyl phosphocholine (DPC) micelles (Cuthbertson et al. 2006).

Although the present study is performed in vacuum the seven residue sequence L<sub>ixx</sub>G<sub>vxx</sub>G<sub>vxx</sub>T is found to be the motif involved in GpA<sub>TM</sub> helix association.

Surprisingly, even though the close distance between Gly residues at the dimer interface is believed to allow hydrogen bonding interactions between the Thr87 side chains (Cuthbertson et al. 2006) or between the Thr87 side chain and the backbone carbonyl of Val84 residue of the opposed helix (Smith et al. 2002), our simulations did not highlight any Hbond. A possible explanation is that, in micelles, Thr87 can be solvated and water molecules may mediate Hbond interactions between the two side chains (Smith et al. 2002). Also, polar residues are destabilizing residues in detergents and because of helix unravelling Hbonds could be favoured with water (MacKenzie and Engelman 1998). Inter-helical H-bonding interactions are not the only determinant factor for inter-helical association. Other factors can compensate for the loss of this particular interaction.



**Fig. 2** GpA<sub>TM</sub> dimer structures. C $\alpha$ -C $\alpha$  contact maps of the GpA<sub>TM</sub> dimer: **a** GpA<sub>R1</sub> derived from MD search; **b** experimental structure 1AFO<sub>16</sub> (1AFO model 16); **c** average structure of all NMR models. The lowest C $\alpha$ -C $\alpha$  distances are in red and the largest distances are in blue. The corresponding structures GpA<sub>R1</sub> (**d**), 1AFO<sub>16</sub> (**e**) and 1AFO average structure (**f**) are shown with the closest residues at the dimer interface



### HER<sup>2-3</sup><sub>TM</sub> dimers

From the 1,296 sampled structures, three families representative of the right-handed, left-handed and parallel dimers with size of 2, 77 and 20, respectively, were obtained after geometric filters. These first results suggest a strong preference for a left-handed association for the TM domains of the HER<sup>2-3</sup> heterodimer.

After clustering (C $\alpha$  RMSD  $\leq$  2 Å), averaging and MD refinement the number of structures representative of each

family is strongly reduced. Only one structure with a right-handed crossing angle was retained and five with a left-handed crossing angle. The dimer with parallel helices resulting from this last step was rejected because of significant distortions at the N-terminal end of both helices. Geometrical properties and energetic data are given in Table 5.

The five left-handed structures are slightly different or relatively different as attested by the C $\alpha$  RMSD values that compare all the structures (between 2.7 and 5.1 Å). The

**Table 4**  $C\alpha$ - $C\alpha$  distances ( $\leq 6$  Å) calculated between the residues of the two helices  $R_i$ - $R_j$  for the MD models predicted for GpA<sub>TM</sub> dimer (right-handed and left-handed) and for the two NMR structures: 1AFO<sub>16</sub> (model 16) and the average structure of the 20 NMR models (1AFO<sub>mean</sub>)

$R_i$ - $R_j$	GpA <sub>R1</sub>	GpA <sub>R2</sub>	GpA <sub>R3</sub>	GpA <sub>R4</sub>	GpA <sub>L1</sub>	GpA <sub>L2</sub>	1AFO <sub>mean</sub>	1AFO <sub>16</sub>		
G79-G79	4.14						4.31	<i>4.48</i>	4.26	<i>5.11</i>
G83-V80	4.28						4.29	<i>4.58</i>	4.78	<i>4.31</i>
G83-G83	4.40						4.43	<i>5.35</i>	4.40	<i>5.19</i>
V80-G79	4.50			5.03			4.53	<i>4.49</i>	4.80	<i>4.29</i>
V80-G83	4.56						4.66	<i>4.66</i>	4.22	<i>4.17</i>
G79-V80	4.64	5.00					4.88	<i>4.88</i>	4.50	<i>4.23</i>
V80-V80	5.19						5.24	<i>5.24</i>	5.27	<i>3.51</i>
G83-M81		4.16			4.79					
G79-M81		4.61			5.33	5.31				
A82-M81		5.86								
V80-M81		5.86				5.60				
A82-A82			4.61							
M81-G79				4.51						
M81-G83				4.58						
M81-V80				5.76						

Values are averaged over the last ns of the simulation. Values from the minimized structures are in italic

crossing angle ranges from 13° and 31°. As shown in Table 5, all these models exhibit energetically stable structural arrangement of the helices leading to stable dimers. Inter-helical interactions differ at the most by 20 kcal/mol and difference in internal energy of the dimers between the lowest structure and the weakest structure is in the same range of value. Energetic factors are not sufficient to distinguish the most representative structure of the TM domain of the HER<sup>2-3</sup> heterodimer. The structure of this domain is unknown and any structural data is available. The only information which can be used to select the best representative structure of the HER<sup>2-3</sup><sub>TM</sub> domain is the presence of the GxxxG-like motif at the dimer interface. As revealed by numerous studies on TM helices of membrane proteins this motif is presumed to drive helix-helix association (Russ and Engelman 2000; Dawson et al. 2002;

Mendrola et al. 2002; Arselin and Giraud 2003; McClain et al. 2003). This hypothesis is confirmed by biological studies on HER<sub>TM</sub> domains (Mendrola et al. 2002; Gerber et al. 2004). Three dimer models conform to this hypothesis with similar free binding energy values. One is the right-handed structure referred to as HER<sup>2-3</sup><sub>R</sub>, the two others are the left-handed structures referred to as HER<sup>2-3</sup><sub>L1</sub> and HER<sup>2-3</sup><sub>L2</sub>. Helix-helix interface of these three structures are detailed below.

#### HER<sup>2-3</sup><sub>TM</sub> dimer interface

The  $C\alpha$ - $C\alpha$  distance maps and the views of HER<sup>2-3</sup><sub>R</sub>, HER<sup>2-3</sup><sub>L1</sub> and HER<sup>2-3</sup><sub>L2</sub> structures are shown in Fig. 3. The contact map relative to the right-handed structure (a) shows that the shortest distances involve the N-terminal

**Table 5** HER<sup>2-3</sup><sub>TM</sub> simulated dimer structures

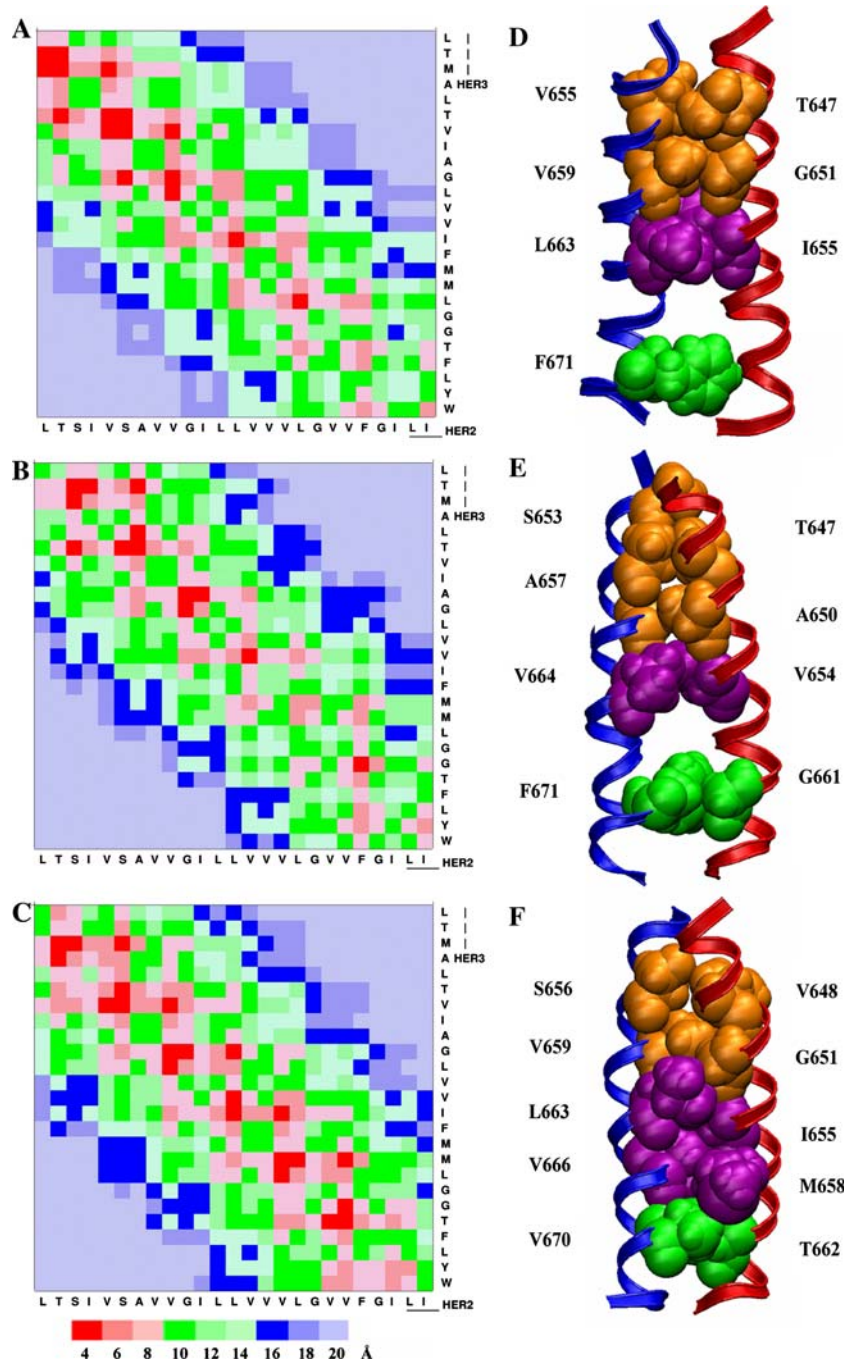
HER <sup>TM</sup> <sub>2-3</sub> models	$\Omega$ (°)	$D$ (Å)	$d_{C\alpha-C\alpha}$ (Å)	$R_i$ - $R_j$	SA (Å <sup>2</sup> )	$E_{tot}$ dimer	$E_{inter}$	$\Delta G^{\circ}_{bind}$
HER <sub>R</sub> <sup>2-3</sup>	-20	8.3	5.6	L663-I665	716	-162	-59	-157
HER <sub>L1</sub> <sup>2-3</sup>	24	8.1	4.4	A657-T647	687	-175	-65	-161
HER <sub>L2</sub> <sup>2-3</sup>	31	7.8	5.5	S656-V648	694	-167	-53	-152
HER <sub>L3</sub> <sup>2-3</sup>	22	8.5	5.2	I673-L659	636	-154	-53	-140
HER <sub>L4</sub> <sup>2-3</sup>	13	8.7	4.8	G668-G660	590	-168	-42	-140
HER <sub>L5</sub> <sup>2-3</sup>	26	8.0	5.1	S656-T647	693	-173	-63	-169

The best representative structures of the HER<sup>2-3</sup><sub>TM</sub> include right-handed coiled-coils (HER<sup>2-3</sup><sub>R</sub>) and left-handed coiled-coils (HER<sup>2-3</sup><sub>L</sub>). Geometric properties and energy terms are averaged over the last ns of the 4 ns MD simulation. Definitions of the geometric parameters, crossing angle  $\Omega$  (°),  $D$  (Å),  $d_{C\alpha-C\alpha}$  (Å), SA (Å<sup>2</sup>) and energy terms (in kcal/mol) are defined as in Table 3. The two residues in closest proximity are given as  $R_i$ - $R_j$  respectively for HER<sup>2</sup><sub>TM</sub> and HER<sup>3</sup><sub>TM</sub>

residues of the helices. The shortest proximities are observed for the following residue pairs: L651T652–T643M644, T652–T647, V655S656–T647V648, V655–M644, S656–G651, V659–V648, V659–G651L652, L663–I655 and L667–L659. Contacting maps of the two left-handed structures are slightly different. The closest proximities are observed between the N-terminal residues for  $HER^{2-3}_{L1}$  (b) while they are observed all along the TM sequences for  $HER^{2-3}_{L2}$  (c). In the case of  $HER^{2-3}_{L1}$ , the shortest distances involve the following residue pairs S656–T647,

A657–T647, G660–A650, G660–G651, I661–A650 at the N-terminus and F671–G661 at the C-terminus. For  $HER^{2-3}_{L2}$ , a strong implication of S656 and G660 of the  $HER^2_{TM}$  and T647 and G651 of  $HER^3_{TM}$  are revealed. One may remark a strong implication of the C-terminal residues and particularly those of the C-terminal GxxxG-like motif of  $HER^2_{TM}$ . All the three structures clearly exhibit the N-terminal GxxxG-like motif at the dimer interface. All the residues involved in short interfacial  $C\alpha$ – $C\alpha$  distances are completely or almost completely buried.

**Fig. 3**  $HER^{2-3}_{TM}$  dimer structures.  $C\alpha$ – $C\alpha$  contact maps of the representative models of  $HER^{2-3}_{TM}$  dimer derived from MD search: right-handed structure  $HER^{2-3}_{R1}$  (a), left-handed structures  $HER^{2-3}_{L1}$  (b) and  $HER^{2-3}_{L2}$  (c). The lowest  $C\alpha$ – $C\alpha$  distances are in red and the largest distances are in blue. The corresponding structures  $HER^{2-3}_{R1}$  (d),  $HER^{2-3}_{L1}$  (e) and  $HER^{2-3}_{L2}$  (f) are shown with the closest residues at the dimer interface. Backbone ribbon of  $HER^2_{TM}$  helix is in blue and that of  $HER^3_{TM}$  helix is in red. At the N-terminus, buried residues of the GxxxG-like motif are colored orange, buried residues at the C-terminus including residues of the GxxxG-like motif of  $HER^2_{TM}$  are shown in green. Other buried residues are in purple



## Concluding discussion

In the present study, we show that sequence information and oligomerization state are sufficient to predict structural models of TM helix assembly.

Because MD simulations in membrane are very time consuming, our study, as most of others that aimed to explore the conformational space of helix–helix assembly are undertaken in vacuum (Adams et al. 1995; Torres et al. 2001; Caballero-Herrera and Nilsson 2003). It is largely admitted that helix–helix contacts are the main factor that govern helix–helix association and that the membrane environment does not exert major effects (White and Wimley 1999). Membrane acts as a modulator of the stability of oligomers (Petrache et al. 2000; Cuthbertson et al. 2006). The conformational space explored in vacuum is very close to that explored in hydrated lipids (Duneau et al. 1999a, b). Thus one may think that the structures predicted in vacuum are close to those in a real membrane provided that they respect the properties of the membrane/solvent environment.

Our approach is based on the global searching method developed for predicting HER2/Neu homodimers (Garnier et al. 1997; Sajot and Genest 2000) and was revisited to explore the whole conformational space without imposing limitation. This is in contrast to several other methods that impose rotational symmetry, helix tilt angle or helix–helix angle (Krogh et al. 2001; Chen et al. 2002; Kim et al. 2003; Kokubo and Okamoto 2004; Akula and Pattabiraman 2005; Adamian and Liang 2006; Beevers and Kukol 2006; Fleishman et al. 2006b; Bu et al. 2007). Another advantage is the MD exploration of a relatively limited number of configurations for a given heterodimer system. Thus, the present approach is very attractive to define rapidly, without a priori, an ensemble of structures representative of the native state of TM heterodimers. Such an approach could be broadened to the prediction of larger oligomeric states by taking into account additional criteria as those deduced from statistical analysis of membrane proteins.

The present MD search protocol is successful in predicting the structure of the GpA<sub>TM</sub> dimer. Any quantitative experimental data is used except the knowledge of the proximity of the GxxxG motifs at the dimer interface.

Obviously, the success of a prediction method is difficult to evaluate without experimental data. In the present work, only geometric criteria are used to propose reasonable structures of two interacting helices. One difficulty is to define a suitable balance between the number of filters and their degree of severity to minimize the risk to predict wrong structures. More the number of filters is large, less is the risk. More accurate prediction imposes the use of experimental data, at least qualitative.

The question of the correctness of a predictive method for describing the structure of helix–helix assembly is of

major importance owing to the difficulty of experimental studies in membrane environment. The biological importance of the HER receptors and the lack of knowledge of homo and hetero dimerization processes lead us to investigate structural models of the TM domain. Structural models of the HER<sup>2–3</sup><sub>TM</sub> domain are proposed, based only on the presence of the GxxxG-like dimerization motif at helix–helix interface (Mendrola et al. 2002). As discussed above, one may think that they are representative of the structures in the membrane environment.

The present method is currently used for the prediction of the structures of the TM domain of all HER heterodimers with the aim to give insights into the hierarchical network of inter receptor interactions that determine signal transduction (Tzahar et al. 1996; Duneau et al. 2007).

## References

- Adamian L, Liang J (2006) Prediction of transmembrane helix orientation in polytopic membrane proteins. *BMC Struct Biol* 6(1):13
- Adams PD, Arkin IT et al (1995) Computational searching and mutagenesis suggest a structure for the pentameric transmembrane domain of phospholamban. *Nat Struct Biol* 2(2):154–162
- Adams PD, Engelman DM et al (1996) Improved prediction for the structure of the dimeric transmembrane domain of glycoporphin A obtained through global searching [published erratum appears in *Proteins* 1997 Jan;27(1):132]. *Proteins* 26:257–261
- Akula N, Pattabiraman N (2005) A systematic search method for the identification of tightly packed transmembrane parallel alpha-helices. *J Biomol Struct Dyn* 22(6):625–634
- Arselin G, Giraud MF (2003) The GxxxG motif of the transmembrane domain of subunit e is involved in the dimerization/oligomerization of the yeast ATP synthase complex in the mitochondrial membrane. *Eur J Biochem* 270(8):1875–1884
- Beevers AJ, Kukol A (2006) Systematic molecular dynamics searching in a lipid bilayer: application to the glycoporphin A and oncogenic ErbB-2 transmembrane domains. *J Mol Graph Model* 25:226–233
- Berman HM, Battistuz T et al (2002) The Protein Data Bank. *Acta Crystallogr D Biol Crystallogr* 58(Pt 6 No 1):899–907
- Beuming T, Weinstein H (2005) Modeling membrane proteins based on low-resolution electron microscopy maps: a template for the TM domains of the oxalate transporter OxIT. *Protein Eng Des Sel* 18(3):119–125
- Bowie JU (1997) Helix packing in membrane proteins. *J Mol Biol* 272:780–789
- Braun R, Engelman DM et al (2004) Molecular dynamics simulations of micelle formation around dimeric glycoporphin A transmembrane helices. *Biophys J* 87(2):754–763
- Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) CHARMM: a program for macromolecular energy minimization, and dynamics calculations. *J Comp Chem* 4:187–217
- Bu L, Im W et al (2007) Membrane assembly of simple helix homooligomers studied via molecular dynamics simulations. *Biophys J* 92(3):854–863
- Bywater RP, Thomas D et al (2001) A sequence and structural study of transmembrane helices. *J Comput Aided Mol Des* 15(6):533–552

- Caballero-Herrera A, Nilsson L (2003) Molecular dynamics simulations of the E1/E2 transmembrane domain of the Semliki Forest virus. *Biophys J* 85(6):3646–3658
- Chen CP, Kernytsky A et al (2002) Transmembrane helix predictions revisited. *Protein Sci* 11(12):2774–2791
- Chothia C, Levitt M et al (1981) Helix to helix packing in proteins. *J Mol Biol* 145:215–250
- Citri A, Skaria KB et al (2003) The deaf and the dumb: the biology of ErbB-2 and ErbB-3. *Exp Cell Res* 284(1):54–65
- Coussens L, Yang-Feng TL et al (1985) Tyrosine kinase receptor with extensive homology to EGF receptor shares chromosomal location with neu oncogene. *Science* 230(4730):1132–1139
- Cuthbertson JM, Bond PJ et al (2006) Transmembrane helix–helix interactions: comparative simulations of the glycoporphin A dimer. *Biochemistry* 45(48):14298–14310
- Dawson JP, Weinger JS et al (2002) Motifs of serine and threonine can drive association of transmembrane helices. *J Mol Biol* 316(3):799–805
- Duneau JP, Crouzy S, Chapron Y, Genest M (1999a) Dynamics of the transmembrane domain of the ErbB-2 receptor. *Theor Chem Acc* 101(1–3):87–91
- Duneau JP, Crouzy S et al (1999b) Molecular dynamics simulations of the ErbB-2 transmembrane domain within an explicit membrane environment: comparison with vacuum simulations. *Biophys Chem* 76(1):35–53
- Duneau JP, Vegh AP et al (2007) A dimerization hierarchy in the transmembrane domains of the HER receptor family. *Biochemistry* 46(7):2010–2019
- Ehsani A, Low J et al (1993) Characterization of a new allele of the human ERBB2 gene by allele-specific competition hybridization. *Genomics* 15(2):426–429
- Eilers M, Shekar SC et al (2000) Internal packing of helical membrane proteins. *Proc Natl Acad Sci USA* 97(11):5796–5801
- Fleishman SJ, Schlessinger J et al (2002) A putative molecular-activation switch in the transmembrane domain of erbB2. *Proc Natl Acad Sci USA* 99(25):15937–15940
- Fleishman SJ, Harrington S et al (2004) An automatic method for predicting transmembrane protein structures using cryo-EM and evolutionary data. *Biophys J* 87(5):3448–3459
- Fleishman SJ, Ben-Tal N (2006a) Progress in structure prediction of alpha-helical membrane proteins. *Curr Opin Struct Biol* 16(4):496–504
- Fleishman SJ, Unger VM et al (2006b) Transmembrane protein structures without X-rays. *Trends Biochem Sci* 31(2):106–113
- Garnier N, Genest D et al (1997) Molecular modeling of c-erbB2 receptor dimerization: coiled-coil structure of wild and oncogenic transmembrane domains. Stabilization by interhelical hydrogen bonds in the oncogenic form. *Biopolymers* 42(2):157–168
- Gerber D, Sal-Man N et al (2004) Two motifs within a transmembrane domain, one for homodimerization and the other for heterodimerization. *J Biol Chem* 279(20):21177–21182
- Gimpelev M, Forrest LR et al (2004) Helical packing patterns in membrane and soluble proteins. *Biophys J* 87(6):4075–4086
- Graus-Porta D, Beerli RR et al (1997) ErbB-2, the preferred heterodimerization partner of all ErbB receptors, is a mediator of lateral signaling. *Embo J* 16(7):1647–1655
- Herzyk P, Hubbard RE (1995) Automated method for modeling seven-helix transmembrane receptors from experimental data. *Biophys J* 69(6):2419–2442
- Jayasinghe S, Hristova K et al (2001) Energetics, stability, and prediction of transmembrane helices. *J Mol Biol* 312(5):927–934
- Kallblad P, Dean PM (2004) Backbone-backbone geometry of tertiary contacts between alpha-helices. *Proteins* 56(4):693–703
- Killian JA (2003) Synthetic peptides as models for intrinsic membrane proteins. *FEBS Lett* 555(1):134–138
- Killian JA, von Heijne G (2000) How proteins adapt to a membrane-water interface. *T.I.B.S. Trends Biochem Sci* 25:429–434
- Kim S, Chamberlain AK et al (2003) A simple method for modeling transmembrane helix oligomers. *J Mol Biol* 329(4):831–840
- Kokubo H, Okamoto Y (2004) Prediction of membrane protein structures by replica-exchange Monte Carlo simulations: case of two helices. *J Chem Phys* 120(22):10837–10847
- Kraus MH, Issing W et al (1989) Isolation and characterization of ERBB3, a third member of the ERBB/epidermal growth factor receptor family: evidence for overexpression in a subset of human mammary tumors. *Proc Natl Acad Sci USA* 86(23):9193–9197
- Krogh A, Larsson B et al (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305(3):567–580
- MacKenzie KR, Engelman DM (1998) Structure-based prediction of the stability of transmembrane helix–helix interactions: the sequence dependence of glycoporphin A dimerization. *Proc Natl Acad Sci USA* 95(7):3583–3590
- MacKenzie KR, Prestegard JH et al (1997) A transmembrane helix dimer: structure and implications. *Science* 276:131–133
- MacKerell JAD, Bashford D, Bellot M, Dunbrack RL, Evansek JD, Field MJ, Fisher S, Gao J, Guo H, Ha S, Joseph McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich B, Smith J, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Karplus M (1992) Self-consistent parametrization of biomolecules for molecular modelling and condensed phase simulations. *Biophys J* 61:A143
- McClain MS, Iwamoto H et al (2003) Essential role of a GXXXG motif for membrane channel formation by *Helicobacter pylori* vacuolating toxin. *J Biol Chem* 278(14):12101–12108
- Mendrola JM, Berger MB et al (2002) The single transmembrane domains of ErbB receptors self-associate in cell membranes. *J Biol Chem* 277(7):4704–4712
- Opella SJ, Marassi FM (2004) Structure determination of membrane proteins by NMR spectroscopy. *Chem Rev* 104(8):3587–3606
- Petrache HI, Grossfield A et al (2000) Modulation of Glycophorin A transmembrane helix interactions by lipid bilayers: molecular dynamics calculations. *J Mol Biol* 302:727–746
- Pinkas-Kramarski R, Soussan L et al (1996) Diversification of Neu differentiation factor and epidermal growth factor signaling by combinatorial receptor interactions. *EMBO J* 15:2452–2467
- Popot JL, Engelman DM (1990) Membrane protein folding and oligomerization: the two-stage model. *Biochemistry* 29:4031–4037
- Rubin I, Yarden Y (2001) The basic biology of HER2. *Ann Oncol* 12(Suppl 1):S3–S8
- Russ WP, Engelman DM (2000) The GxxxG motif: a framework for transmembrane helix–helix association. *J Mol Biol* 296(3):911–919
- Ryckaert JP, Ciccotti G et al (1977) Numerical integration of cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J Comput Phys* 23:327
- Sajot N, Genest M (2000) Structure prediction of the dimeric neu/ErbB-2 transmembrane domain from multi-nanosecond molecular dynamics simulations. *Eur Biophys J* 28(8):648–662
- Schlessinger J (2000) Cell signaling by receptor tyrosine kinases. *Cell* 103(2):211–225
- Senes A, Engel DE et al (2004) Folding of helical membrane proteins: the role of polar, GxxxG-like and proline motifs. *Curr Opin Struct Biol* 14(4):465–479
- Smith SO, Eilers M et al (2002) Implications of threonine hydrogen bonding in the glycoporphin A transmembrane helix dimer. *Biophys J* 82(5):2476–2486
- Tanner KG, Kyte J (1999) Dimerization of the extracellular domain of the receptor for epidermal growth factor containing the mem-

- brane-spanning segment in response to treatment with epidermal growth factor. *J Biol Chem* 274(50):35985–35990
- Tomita M, Furthmayr H et al (1978) Primary structure of human erythrocyte glycophorin A. Isolation and characterization of peptides and complete amino acid sequence. *Biochemistry* 17(22):4756–4770
- Torres J, Kukol A et al (2001) Mapping the energy surface of transmembrane helix–helix interactions. *Biophys J* 81(5):2681–2692
- Treutlein HR, Lemmon MA et al (1992) The glycophorin A transmembrane domain dimer: sequence-specific propensity for a right-handed supercoil of helices. *Biochemistry* 31:12726–12733
- Tzahar E, Waterman H et al (1996) A hierarchical network of interreceptor interactions determines signal transduction by Neu differentiation factor/neuregulin and epidermal growth factor. *Mol Cell Biol* 16(10):5276–5287
- White SH (2004) The progress of membrane protein structure determination. *Protein Sci* 13(7):1948–1949
- White SH, Wimley WC (1999) Membrane protein folding and stability: physical principles. *Annu Rev Biophys Biomol Struct* 28:319–365
- White SH, Ladokhin AS et al (2001) How membranes shape protein structure. *J Biol Chem* 276(35):32395–32398
- Yamamoto T, Ikawa S et al (1986) Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth factor receptor. *Nature* 319:230–234
- Zoete V, Meuwly M et al (2005) Study of the insulin dimerization: binding free energy calculations and per-residue free energy decomposition. *Proteins* 61(1):79–93

## Conclusion

Dans cet article, j'ai présenté la méthodologie de recherche conformationnelle appliquée à l'étude des hétérodimères HER2-HER3. La méthode est parfaitement validée puisqu'elle décrit avec succès la structure du dimère de la GpA (rmsd sur les atomes C $\alpha$  de 1.5 Å par rapport à la structure expérimentale (code PDB 1AFO) résolue par RMN par MacKenzie et collaborateurs (MacKenzie et al. 1997).

Afin d'évaluer la pertinence de mes modèles nous avons comparé les distances inter-hélices, de nos six structures GpA simulées, aux contraintes NOE utilisées par MacKenzie et collaborateurs (MacKenzie et al. 1997) pour obtenir les 20 modèles RMN (code 1AFO). Le **tableau 3.1** résume les pourcentages des distances inter-hélices respectées ou violées pour nos modèles. Seule la structure prédite (GpA<sub>R1</sub>) présente peu de violations de ces contraintes de distances NOE. Toutefois ces violations sont comprises entre 0.5 Å et 2.8 Å, alors que pour les autres structures ces violations dépassent souvent les 10 Å. Ces distances sont assez similaires à celles des modèles de Smith et collaborateurs (Smith et al. 2002; Smith et al. 2001). Ces valeurs sont assez correctes dans la mesure où nos simulations sont effectuées dans le vide.

Structure	Pourcentage des contraintes NOE	
	respectées	avec violation $\leq \pm 0.5\text{Å}$
GpA <sub>R1</sub>	66	8
GpA <sub>R2</sub>	16	8
GpA <sub>R3</sub>	0	0
GpA <sub>R4</sub>	0	24
GpA <sub>L1</sub>	0	8
GpA <sub>L2</sub>	0	8

**Tableau 3.1** : Tableau récapitulatif des pourcentages des distances inter-hélices respectées ou violées pour les 6 structures GpA simulées. Ces distances correspondent aux 12 contraintes NOE utilisées pour prédire les modèles RMN (1AFO).

Les modèles représentatifs du domaine TM de l'hétérodimère HER2-HER3 sont ceux pour lesquels on retrouve les motifs de dimérisation à l'interface.

**2<sup>e</sup> PARTIE : ASSOCIATION HETERODIMERIQUE DES HELICES  
TRANSMEMBRANAIRES DES RECEPTEURS HER : ETUDE PAR SIMULATION  
DE DYNAMIQUE MOLECULAIRE EN L'ABSENCE D'ENVIRONNEMENT  
MEMBRANAIRE.**



## **Introduction**

Les 4 récepteurs ErbB sont soumis à un réseau d'interactions très complexe. Des études expérimentales montrent que les interactions récepteur-récepteur obéissent à une certaine sélectivité qui définit une hiérarchie des interactions et également la compétition entre récepteurs.

Le domaine TM des 4 récepteurs ErbB participe aux interactions récepteur-récepteur et pourrait être un des moteurs de la hiérarchie des interactions (Duneau et al. 2007). La connaissance des modes d'interaction des hélices TM des différents hétérodimères pourrait aider à comprendre le rôle de ce domaine dans la hiérarchie des interactions des récepteurs entiers. Cette étude est développée dans la partie suivante et est présentée sous forme d'article.

*Article*

**Insight into the recognition patterns of the ErbB receptor family transmembrane domains. Heterodimerization models through MD search.**

Oumarou Samna Soumana, Norbert Garnier and Monique Genest\*

Centre de Biophysique Moléculaire, UPR 4301, CNRS, rue Charles Sadron, 45071 Orléans Cedex 02, affiliated to the University of Orléans, France.

Keywords: ErbB receptor family, TM domain, heterodimerization, motifs, MD computational search

\* to whom correspondence should be addressed.

Tel : 33 (0)2 38 25 55 68

Fax : 33 (0)2 38 63 15 17

Email : [m.genest@cnrs-orleans.fr](mailto:m.genest@cnrs-orleans.fr)

## Summary

ErbB receptors undergo complex interaction network defining hierarchical and competition relationships. Dimerization motifs in the transmembrane domain of these receptors are supposed to play a role in modulating the specificity and the selection of the partners during signal transduction. We investigate structural models of ErbB transmembrane heterodimers by using MD conformational search method to give insights into the hierarchy of receptor interactions and to clarify the role of dimerization motifs. We show that left-handed structures of the TM heterodimers are preponderant over right-handed structures. All heterotypic heterodimers undergo two modes of association either *via* the N-terminal motifs or *via* the C-terminal motif. The ErbB3 TM domain that impairs the second dimerization motif can associate with the other partners owing to the presence of the C-terminal Gly residues. Some discrepancies exist between the hierarchy of interactions defined experimentally and from our theoretical models. Nevertheless the TM association of ErbB2 with ErbB3 and with ErbB1 is found favorable placing these heterodimers at the first ranks of the scale of affinity. A molecular-switch model allowing the transition between the two dimerization states may apply to the heterodimers and could help interpreting receptor competition for the formation of homodimers and heterodimers.

## **Introduction**

The epidermal growth factor receptor (EGFR, ErbB1) belongs to a subclass of four receptors including ErbB2 in human and Neu in rodents, ErbB3 and ErbB4 (Ferguson et al. 2003; Yarden and Sliwkowski 2001). The ErbB receptors tyrosine kinase (RTK) are involved in a wide range of differentiation and growth events. They are involved in the progression of carcinoma types (Citri and Yarden 2006; Salomon et al. 1995; Yarden and Sliwkowski 2001). ErbB receptors are transmembrane (TM) glycoproteins that have three different domains: a variable extracellular ligand-binding domain at the N-terminal with two cysteine-rich sequences, a transmembrane spanning domain connected to the intracellular domain that contains a tyrosine kinase domain flanked by a large hydrophilic tail (Olayioye et al. 2000).

Ligand binding to the extracellular domain enhances the dimerization of the receptors or stabilizes existing but unstable dimers (Lemmon et al. 1997; Sako et al. 2000) and results in the cross-phosphorylation of the dimerized receptors (Olayioye et al. 2000; Schlessinger 2000). Pronounced differences in the mode of ligand binding have been revealed leading to homo oligomers and hetero oligomers in various combinations involved in signalization across cell membrane (Linggi and Carpenter 2006; Warren and Landgraf 2006). All possible complexes between the receptors have been identified in different systems (Gullick 2001; Olayioye et al. 2000; Schlessinger 2000).

Relevant protein-protein interactions are beginning to be understood and significant insights into the molecular mechanisms of ErbB dimerization comes from the high-resolution structures of soluble ectodomains (Bouyain et al. 2005; Cho and Leahy 2002; Ferguson et al. 2003; Garrett et al. 2003; Garrett et al. 2002; Ogiso et al. 2002). Ectodomain contacts between the two monomers occur but their spatial proximity alone is not sufficient for receptor dimerization and efficient phosphorylation (Ferguson et al. 2003; Mattoon et al. 2004). Activation can be achieved without extracellular domains (Bargmann and Weinberg 1988; Di Fiore et al. 1987) and dimers can form in the presence or the absence of ligand (Moriki et al. 2001) suggesting that the basic ligand-induced dimerization mechanism is more complex as initially thought.

A series of findings indicate that regions other than the ectodomain contribute significantly to interactions. Contact regions exist as those in the kinase domain through two lobes termed N and C lobes (Zhang et al. 2006). The cytojuxtamembrane region (JM) of each ErbB receptor

contains basic residues that are reported to influence dimerization (Aifa et al. 2005; Aifa et al. 2006). Dimerization may also involve direct contacts between the helical TM structures of the receptors. The potential implication of the TM domain is illustrated by the Neu mutation in the ErbB2 that provokes constitutive dimerization and kinase activation (Bargmann et al. 1986; Bargmann and Weinberg 1988, 1988). Close association between these domains was thought to be permitted by a dimerization motif originally described by Sternberg and Gullick (Sternberg and Gullick 1990). This motif is the GxxxG sequence (x is any residue), where Gly can be replaced by a small residue (Ala, Ser or Thr) (Dawson et al. 2003; Dawson et al. 2002; Eilers et al. 2002; Kleiger et al. 2002; Senes et al. 2000).

All ErbB receptors have a single TM domain that exhibits a conserved and very common GxxxG-like motif at both the N-terminal and the C-terminal ends, except ErbB3 for which this motif is present only at the N-terminus. The degree of self-association of the ErbB TM spans correlate with the presence of this motif (Mendrola et al. 2002). It is suggested that the Neu mutation (V659E) when introduced into ErbB2 induces a shift between two packing modes which may represent active and inactive configurations of the receptor (Fleishman et al. 2002). In the case of the heterodimer ErbB1-ErbB2, the two motifs could exert a separate function in TM domain assembly, one for homo dimerization and the other for hetero dimerization (Gerber et al. 2004). Helix-helix interactions are weak in detergent micelles (Stanley and Fleming 2005) but exist and require stabilizing interactions by other parts of the receptor. The association of the receptors comes from the sum of many interactions including weak contacts distributed throughout all the domains. In addition, ligand binding also induces a rearrangement of each receptor subunit owing to the rotation of the TM domains (Bell et al. 2000; Moriki et al. 2001).

Receptor-receptor interactions display distinct selectivity that defines hierarchical and competitive relationships (Karunakaran et al. 1996; Prigent and Gullick 1994). Biochemical experiments undertaken on the four human ErbB receptors (Tzahar et al. 1997) revealed that from the six possible heterotypic interactions within the family, three shows relatively strong inter receptor interactions. ErbB2 acts as a superior partner of the ligand-bound ErbB1, ErbB3, ErbB4 and its major partner is the kinase defective member of the family ErbB3. Heterodimers between ErbB3 and ErbB4 are not detectable. These results agree on many aspects with the study undertaken by Riese et al. (Riese et al. 1995). A correlation between the TM domain ( $HER_{TM}$ ) and the whole receptor association preferences has been recently demonstrated using FRET (Fluorescence Resonance Energy Transfer) (Duneau et al. 2007).  $HER_{TM}^2$  was found the preferential partner of other receptors with a marked preference for

HER<sup>3</sup><sub>TM</sub> and HER<sup>1</sup><sub>TM</sub>. The association of HER<sup>4</sup><sub>TM</sub> with HER<sup>2</sup><sub>TM</sub> is less favorable. Other heterodimers involving HER<sup>1</sup><sub>TM</sub> are observed and HER<sup>3-4</sup><sub>TM</sub> species are detected.

Thus, the TM domain of the ErbB receptors is presented as a selector of the population of dimer species, and is proposed to play a major role in modulating the affinity, specificity, and selection of partners by receptors during signal transduction.

Even if the TM domain does not provide a driving force to stabilize the receptor dimers it is important to give light on how the TM domains associate in the context of heterodimer formation. The details of the interactions between the TM domains and the knowledge of their structure could help to understand the role of the juxtamembrane segment, at the junction of the TM domain and the N-lobe of the kinase domain. This lobe is believed to maintain a certain flexibility allowing switch positions to activate each receptor.

To address this question of inter receptor interactions at the level of the TM domain we have analyzed the six possible heterotypic interactions. Global search methods combined with molecular dynamic (MD) simulations are used to identify possible structures of the heterodimers.

MD search protocol, recently described in (Samna Soumana et al. 2007) successfully predicts the right-handed structure of the TM dimer of glycophorin A (MacKenzie et al. 1997). Models of the HER<sup>2-3</sup><sub>TM</sub> heterodimer have been presented showing that right-handed and left-handed coiled coil structures conform to the presence of the dimerization motifs at the dimer interface. The present study is extended to all heterotypic interactions and clearly shows a net preference for a left-handed structure of TM helix pairs over a right-handed structure for all heterodimers. We show that the most representative MD models strongly correlate with the presence of either one of the two dimerization motifs at the contacting interface. Existence of two dimerization modes for the ErbB TM domains hypothesized to be involved in the complex interactions network of the receptors is discussed in relation to experimental data. A hierarchy of TM interactions is proposed and compared to the dimerization hierarchy of the whole receptors and of the TM domains defined experimentally (Duneau et al. 2007; Tzahar et al. 1996).

## **Methods**

### **TM domain systems**

The six helical heterodimers within the ErbB family are studied. Heterodimers are referred to as  $HER^{i-j}_{TM}$  ( $i, j = 1-4, i \neq j$ ) in the following. Each monomer includes the 25 residues of the hydrophobic sequence of the TM domains of each receptor ErbB1: (Ullrich et al. 1984), ErbB2:(Coussens et al. 1985; Yamamoto et al. 1986), ErbB3: (Kraus et al. 1989), ErbB4: (Plowman et al. 1993). The sequences are given in Table 1. Each heterodimer comprises two helices, each built with the amino terminus acetylated and the carboxyl terminus methylaminated to mimic the peptide bond.

**Table 1**

TM domain	Sequence	Residues
HER1	PSIAT <u><b>GMV</b></u> GALLLLLVVALGIGLFM	644-668
HER2	LT <u><b>SIVS</b></u> AVVGILLVVVLGVVFGILI	651-675
HER3	L <u><b>TMALT</b></u> VIAGLVVIFMMLGGTFLYW	642-666
HER4	PLIAAG <u><b>VIGGL</b></u> FILVIVGLTFAVYV	651-675

**Table 1:** Sequences of the TM domain of the four ErbB receptors: HER<sup>1</sup><sub>TM</sub> (Ullrich et al. 1984), HER<sup>2</sup><sub>TM</sub> (Coussens et al. 1985; Yamamoto et al. 1986), HER<sup>3</sup><sub>TM</sub> (Kraus et al. 1989) and HER<sup>4</sup><sub>TM</sub> (Plowman et al. 1993). The regions corresponding to the putative GxxxG dimerization motif in each TM domain are shaded gray. SmxxxSm interactions sites are in bold or underlined (see text).



## Computational search

The program CHARMM Version 26 (Brooks 1983) and the all hydrogen PARAM22 force field (MacKerell et al. 1998) were used in the present study. All calculations were carried out in vacuo and performed on a multi-processor platform Linux Beowulf cluster with dual 3.06 GHz Intel Xeon processors.

Conformational sampling and molecular dynamics protocol are detailed in a previous paper (Samna Soumana et al. 2007). The main steps of the MD search are briefly summarized in the following and in the flowchart in Figure 1.

Parallel dimers were first generated with two helices (H1 and H2) built with the standard torsion angles. For a complete search over the two-body rotational interaction space, helices were rotated about their own axis by  $10^\circ$  steps for a complete rotation. A total of 1296 ( $36 \times 36$ ) initial configurations are generated for each heterodimer. The two helices were first separated by a distance  $D_h$  between the helix axes corresponding to the minimum distance for which they are in favorable interactions. This is obtained by combining energy minimization and progressive approach of the two helices maintained in a parallel orientation. The 1296 initial configurations were submitted to this first step before starting MD simulations.

MD simulations were performed at constant temperature ( $T = 300\text{K}$ ). Thermalization was performed by increments of 50K by maintaining the two helices close to the  $\alpha$  structure while restricting a vertical shift of one helix relatively to the second helix. Backbone hydrogen bonds ( $\alpha$  Hbonds) and crossed distances restraints (see Table 2 for definition) were applied and the force constant was gradually reduced in the course of the equilibration period.

The production period was continued for 3 ns without any restraint. Atomic coordinates and energy terms were saved every picosecond (ps). The last nanosecond (ns) of the trajectory served for averaging and analyses. First, the quality of the 1296 structures calculated for each type of dimer was evaluated by using geometric filters. The six geometric parameters used are given in Table 2. A structure is retained when at least 80% of the 250 configurations are structurally correct (all the parameters are respected). The remaining structures form three families according to the crossing angle value (left-handed and right-handed structures and dimers with parallel helices).  $C\alpha$  RMSD clustering was applied to each family ( $\text{RMSD} \leq 2 \text{ \AA}$ )

defining clusters of similar structures. Energy minimization and MD refinement was systematically applied to the average structures representative of each cluster.

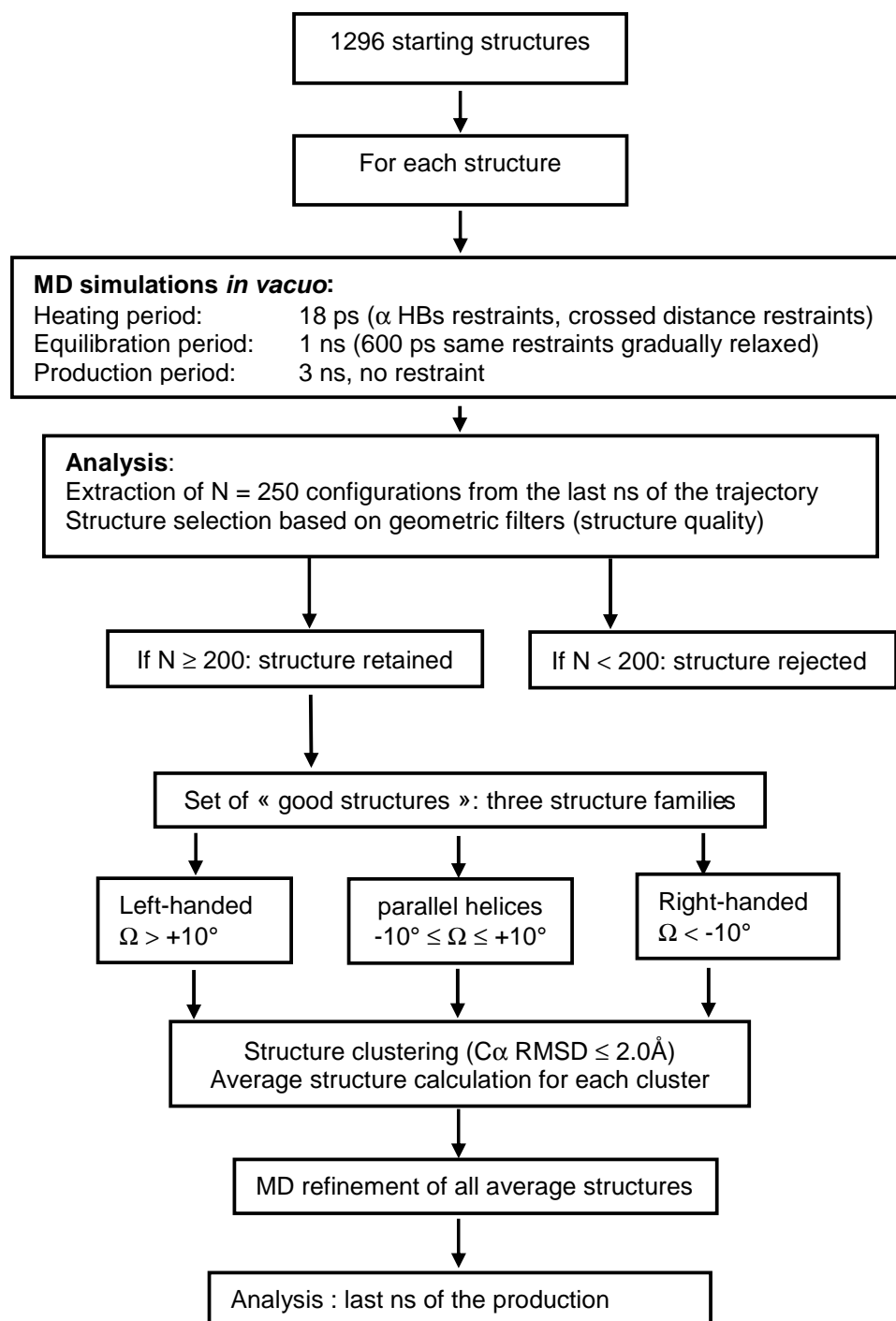
The structures are analyzed in terms of energy and interface surface area (SA) and average values are calculated over the last ns of the trajectory. In addition, the gas-phase contribution to the binding free energy  $\Delta G^{\circ}_{\text{bind}}$  was calculated (Zoete et al. 2005).  $\Delta G^{\circ}_{\text{bind}}$  is an indicator of the affinity of the two monomers.

**Table 2**

Parameters	$L_{H\alpha}$ (Å)		$\Delta dc$ (Å)	$\Delta Z_{Nter}$	$\Delta Z_{Cter}$ (Å)	$\Omega$ (°)	
	min	max				min	max
HER <sub>TM</sub>	32.0	38.0	$\leq 3.5$		$\leq 3.5$	-65°	+45°

**Table 2:** Parameters defining the dimer structure quality. (Samna Soumana et al. 2007).  $L_H$  is the length of a  $\alpha$  canonical helix ( $L_{H\alpha} = 35.25\text{Å}$  for a 25 residue helix).  $L_{H\alpha}$  is the distance between the  $C_\alpha$  atoms of the first residue  $C_\alpha^{H1}Nter$  and the last residue  $C_\alpha^{H2}Cter$ .  $\Delta dc$ : difference between the crossed distances ( $C_\alpha^{H1}Nter-C_\alpha^{H2}Cter$  and  $C_\alpha^{H2}Nter-C_\alpha^{H1}Cter$ ). Crossed distance: distance between the  $C_\alpha$  atoms of the first residue  $C_\alpha^{H1}Nter$  of one helix (H1 or H2) and the  $C_\alpha$  atoms of the last residue  $C_\alpha^{H2}Cter$  of the second helix (H2 or H1).  $\Delta Z(Nter) = |Z(C_\alpha^{H1}Nter) - Z(C_\alpha^{H2}Nter)|$ : difference in Z positions of the  $C_\alpha$  atoms of the first residue of H1 and H2. Same definition for  $\Delta Z(Cter)$  but for the  $C_\alpha$  atoms of the last residue of H1 and H2.  $\Omega$  (°) is the crossing angle calculated following the method of Chothia (Chothia et al. 1981) included in CHARMM. Crossing angle values are from Bowie (Bowie 1997) and statistical analysis of membrane proteins.

**Figure 1**



**Figure 1:** Flowchart giving an overview of the main steps of the search protocol and MD simulations applied to the prediction of heterodimerization of the TM domain of the ErbB receptor family. Details of the method are given in our previous work (Samna Soumana et al. 2007).

## Results

### Global analysis of the $HER^{i-j}_{TM}$ models

From the 1296 sampled structures of each of the different heterodimers, three families representative of the right-handed, left-handed and parallel dimers were obtained. The number of the selected structures for the different families is given in Table 3. It is clearly shown a net preference for dimers with helices in left-handed interactions. Very few right-handed structures result from geometric filtering selection except for the two heterodimers  $HER^{2-1}_{TM}$  and  $HER^{2-4}_{TM}$ . Helices in parallel interactions (absolute value of the crossing angle lesser than  $10^\circ$ ) are largely represented for the heterodimers formed by  $HER^4_{TM}$  with  $HER^2_{TM}$  or  $HER^3_{TM}$ . After the last step including structure clustering and MD refinement, the resulting structures, named “good“ structures, represent a percentage ranging from 5% to 13% of the total structures.

The set of “good” structures is characterized by the distribution of the crossing angles as shown in Figure 2. Helix-helix angles are centered around  $25^\circ$ - $35^\circ$  for the left-handed structures and for the few structures defining right-handed coiled coils the crossing angle is between  $-30^\circ$  and  $-10^\circ$ . The two peaks around  $\pm 5^\circ$  correspond to the large population of the dimers with parallel helices found for the heterodimers  $HER^{2-4}_{TM}$  and  $HER^{3-4}_{TM}$ .

An overview of dimer energy of all the selected MD models indicates that, in average, the TM heterodimers  $HER^{1-3}$  are the most stable.  $HER^{2-3}_{TM}$  and  $HER^{3-4}_{TM}$  models are less stable in average than those formed by  $HER^1_{TM}$  and  $HER^3_{TM}$  but are more stable than the three others. The  $HER^2$ - $HER^4$  TM association is the less favorable as revealed by the stability of the dimers which is half that calculated for the TM helix pair  $HER^{1-3}$ .

The magnitude of helix-helix interactions could be considered as indicator for a preferential partner. In average, interactions between the TM domains are in the same order of magnitude for the whole set of MD models ranging from -84 kcal/mol for the strongest interactions to -42 kcal/mol for the weakest interactions. For all  $HER_{TM}$  heterodimers the van der Waals interactions are dominant and represent 70% or more of the total energy calculated between the two helices. Electrostatic interactions are negative and greatly vary depending on the nature of the residues at the interface. These interactions strongly modulate the association of TM helix pairs.

Another important parameter for evaluating the affinity of the two monomers is the value of the binding free energy  $\Delta G^{\circ}_{\text{bind}}$ . In average, if we consider all the MD selected models, the three TM heterodimers, namely  $\text{HER}^{2-3}_{\text{TM}}$ ,  $\text{HER}^{1-3}_{\text{TM}}$  and  $\text{HER}^{3-4}_{\text{TM}}$ , are those of best affinity. The three others,  $\text{HER}^{2-4}_{\text{TM}}$ ,  $\text{HER}^{1-4}_{\text{TM}}$  and  $\text{HER}^{2-1}_{\text{TM}}$  occupy the three last ranks of the interactions scale.

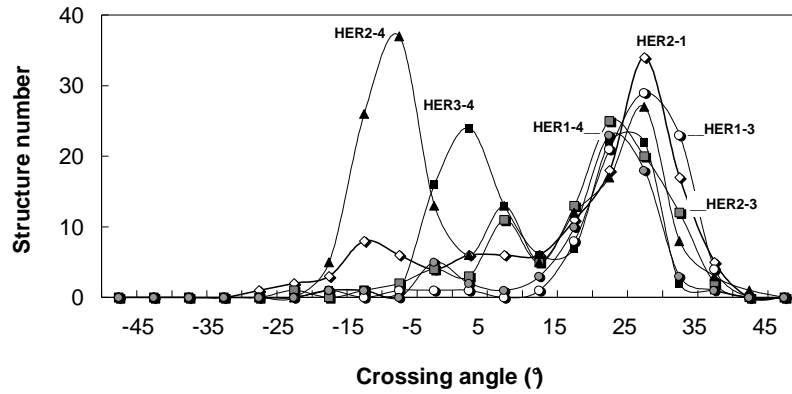
Contacting surface area between the two helices informs on the way the two helices interact. For the whole set of correct structures defined for the 6 types of heterodimers, the contacting surface area varies from  $640 \text{ \AA}^2$  ( $\text{HER}^{2-4}_{\text{TM}}$  and  $\text{HER}^{1-4}_{\text{TM}}$ ) and  $685 \text{ \AA}^2$  ( $\text{HER}^{3-4}_{\text{TM}}$  and  $\text{HER}^{1-3}_{\text{TM}}$ ). Models of  $\text{HER}^{2-1}_{\text{TM}}$  and  $\text{HER}^{2-3}_{\text{TM}}$  exhibit, in average, a contacting surface of  $655 \text{ \AA}^2$  and  $669 \text{ \AA}^2$  respectively.

**Table 3**

HER <sub>TM</sub> heterodimer	Structure family	HER <sup>2-1</sup>	HER <sup>2-3</sup>	HER <sup>2-4</sup>	HER <sup>3-4</sup>	HER <sup>1-3</sup>	HER <sup>1-4</sup>
Nb of selected structures after step 1	L	91	77	73	60	86	58
	P	22	20	69	54	3	8
	R	14	2	31	2	1	2
% “good” structures		10	8	13	9	7	5
after step 2	L	11	5	5	9	9	5
	P	2	1	3	1	0	0
	R	1	1	0	1	1	1

**Table 3:** Summary of the simulations. Selected structures for the different HER<sub>TM</sub> heterodimers after geometric filters (step1) and after structure clustering and final MD refinement (step 2). Three families of dimer models are found: left-handed coiled coil (L); right-handed coiled coil (R) and dimers with parallel helices (P). Percentage of “good” structures is calculated over the 1296 initial configurations simulated for each heterodimer.

**Figure 2**



**Figure 2:** Distribution of the crossing angles  $\Omega$  ( $^\circ$ ) calculated for the  $\text{HER}_{\text{TM}}$  heterodimers models resulting from geometric filters. Left-handed structures are dominant with  $\Omega$  ranging between  $20^\circ$  and  $35^\circ$ .  $\text{HER}_{\text{TM}}^{2-4}$  and  $\text{HER}_{\text{TM}}^{3-4}$  heterodimers are characterized by a large number of parallel dimers  $\Omega = \pm 5^\circ$ . Very few right-handed structures are formed between the TM domains.



## **GxxxG-like motifs mediate HER<sub>TM</sub> association**

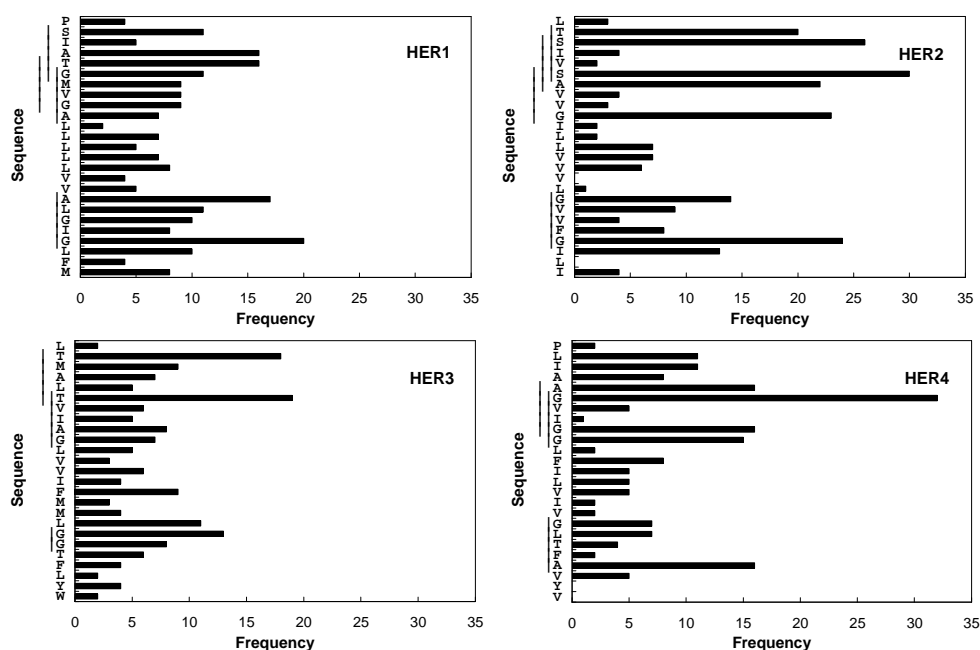
The presence of SmxxxSm motifs (Sm is a small residue Gly, Ala, Ser or Thr and x is any residue) in all the TM domains of the receptors motivates an analysis of their position at the dimer interface. For that, we have undertaken the analysis of the residues implicated in short contacts (C $\alpha$ -C $\alpha$  distance between the TM monomers  $\leq 6$  Å). This analysis uses all the “good” structures obtained after the last selection step. The results are shown in Figure 3.

The plots clearly evidence that the residues with a high occurrence at the dimer interface are the residues of the GxxxG motif or more largely those of the SmxxxSm motif. In the case of HER<sup>1</sup><sub>TM</sub>, Ser645, Ala647, Thr648, Gly649 at the N-terminus and Ala661 and Gly665 at the C-terminus are the most frequent at the dimer interface. For HER<sup>2</sup><sub>TM</sub>, all the small residues, particularly the polar residues Ser and Thr at the N-terminus have a high frequency of occurrence similarly to the Gly residues of the C-terminus. The same features are observed for HER<sup>4</sup><sub>TM</sub> with a strong implication of Gly and Ala at the interface. In the case of HER<sup>3</sup><sub>TM</sub>, the polar residues Thr643 and Thr647 included in the N-terminal motifs are also often present at dimer interfaces. Despite the fact that HER<sup>3</sup><sub>TM</sub> is the only domain of the receptor family without putative dimerization motif at the C-terminal side, the two Gly residues 660-661 localized at the same level than the second dimerization motif of the other HER<sub>TM</sub> domains, frequently occur in close proximity with the facing helices.

The predominance of the small residues of the TM domains at helix-helix interface is recognized as a main property of helices in membrane proteins (Adamian and Liang 2001; Eilers et al. 2000; Gratkowski et al. 2001; Javadpour et al. 1999; Walters and DeGrado 2006; Zhou et al. 2001).

GxxxG-like motifs are mediators of helix-helix association. The fact that we are able to demonstrate that ErbB TM domain interactions are driven by such motifs makes reliable the search of representative structural models of HER<sub>TM</sub> heterodimers.

**Figure 3**



**Figure 3: Statistics on inter helix-helix distances**

Number of times the residue  $R_k$  of the monomer  $HER_{TM}^i$  ( $i=1-4$ ) is close to any residue  $R_l$  of the second monomer  $HER_{TM}^j$  ( $j=1-4, j \neq i$ ): distance  $C_\alpha R_k - C_\alpha R_l \leq 6 \text{ \AA}$ . The analyses use the structures of all the heterodimers  $HER_{TM}^{i-j}$  (6 types) resulting from the final selection (see Table 1). The number of residue pairs examined is 223 for  $HER^1$ , 238 for  $HER^2$ , 170 for  $HER^3$  and 187 for  $HER^4$ . The  $G_{xxx}G$ -like motif of the  $HER_{TM}$  sequences and the extended motifs  $Sm_{xxx}Sm$  are underlined.  $HER_{TM}^3$  exhibits only one  $G_{xxx}G$  motif at the N-terminal side. The two Gly residues at a position similar to that of the C-terminal motif found in all others sequences are highlighted.

## Structures and characteristics of the heterodimers

All the  $HER^{i,j}_{TM}$  heterodimers models will not be detailed in the following. For sake of clarity, only the structures having the best interest in term of dimer stability or in term of helix-helix interactions, with a net implication of the N-terminal or the C-terminal dimerization motif are given. In addition, the models detailed below are those for which the structure of the two helices is the closest to the regular  $\alpha$  structure. The energetic and geometrical characteristics of these dimer models are collected in Table 4.

Hydrogen bonds are not included in the data. HBs involve essentially the Thr and Ser side chains but their life time is generally very short and do not contribute to dimer stabilization.

### $HER^2_{TM}$ and its partners

Among the 14 structures retained for the  $HER^{2-1}_{TM}$  heterodimers, three structures are presented. One is a right-handed structure with the two helices crossed with a small angle. This structure is the most stable with helix-helix interactions mediated by the N-terminal motif. The shortest residue-residue distance is observed between the two facing Gly residues 660 and 652 of  $HER^1_{TM}$  and  $HER^2_{TM}$  respectively. Interactions between the two monomers are the most favorable of all other models because of strong electrostatic interactions (about 30% of the helix-helix interactions). It is to note that it is the only structure which exhibits a strong interhelical HB between the carbonyl group of Thr648 on  $HER^1_{TM}$  and the side chain of Ser656 on  $HER^2_{TM}$ . The second structure exhibits parallel helices in tight contacts all along the dimer interface giving thus the largest contacting surface of all the selected models. The polar residues Ser645 and Thr648 of  $HER^1_{TM}$  and Thr652 and Ser656 of  $HER^2_{TM}$  pertaining to the SmxxxSm motifs are in close proximities and Gly residues of the C-terminal motifs are buried within the dimer interface. For this model, the contribution of the C-terminal motif to the association is slightly stronger than that of the N-terminus. The Phe residues, with their side chain oriented towards the dimer interface increase the van der Waals interactions.

The left-handed structures constitute the main family of the  $HER^{2-1}_{TM}$  heterodimers. One structure with helices crossed at  $22^\circ$  (the crossing angle is between  $20^\circ$  and  $40^\circ$  for the set of left-handed structures) represents this family. The stability of this heterodimer and the magnitude of helix-helix interactions are very similar to the two others. The shortest residue-

residue distances are observed at the C-terminal extremity and, in average, the two motifs contribute weakly to the association.

For these HER<sup>2-1</sup><sub>TM</sub> MD models, the GxxxG-like motif, either at the N-terminus or at the C-terminus exerts favorable but weak inter molecular interactions. Calculated energy values for the “true” GxxxG motif do not exceed -5 kcal/mol. Motif-motif interactions results from the sum of dominant van de Waals interactions and very often of positive electrostatic interactions. This is a general characteristic of the energetic contributions of GxxxG dimerization motifs for all the heterodimers.

The HER<sup>2</sup>-HER<sup>3</sup> TM association leads to a number of “good” structures relatively low. MD models have been presented in our previous paper and only two of them (referred to as L1 and L4) (Samna Soumana et al. 2007) are recalled here. For most of the models the N-terminal motif is a mediator of the interactions. However, several models show that helix-helix contacts can be formed through the C-terminal side. This result is not expected owing to the absence of dimerization motif at this extremity for HER<sup>3</sup><sub>TM</sub>. As shown below, all heterodimers formed with HER<sup>3</sup><sub>TM</sub> show the presence of the residues Gly660 and Gly661 at the dimer interface.

One of the models presented exhibits the best stability with an energy value equal -175 kcal/mol in average. The two residues of the N-terminal motif, Ala657 and Thr647 respectively on HER<sup>2</sup><sub>TM</sub> and HER<sup>3</sup><sub>TM</sub>, are in a close approach. Helix-helix interactions and the binding free energy are the lowest of all the MD models. The other left-handed model is slightly less stable and shows short Gly-Gly inter helical distance at the C-terminus. Helix-helix interactions are much less favorable than for the previous structure. These two models are much more stable than the HER<sup>2-1</sup><sub>TM</sub> heterodimers and this is observed for most of the other selected models (data not shown). Figure 4 represents these two structures and highlights the dimer interface. It is to note that the two modes of association reveal a different orientation of the helices. This will be discussed in the following.

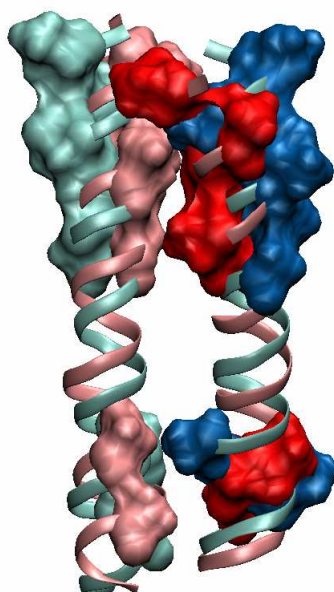
HER<sup>4</sup><sub>TM</sub> appears as very bad partner of HER<sup>2</sup><sub>TM</sub> as stated by the stability of the heterodimers and inter helical interactions that are slightly lower than those calculated for the other types of heterodimers. HER<sup>4</sup><sub>TM</sub> is the only TM span without polar residues at the N-terminus leading to low electrostatic interactions between the helices. Helix-helix association is essentially governed by the van der Waals interactions for this heterodimer. Much of the models exhibit

helices far from a regular helix (limit of the geometric criteria). Helix distortions are probably induced by the presence of bulky side chains (Phe, Tyr) which are aligned along the same helix face composed of small residues.

Two structures relatively different corresponding to the HER<sup>2</sup>-HER<sup>4</sup> TM association are detailed. One is a left-handed structure for which the crossing angle is the greatest (40°) and helix-helix contacts are the shortest. The second model is constituted by two parallel helices in weak interactions. The two Gly residues 668 and 672 on HER<sup>2</sup><sub>TM</sub> and Gly668 on HER<sup>4</sup><sub>TM</sub> favor tight contacts at the C-terminal side. For this parallel association, the bulky residues Phe and Tyr are oriented out of the interface and electrostatic interactions are null. For these two models, the distance between the two helix axes is short compared to the other types of the heterodimers.

In conclusion, HER<sup>2</sup><sub>TM</sub> associates very favorably with HER<sup>1</sup><sub>TM</sub> and HER<sup>3</sup><sub>TM</sub> forming stable dimers. The heterodimers formed with HER<sup>4</sup><sub>TM</sub> are much less stable. Helix-helix interactions are not correlated to the stability of the dimer but are correlated to the binding free energy.

**Figure 4**



**Figure 4:** HER<sup>2-3</sup><sub>TM</sub> heterodimer. The two structures described in Table 4 are superimposed (C $\alpha$  atoms of the backbone). Small residues including Gly, Ala, Ser and Thr are represented by their vdW surface. The association mediated by the N-terminal motif shows the contacting surface between the residues of HER<sup>2</sup><sub>TM</sub> (pink) and the residues of HER<sup>3</sup><sub>TM</sub> (red). For this association Gly of the C-terminal motif of HER<sup>2</sup><sub>TM</sub> and at the C-terminal side of HER<sup>3</sup><sub>TM</sub> are out of the interface. When helices associate *via* the C terminus, Gly residues are oriented within the interface (HER<sup>2</sup><sub>TM</sub>: light blue, HER<sup>3</sup><sub>TM</sub>: blue) and the small residues of the N-terminus are out of the interface.

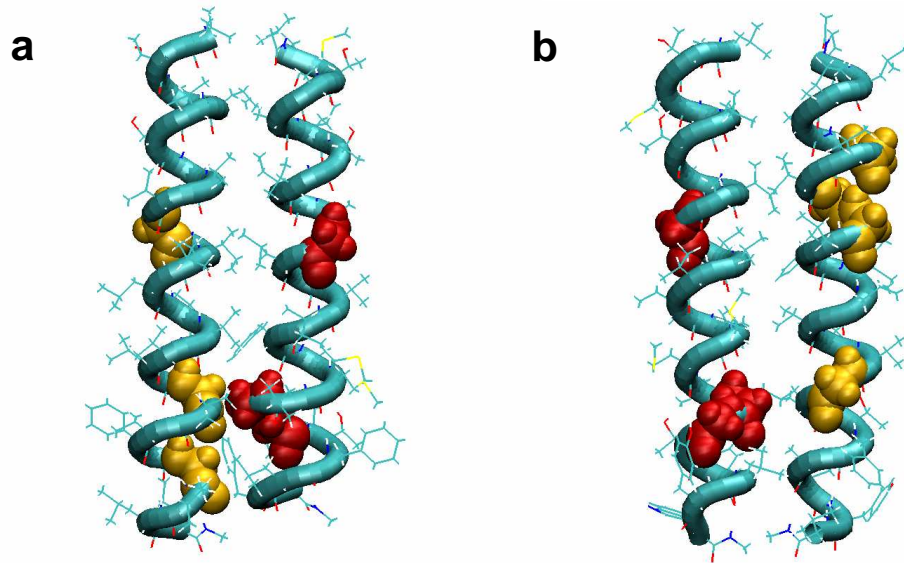
### **HER<sup>3</sup><sub>TM</sub> and its partners**

The TM domain of the ErbB3 receptor associates favorably with the other domains of the family. In terms of dimer stability, it is HER<sup>1</sup><sub>TM</sub> which associates the most favorably followed by HER<sup>2</sup><sub>TM</sub> and then by HER<sup>4</sup><sub>TM</sub>. The structures the most representative of the association with HER<sup>3</sup><sub>TM</sub> are left-handed coiled coils. When helix-helix association is mediated by the N-terminal motif, Thr647 is the most frequently involved at the dimer interface. At the C-terminal side, Gly660 or Gly661 are often involved in the interactions with the second helix although they are not included in a dimerization motif. This is shown in Figure 5. These two Gly residues are followed by a Thr residue the side chain of which participates, but rarely, to HB's interactions with residues of the opposite helix. For the three heterodimers formed with HER<sup>3</sup><sub>TM</sub>, helix-helix interactions are generally the strongest when residues of the N-terminal motif are in a close approach. In the case of HER<sup>1</sup>-HER<sup>3</sup> TM association, electrostatic interactions can be relatively strong compared to the van der Waals interactions (ratio 1:2). This is also observed for HER<sup>2-3</sup><sub>TM</sub> and HER<sup>3-4</sup><sub>TM</sub> heterodimers owing to the presence of residue Thr647 at the interface.

The shortest C $\alpha$ -C $\alpha$  distances within the interface are comprised between 4.3 and 4.8 Å while for all other models the closest residues are at the most separated by a distance of 5.5 Å. All the heterodimer models have their helices at a distance around 8.2 Å in average which is slightly shorter than found for left-handed coiled coils in membrane proteins.

Very negative values calculated for the binding free energy are correlated to strong helix-helix interactions. More the contacting surface is large, more stable is the dimer due to strong van der Waals interactions. Such a correlation between strong association and large contacting area is better evidenced for the left-handed structures. Obviously the magnitude of the interface area depends on the orientation of the two helices relatively to each other.

**Figure 5**



**Figure 5:** Gly residues at the dimer interface of the heterodimers formed with HER<sup>3</sup><sub>TM</sub>. a) HER<sup>2-3</sup><sub>TM</sub> heterodimer. b) HER<sup>3-4</sup><sub>TM</sub> heterodimer. The Gly residues are represented by van der Waals surfaces. Gly on HER<sup>3</sup><sub>TM</sub> are in red, Gly on HER<sup>2</sup><sub>TM</sub> and HER<sup>4</sup><sub>TM</sub> are colored yellow. Views of the helices from the N-terminus (top).



## **HER<sup>4</sup><sub>TM</sub> and its partners**

The behavior of two partners of HER<sup>4</sup><sub>TM</sub> has been described previously showing that the association with HER<sup>3</sup><sub>TM</sub> is very favorable while that with HER<sup>2</sup><sub>TM</sub> is not. In the case of the association with HER<sup>1</sup><sub>TM</sub> the heterodimers exhibit intermediate stability. The two structures described in Table 4 are left-handed coiled coils with helices almost regular. Helices are not vertically shifted and in both models the small Ala and Gly residues of either the N-terminal motif or the C-terminal motif are in contact at the interface. As noted before, the particularity of the TM domain HER<sup>4</sup> is the absence of polar residues at the N-terminal side that may explain weak helix-helix interactions at the level of the dimerization motif.

**TABLE 4**

	HER <sup>i-j</sup> <sub>TM</sub>	Motif	$\Omega$	$D_{Hi-Hj}$	$D_{C\alpha_i-C\alpha_j}$	$R_{Hi-RHj}$	SA	$E_{tot}$ dimer	$E_{inter}$	$\Delta G^{\circ}_{bind}$	$E_{inter}$ Motif	
											N ter	Cter
HER <sup>2</sup> <sub>TM</sub> partners	HER <sup>2-1</sup>	N ter	-12	8.4	4.3	G660-G652	632	-140	-66	-152	-5	-2
		N ter	0	8.6	4.2	T652-S645	741	-134	-76	-151	-1.2	-3
		C ter	22	8.1	4.7	G672-F667	721	-128	-69	-144	-1	-2
	*HER <sup>2-4</sup>	N ter	40	6.8	4.1	A657-G659	627	-94	-57	-149	-0.8	-7.4
		C-ter	-10	5.5	4.7	V665-V665	595	-81	-43	-136	-0.7	-2.6
	HER <sup>2-3</sup>	N ter	24	8.1	4.4	A657-T647	687	-175	-65	-161	-4.1	-4.1
C ter		13	8.7	4.8	G668-G660	590	-168	-42	-140	-1.7	-1.7	
HER <sup>3</sup> <sub>TM</sub> partners	HER <sup>1-3</sup>	N ter	23	8.1	4.7	M650-T647	709	-198	-69	-185	-7.1	-7.1
		C ter	21	8.2	4.8	L666-G661	672	-207	-62	-178	-3.8	-3.8
	*HER <sup>3-4</sup>	N ter	27	7.8	4.8	T647-G656	734	-151	-72	-180	-4.6	-4.6
		C ter	23	8.2	4.3	G660-G668	673	-138	-45	-139	-1.4	-1.4
	HER <sup>1-4</sup>	N ter	21	9.1	5.3	A647-A655	641	-108	-46	-141	-1.5	-1.9
		C ter	28	7.8	4.2	G665-A672	603	-102	-42	-129	-2.5	-4.9

**Table 4:** Characteristics of the HER<sup>i-j</sup><sub>TM</sub> heterodimers models.

All the selected structures are not detailed. Only an example of association *via* the N-terminal motif and the C-terminal motif is given for the different heterodimers (see text). Right-handed structures are characterized by a negative value of the crossing angle ( $\Omega$ ). A positive value characterizes a left-handed structure. Helix-helix angle is between  $-10^{\circ}$  and  $+10^{\circ}$  for parallel dimers. Crossing angle  $\Omega$  ( $^{\circ}$ ) is calculated following a method by Chothia (Chothia et al. 1981) included in CHARMM.  $D_{Hi-Hj}$  ( $\text{\AA}$ ): inter-helical distance between the two helices;  $D_{C\alpha_i-C\alpha_j}$  ( $\text{\AA}$ ): minimum distance between  $C\alpha$  atoms of opposed residues  $R_{Hi-RHj}$ . SA ( $\text{\AA}^2$ ): contacting surface area at the dimer interface. The total energy of the dimer ( $E_{tot}$  dimer) is the sum of the intramolecular energy of the two helices and helix-helix interactions  $E_{inter}$ .  $\Delta G^{\circ}_{bind}$  is mean value of the binding free energy in the gas-phase (see text for definition).  $E_{inter}$  motif is calculated between the dimerization motif GxxxG (see Figure 1) for the N-terminal motif and the C-terminal motif). Energy terms are in kcal/mol. (\*): heterodimers formed with HER<sup>4</sup><sub>TM</sub>

## Discussion

Receptor-receptor interactions in the ErbB family display distinct selectivity defining hierarchy and competition relationships (Prigent et al. 1992; Riese et al. 1995; Tzahar et al. 1996). Until recently, activation of the ErbB receptors was attributed to ligand-induced dimerization of the extracellular domains of the receptor, but numerous data tend to demonstrate that receptors can associate prior to ligand binding. Other dimerization sites, distinct from extracellular sites can mediate homo and heterodimerization. Specifically, the two GxxxG consensus motifs in the TM span that are able to induce intrinsic dimerization (Mendrola et al. 2002) may explain the hierarchy, the competition and directionality of receptor-receptor contacts. The power of heterodimerization of the TM spans has been recently studied by fluorescence resonance energy transfer (FRET) (Duneau et al. 2007) and it was shown that the ability of the TM helix pairs to form dimers varies from non interactive pairs to strong dimerization. A hierarchy of TM interactions exists and  $HER^2_{TM}$  was shown as the TM span having the largest tendency to heterodimerize with  $HER^3_{TM}$  and  $HER^1_{TM}$ . However, a direct implication of the dimerization motifs within the different heterodimers has not been reported. These experiments need to be complemented by structural bases and, for this purpose, we investigated structural models of TM domain of ErbB heterodimers by using conformational search method combined to MD simulations (Samna Soumana et al. 2007). Our aim is to give insight into the hierarchy of interactions of these domains with the hope to evidence the role of the consensus motifs in the association.

### *Hierarchy of interactions*

Surprisingly and contrarily to what is observed from FRET measurements we find that  $HER^3_{TM}$  exhibits the largest tendency to associate with the three other partners.  $HER^3_{TM}$  is the only domain with impaired C-terminal motif but the best representative heterodimer models including this domain show that the two Gly residues localized at the C-terminal side can serve to mediate interactions.  $HER^{1-3}_{TM}$  heterodimers are more stable than  $HER^{2-3}_{TM}$  heterodimers. The association of  $HER^3_{TM}$  with  $HER^4_{TM}$  is a little less favorable and the stability of the heterodimers is of the same order of magnitude than those formed by  $HER^2_{TM}$  and  $HER^1_{TM}$ .  $HER^{1-4}_{TM}$  and  $HER^{2-4}_{TM}$  are the less stable heterodimers indicating a low affinity of the monomers which is in agreement with FRET experiments (Duneau et al. 2007).

The most striking finding concerns  $HER^{3-4}_{TM}$ . Theoretical results show that the stability of this heterodimer is approximately that calculated for  $HER^{1-2}_{TM}$ . Experimental studies place  $HER^{3-4}_{TM}$  at the last rank in the hierarchy of TM interactions while  $HER^{1-2}_{TM}$  is at the first rank owing to the highest affinity.

Strong interactions calculated between  $HER^3_{TM}$  and  $HER^4_{TM}$  can be explained by the strong hydrogen bonds between the N-terminal residues and especially those involving Thr643 of  $HER^3_{TM}$  that increase significantly the inter helical interactions (data not shown). It is to note that this heterodimer counts the greatest number of aromatic residues. For the two representative models the Phe and Tyr rings are stacked at the C-termination and inter helical hydrogen bonds involving one hydroxyl group of Tyr residues and a carbonyl group of the opposed helix are sparsely formed. These interactions do not contribute significantly to stabilize the dimer.

Studies undertaken on the whole receptors indicate that inter ErbB receptor cross talks obey to a hierarchy almost similar to that observed for the TM domains. This is shown in Table 5 along with the hierarchy of interactions deduced from our theoretical study.

According to Tzahar (Tzahar et al. 1996) ErbB2 receptor interacts with the other three proteins with the most potent interaction with ErbB3 which is a kinase defective member of the family. The superior ability for ErbB2 to form dimers with ErbB1 and ErbB3 is also described for the TM domains. Heterodimerization of ErbB3 and ErbB4 receptors was not detected using affinity labeling and coimmuno precipitation experiments (Tzahar et al. 1996) but interactions exist as demonstrated by Riese and al. (Riese et al. 1995). Apparent discrepancies are explained by the differences in ErbB expression levels and cellular environments. Otherwise, simulations studies using a detailed kinetic model applied to ErbB receptor dimerization show that ErbB3-ErbB2 heterodimers are more abundant than ErbB2-ErbB1 heterodimers themselves more abundant than those formed by ErbB1 and ErbB3, this when receptors are subjected to ligand stimulation (Shankaran et al. 2006). The heterodimerization of the ErbB4 receptor was not reported in this study.

Some discrepancies exist between experimental data on the whole receptors and the TM domains on one hand and between experimental and theoretical data in the other hand. Obviously, the comparison is difficult. Experiments on the whole receptors take into account the environment of the cell and the presence of other receptors and their ligands. Experiments on the TM domains, as evidence, ignore the effects of the extra and intracellular domains. In addition, these experiments use synthesized peptides containing minor changes in the

sequence in order to enhance FRET signals and are reconstituted in micelles. Theoretical data reflect exclusively the intrinsic properties of the exact sequence of TM domains without any environmental constraint.

Moreover, the scale of affinity of the TM spans defined from modeling study depends on the choice of the energetic parameter. One parameter can be the stability of the heterodimers giving a global view of the association including both inter molecular and intra molecular interactions which include the effects the deformability of the helices needed for optimal association. A second parameter can be solely the magnitude of helix-helix interactions that highlight the affinity of the two monomers and the importance of the recognition motifs. In both cases it is shown that  $HER^3_{TM}$  is the TM domain which associates the most favorably with the other partners (Table 5). The affinity of  $HER^2_{TM}$  with  $HER^1_{TM}$  is close to that of  $HER^3_{TM}$  with  $HER^4_{TM}$ .  $HER^{2-4}_{TM}$  and  $HER^{1-4}_{TM}$  are the heterodimers of lower affinity as observed experimentally.

**Table 5**

Affinity scale Rank		1	>	2	>	3	>	4	>	5	>	6
Whole receptor (◆)	Experiments	ErbB2-3		ERbB2-1		ErbB2-4		ErbB1-3		ErbB1-4		ERbB3-4?
TM domain (■)		HER <sup>2-1</sup>		HER <sup>2-3</sup>		HER <sup>1-3</sup>		HER <sup>2-4</sup>		HER <sup>1-4</sup>		HER <sup>3-4</sup>
Whole receptor (●)	Kinetic model	ErbB2-3		ERbB2-1		ErbB1-3		–		–		–
TM domain (this work)		(*)	HER <sup>1-3</sup>		HER <sup>2-3</sup>		HER <sup>3-4</sup>		HER <sup>2-1</sup>		HER <sup>1-4</sup>	
	(§)	HER <sup>2-1</sup>		HER <sup>1-3</sup>		HER <sup>3-4</sup>		HER <sup>2-3</sup>		HER <sup>2-4</sup>		HER <sup>1-4</sup>

**Table 5:** Hierarchy of interactions of the ErbB receptors defined from experiments for the whole ErbB receptors (◆) (Tzahar et al. 1996) and for the TM domains (■) (Duneau et al. 2007). Data from Shankaran (●) (Shankaran et al. 2006) are obtained from a kinetic model only for the three receptors ErbB1, 2, 3. The theoretical hierarchy is defined from the best representative structures given in Table 4 using the scale of dimer stability (\*) and the magnitude of helix-helix interactions (§).

### *The recognition motifs*

The main finding of this study is that the SmxxxSm motifs are recognition patterns of the TM domain of the ErbB family. For all the heterodimers, TM helix pairs form stable structures when the small residues, namely Gly, Ala, Ser, Thr, are dominant at the dimer interface. This is a characteristic of the packing of helices observed in membrane proteins (Adamian and Liang 2001; Javadpour et al. 1999).

The association *via* the N-terminus generally leads to the more stable state of the dimer and helix-helix interactions are the strongest in correlation with the best affinity of the monomers. In the case of the HER<sup>1-3</sup><sub>TM</sub> heterodimer the two modes of association either through the N-terminus or the C-terminus are very similar in term of dimer stability and affinity of the monomers.

According to experimental studies, these two motifs may have a different role in the dimerization process. Mendrola and collaborators (Mendrola et al. 2002) demonstrated from TOXCAT measurements that the TM domains of the ErbB receptors self-associate strongly in the absence of the extracellular domain. For ErbB2/Neu it was hypothesized that one of the two motifs of the TM domain could promote the formation of an active conformation. Computational studies have shown that this TM domain may undergo dimerization *via* either one of the two motifs (Fleishman et al. 2002). Supposing a right-handed structure for the two dimerization states, a molecular-switch model for the activation of the receptor was proposed. Movements of sliding and of rotation of each monomer with respect to the other would allow, theoretically, switching between the two dimerization modes that would correspond to active and inactive states.

This molecular-switch model could be applied to the ErbB1 and ErbB2 receptors. Indeed, recent biological studies have shown that the N-terminal motif of the TM domains of these receptors does not play a uniform role in dimerization process (Gerber et al. 2004). While the N-terminal motif of the TM domain of ErbB2 appears to be involved in homodimerization that of ErbB1 seems to be insufficient for the formation of homodimers. It is suggested that the two distinct recognition motifs of the TM domain of ErbB1 are specific for different interaction, the N-terminal motif being necessary for heterodimerization with the TM domain of ErbB2 whereas the C-terminal motif is needed for homodimerization. These two distinct functions could be achieved through a mechanism similar to the molecular-switch model proposed for self-association of HER<sup>2</sup><sub>TM</sub>.

The present study shows that the two dimerization modes exist for all the TM heterodimers even for those formed with  $HER^3_{TM}$  that impairs the second recognition motif. Although left-handed structures are found preponderant over right-handed structures for all types of association, one may hypothesized that this molecular-switch model may apply to all the TM domains of ErbB heterodimers. This has been already suggested by Fleishman et al (Fleishman et al. 2002).

The Figure 4 which shows the dimer interface for the two modes of association of the  $HER^{2-3}_{TM}$  heterodimer illustrates this hypothesis. When the two segments associate *via* the N-terminus, the small residues Ser653, Ala657, Gly660 on  $HER^2_{TM}$  and Thr647 and Ala650 on  $HER^3_{TM}$  are completely or partially buried within the interface while the Gly residues of the second motif on  $HER^2_{TM}$  and one of the two Gly residues on  $HER^3_{TM}$  are fully accessible. On the contrary, when the two TM spans associate *via* the C-terminus, the four Gly residues, 668 and 672 on  $HER^2_{TM}$  and 660 and 661 on  $HER^3_{TM}$ , are completely buried (accessible surface null) while all the small residues of the N-terminus are fully accessible except Ala645 on  $HER^3_{TM}$ .

The same observation can be made for the  $HER^{2-1}_{TM}$  heterodimers. Ser653, Ser656, and Gly660 on  $HER^2_{TM}$  and Thr648 and Gly652 on  $HER^1_{TM}$  are hidden at the interface in the case of the N-terminal association while Gly and Ala of the GxxxG motif at the C-terminus of the two monomers are fully exposed. These same Gly residues (672 on  $HER^2_{TM}$  and 665 on  $HER^1_{TM}$ ) are oriented towards the interface and totally inaccessible while almost all the small residues at the N-terminus are fully accessible when the association is mediated *via* the C-terminus. Similar characteristics are observed for the  $HER^{2-4}_{TM}$  and  $HER^{1-4}_{TM}$  heterodimer models. Such correlation between accessible residues at the C-terminus and buried residues at the N-terminus, and reciprocally, is less pronounced for  $HER^{3-4}_{TM}$  and  $HER^{1-3}_{TM}$  but this remark could be a consequence of helix deformations seen in the models.

Thus, the present study shows that the two dimerization modes are associated to different orientations of the helices. The fact that when one motif is within the dimer interface while the other is oriented outside and become accessible for other interactions appears to be a common characteristics of the  $HER_{TM}$  heterodimers. This result allows suggesting that these two modes of interactions could serve to the ErbB2 receptor to exert the function of lateral mediator as hypothesized from experiments (Shankaran et al. 2006).

Further analyses of the MD models in membrane environment could help to understand such mechanisms.



## **Conclusion**

Dimerization of the ErbB receptor family is governed by complex interactions that involve multiple parts of the receptor. Interactions between the extracellular domains of the receptors are not sufficient to stabilize dimers and additional interactions in the TM spans may aid in their stabilization.

The present study allows a better understanding of the nature of the interactions within heterodimers. The prediction of TM heterodimerization models leads to the description of the different modes of association and to an explicit identification of interfaces. The two modes of association mediated either by one of the two dimerization motifs of the TM domain may play a role in competition processes between the receptors.

MD models for TM heterodimerization add to the understanding of the complex interaction network of the ErbB receptors and can serve to design new experiments.

## References

1. Ferguson, K. M.; Berger, M. B.; Mendrola, J. M.; Cho, H. S.; Leahy, D. J.; Lemmon, M. A. *Mol Cell* 2003, 11, 507-517.
2. Yarden, Y.; Sliwkowski, M. X. *Nat Rev Mol Cell Biol* 2001, 2, 127-137.
3. Salomon, D. S.; Brandt, R.; Ciardiello, F.; Normanno, N. *Crit Rev Oncol Hematol* 1995, 19, 183-232.
4. Citri, A.; Yarden, Y. *Nat Rev Mol Cell Biol* 2006, 7, 505-516.
5. Olayioye, M. A.; Neve, R. M.; Lane, H. A.; Hynes, N. E. *EMBO J* 2000, 19, 3159-3167.
6. Lemmon, M. A.; Bu, Z.; Ladbury, J. E.; Zhou, M.; Pinchasi, D.; Lax, I.; Engelman, D. M.; Schlessinger, J. *EMBO J* 1997, 16, 281-294.
7. Sako, Y.; Minoghchi, S.; Yanagida, T. *Nat Cell Biol* 2000, 2, 168-172.
8. Schlessinger, J. *Cell* 2000, 103, 211-225.
9. Linggi, B.; Carpenter, G. *Trends Cell Biol* 2006, 16, 649-656.
10. Warren, C. M.; Landgraf, R. *Cell Signal* 2006, 18, 923-933.
11. Gullick, W. J. *Endocr Relat Cancer* 2001, 8, 75-82.
12. Garrett, T. P.; McKern, N. M.; Lou, M.; Elleman, T. C.; Adams, T. E.; Lovrecz, G. O.; Zhu, H. J.; Walker, F.; Frenkel, M. J.; Hoyne, P. A.; Jorissen, R. N.; Nice, E. C.; Burgess, A. W.; Ward, C. W. *Cell* 2002, 110, 763-773.
13. Ogiso, H.; Ishitani, R.; Nureki, O.; Fukai, S.; Yamanaka, M.; Kim, J. H.; Saito, K.; Sakamoto, A.; Inoue, M.; Shirouzu, M.; Yokoyama, S. *Cell* 2002, 110, 775-787.
14. Garrett, T. P.; McKern, N. M.; Lou, M.; Elleman, T. C.; Adams, T. E.; Lovrecz, G. O.; Kofler, M.; Jorissen, R. N.; Nice, E. C.; Burgess, A. W.; Ward, C. W. *Mol Cell* 2003, 11, 495-505.
15. Cho, H. S.; Leahy, D. J. *Science* 2002, 297, 1330-1333.
16. Bouyain, S.; Longo, P. A.; Li, S.; Ferguson, K. M.; Leahy, D. J. *Proc Natl Acad Sci U S A* 2005, 102, 15024-15029.
17. Mattoon, D.; Klein, P.; Lemmon, M. A.; Lax, I.; Schlessinger, J. *Proc Natl Acad Sci U S A* 2004, 101, 923-928.
18. Bargmann, C. I.; Weinberg, R. A. *Proc Natl Acad Sci USA* 1988, 85, 5394-5398.
19. Di Fiore, P. P.; Pierce, J. H.; Kraus, M. H.; Segatto, O.; King, C. R.; Aaronson, S. A. *Science* 1987, 237, 178-182.
20. Moriki, T.; Maruyama, H.; Maruyama, I. N. *J Mol Biol* 2001, 311, 1011-1026.

21. Zhang, X.; Gureasko, J.; Shen, K.; Cole, P. A.; Kuriyan, J. *Cell* 2006, 125, 1137-1149.
22. Aifa, S.; Aydin, J.; Nordvall, G.; Lundstrom, I.; Svensson, S. P.; Hermanson, O. *Exp Cell Res* 2005, 302, 108-114.
23. Aifa, S.; Miled, N.; Frikha, F.; Aniba, M. R.; Svensson, S. P.; Rebai, A. *Proteins* 2006, 62, 1036-1043.
24. Bargmann, C. I.; Hung, M. C.; Weinberg, R. A. *Nature* 1986, 319, 226-230.
25. Bargmann, C. I.; Weinberg, R. A. *EMBO J* 1988, 7, 2043-2052.
26. Sternberg, M. J.; Gullick, W. J. *Protein Eng* 1990, 3, 245-248.
27. Senes, A.; Gerstein, M.; Engelman, D. M. *J Mol Biol* 2000, 296, 921-936.
28. Eilers, M.; Patel, A. B.; Liu, W.; Smith, S. O. *Biophys J* 2002, 82, 2720-2736.
29. Kleiger, G.; Grothe, R.; Mallick, P.; Eisenberg, D. *Biochemistry* 2002, 41, 5990-5997.
30. Dawson, J. P.; Melnyk, R. A.; Deber, C. M.; Engelman, D. M. *J Mol Biol* 2003, 331, 255-262.
31. Dawson, J. P.; Weinger, J. S.; Engelman, D. M. *J Mol Biol* 2002, 316, 799-805.
32. Mendrola, J. M.; Berger, M. B.; King, M. C.; Lemmon, M. A. *J Biol Chem* 2002, 277, 4704-4712.
33. Fleishman, S. J.; Schlessinger, J.; Ben-Tal, N. *Proc Natl Acad Sci U S A* 2002, 99, 15937-15940.
34. Gerber, D.; Sal-Man, N.; Shai, Y. *J Biol Chem* 2004, 279, 21177-21182.
35. Stanley, A. M.; Fleming, K. G. J. *Mol Biol* 2005, 347, 759-772.
36. Bell, C. A.; Tynan, J. A.; Hart, K. C.; Meyer, A. N.; Robertson, S. C.; Donoghue, D. J. *Mol Biol Cell* 2000, 11, 3589-3599.
37. Prigent, S. A.; Gullick, W. J. *Embo J* 1994, 13, 2831-2841.
38. Karunakaran, D.; Tzahar, E.; Beerli, R. R.; Chen, X.; Graus-Porta, D.; Ratzkin, B. J.; Seger, R.; Hynes, N. E.; Yarden, Y. *Embo J* 1996, 15, 254-264.
39. Tzahar, E.; Pinkas-Kramarski, R.; Moyer, J. D.; Klapper, L. N.; Alroy, I.; Levkowitz, G.; Shelly, M.; Henis, S.; Eisenstein, M.; Ratzkin, B. J.; Sela, M.; Andrews, G. C.; Yarden, Y. *EMBO J* 1997, 16, 4938-4950.
40. Riese, D. J.; van Raaij, T. M.; Plowman, G. D.; Andrews, G. C.; Stern, D. F. *Mol Cell Biol* 1995, 15, 5770-5776.
41. Duneau, J. P.; Vegh, A. P.; Sturgis, J. N. *Biochemistry* 2007, 46, 2010-2019.
42. Samna Soumana, O.; Garnier, N.; Genest, M. *Eur Biophys J* 2007, DOI 10.1007/s00249-007-0195-6.
43. MacKenzie, K. R.; Prestegard, J. H.; Engelman, D. M. *Science* 1997, 276, 131-133.

44. Tzahar, E.; Waterman, H.; Chen, X.; Levkowitz, G.; Karunagaran, D.; Lavi, S.; Ratzkin, B. J.; Yarden, Y. *Mol Cell Biol* 1996, 16, 5276-5287.
45. Ullrich, A.; Coussens, L.; Hayflick, J. S.; Dull, T. J.; Gray, A.; Tam, A. W.; Lee, J.; Yarden, Y.; Libermann, T. A.; Schlessinger, J.; et al. *Nature* 1984, 309, 418-425.
46. Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Saito, T.; Toyoshima, K. *Nature* 1986, 319, 230-234.
47. Coussens, L.; Yang-Feng, T. L.; Liao, Y. C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg, P. H.; Libermann, T. A.; Schlessinger, J.; Francke, U.; et al. *Science* 1985, 230, 1132-1139.
48. Kraus, M. H.; Issing, W.; Miki, T.; Popescu, N. C.; Aaronson, S. A. *Proc Natl Acad Sci U S A* 1989, 86, 9193-9197.
49. Plowman, G. D.; Culouscou, J. M.; Whitney, G. S.; Green, J. M.; Carlton, G. W.; Foy, L.; Neubauer, M. G.; Shoyab, M. *Proc Natl Acad Sci U S A* 1993, 90, 1746-1750.
50. Brooks, B. R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., Karplus, M. *J Comp Chem* 1983, 4, 187-217.
51. MacKerell, A. D., Jr.; Brooks, B.; Brooks, C. L., III.; Nilsson, L.; Roux, B.; Won, Y.; Karplus, M. In the *Encyclopedia of Computational Chemistry* 1998, 1, 271-277.
52. Zoete, V.; Meuwly, M.; Karplus, M. *Proteins* 2005, 61, 79-93.
53. Walters, R. F.; DeGrado, W. F. *Proc Natl Acad Sci U S A* 2006, 103, 13658-13663.
54. Eilers, M.; Shekar, S. C.; Shieh, T.; Smith, S. O.; Fleming, P. J. *Proc Natl Acad Sci U S A* 2000, 97, 5796-5801.
55. Adamian, L.; Liang, J. *J Mol Biol* 2001, 311, 891-907.
56. Javadpour, M. M.; Eilers, M.; Groesbeek, M.; Smith, S. O. *Biophys J* 1999, 77, 1609-1618.
57. Gratkowski, H.; Lear, J. D.; DeGrado, W. F. *Proc Natl Acad Sci U S A* 2001, 98, 880-885.
58. Zhou, F. X.; Merianos, H. J.; Brunger, A. T.; Engelman, D. M. *Proc Natl Acad Sci U S A* 2001, 98, 2250-2255.
59. Prigent, S. A.; Lemoine, N. R.; Hughes, C. M.; Plowman, G. D.; Selden, C.; Gullick, W. J. *Oncogene* 1992, 7, 1273-1278.
60. Shankaran, H.; Wiley, H. S.; Resat, H. *Biophys J* 2006, 90, 3993-4009.
61. Chothia, C.; Levitt, M.; Richardson, D. *J Mol Biol* 1981, 145, 215-250.
62. Bowie, J. U. *J Mol Biol* 1997, 272, 780-789.

## Conclusion

La prédiction des modèles des hétérodimères transmembranaires montre l'importance des motifs dimérisants de type GxxxG. Les motifs N-terminal et C-terminal présents dans la séquence du domaine TM sont capables de médier l'association de tous les domaines. Les deux modes d'association entraînent des orientations différentes des hélices. Lorsqu'un motif est présent à l'interface du dimère, l'autre motif est orienté vers l'extérieur de l'interface et devient donc accessible pour d'autres interactions.

La visualisation des structures des modèles des hétérodimères illustre la différence des interfaces (**figure 4.1**). La différence d'orientation des hélices est bien illustrée dans le cas des hétérodimères  $HER^{2-3}_{TM}$ ,  $HER^{2-4}_{TM}$ ,  $HER^{2-1}_{TM}$  et  $HER^{1-4}_{TM}$ . Elle semble moins marquée pour les hétérodimères  $HER^{3-4}_{TM}$  et  $HER^{1-3}_{TM}$ .

Les études expérimentales de Duneau et collaborateurs suggèrent fortement que la hiérarchie des interactions des domaines TM est corrélée à celle des récepteurs entiers. Il existe une compétition entre les différents récepteurs. Le domaine TM des récepteurs ErbB pourrait participer à ces processus de compétition par la voie de diverses combinaisons possibles de leur mode d'association en N-terminal et C-terminal.

Par ailleurs, le récepteur ErbB2 est supposé être un médiateur latéral des interactions entre récepteurs (Shankaran et al. 2006). Nos modèles simulés pourraient être exploités dans ce sens pour montrer quel est le rôle de ce domaine TM dans les processus latéraux d'interactions et dans la compétition entre les différents récepteurs. Le milieu membranaire peut jouer un rôle modulateur dans ces processus.

J'ai commencé à aborder cet aspect du milieu membranaire en étudiant l'insertion des différents monomères dans une bicouche de POPC et leur comportement structural. L'étape la plus importante est celle des hétérodimères avec pour objectif une meilleure caractérisation des modes d'association N et C-terminaux.

Il est bien évident que ces études dans le milieu membranaire ne permettront pas de répondre complètement à la question du rôle du domaine TM des récepteurs ErbB dans les processus latéraux des interactions et de l'échelle hiérarchique des interactions mais elles pourront susciter de nouvelles expériences.

La hiérarchie d'interactions des domaines TM des récepteurs HER est loin d'être établie de façon définitive. Des études supplémentaires sur ces domaines TM sont donc nécessaires pour la compréhension du mode d'association des hélices transmembranaires.

## CONCLUSION GENERALE

*Les études de simulations numériques de dynamique moléculaire sur les domaines transmembranaires (TM) des récepteurs à activité tyrosine kinase de la famille de l'EGFR ont permis d'accéder à des informations structurales sur l'association hétérodimérique de ces récepteurs.*

*La première partie de ce travail a consisté à la mise au point d'une méthode théorique de recherche conformationnelle pour explorer les associations des domaines TM. Cette stratégie a été divisée en trois phases. La première phase a consisté en la construction des dimères des domaines TM en balayant l'ensemble des interfaces dimériques. La deuxième étape combinait des techniques de mécanique et de dynamique moléculaires réalisées in vacuo ce qui fait l'originalité de la méthode par rapport aux études antérieures. Cette phase est suivie d'une étape de sélection des structures, basée uniquement sur des critères géométriques déduits d'analyses statistiques sur les protéines membranaires de structure tri-dimensionnelle connue.*

*Pour être validée, cette méthode a été tout d'abord appliquée à l'homodimère du domaine TM de la glycophorine A (GpA) dont la structure tri-dimensionnelle a été résolue par RMN (MacKenzie et al. 1997). Nos simulations ont permis de proposer un nombre très restreint (six) de modèles théoriques d'association des hélices TM parmi lesquelles une structure est en très bon accord avec les données expérimentales. Le RMSD sur les atomes Ca est de 1.5 Å entre cette structure et la structure RMN. De plus, 66% des contraintes RMN sont respectées pour la structure simulée.*

*Cette validation de la méthode de prédiction a permis de poursuivre les études des hétérodimères du domaine TM des récepteurs de la famille ErbB. Nos études montrent une nette préférence pour l'association gauche des hélices TM des récepteurs ErbB, et cela quel que soit l'hétérodimère considéré. De plus, l'association de ces domaines implique principalement des petits résidus de type Gly, Ala, Ser et Thr. La plupart de ces résidus composent des motifs de dimérisation du type GxxxG ou plus largement du type SmxxxSm.*

*Les résultats obtenus nécessiteraient d'être optimisés pour choisir de manière plus catégorique lequel de ces modèles est préférentiellement formé. Toutefois, on peut également supposer qu'il existe un équilibre entre les différents modèles proposés. Afin d'affiner encore*

plus les modèles, d'autres études théoriques de simulations de dynamique moléculaire sont envisagées. Il s'agit par exemple d'étudier dans un environnement membranaire constitué de molécules de POPC, le domaine TM des différents monomères ErbB et les modèles des hétérodimères sélectionnés. Les simulations des monomères ErbB dans ce modèle de membrane sont déjà effectuées. Celles des hétérodimères viennent de démarrer. Ces études devraient permettre d'une part de comprendre le mode d'insertion des hélices ErbB dans un environnement membranaire et aussi de rendre compte de l'effet de cet environnement sur les domaines TM des récepteurs ErbB et leur association. Ces études devraient aussi permettre d'affiner la hiérarchie d'interactions des domaines TM. Il est envisagé de poursuivre les études pour tenter de déterminer les énergies libres d'interactions. Ainsi, les résultats obtenus pourront être comparés aux données expérimentales.

Toutefois, pour définir un modèle préférentiel d'association ou pour caractériser un équilibre entre plusieurs modes d'association, des études expérimentales supplémentaires seront nécessaires. Ces études pourraient être effectuées sur des mutations au sein du domaine transmembranaire par exemple au niveau des petits résidus (au niveau d'un des motifs de dimérisation) qui occupent l'interface des dimères gauches sélectionnés afin de bloquer cette association. Si l'activité de ces récepteurs est inhibée *in vitro* et/ou *in vivo*, alors, on pourrait conclure que l'association gauche des hélices des récepteurs ErbB est la plus favorable. De plus, le motif de dimérisation le plus impliqué pourrait être identifié.

On pourrait aussi envisager d'appliquer cette méthode de prédiction de modèles aux protéines membranaires formées de plus de deux hélices TM. L'avantage de cette procédure est de ne pas faire appel à des contraintes d'angle de croisement et/ou de symétrie contrairement aux méthodes classiques de prédiction de structures.

La complexité des interactions entre les récepteurs ErbB et de leur mécanisme d'action est loin d'être complètement élucidée. Beaucoup d'équipes dans le monde déploient des efforts importants pour comprendre le fonctionnement de cette machinerie.

Les travaux effectués au cours de cette thèse représentent une part infime dans la réflexion globale des mécanismes impliqués dans la fonction des récepteurs ErbB. Mais on peut espérer que les avancées des techniques expérimentales et théoriques permettront de caractériser les structures des récepteurs entiers et leur association afin d'aider au développement de nouveaux agents thérapeutiques contre les cancers.



# **ANNEXES**

**ANNEXE 1 : ROLE DES RECEPTEURS ErbB DANS LES PROCESSUS  
CANCEREUX**

## I Introduction

Les mutations et les dysfonctionnements de gènes codant pour des RTK entraînent de nombreux changements dans leurs structures et leurs fonctions ce qui contribue à l'apparition de diverses pathologies telles que le psoriasis, l'athérosclérose, ainsi que de nombreux cancers comme des cancers du poumon, du colon, du sein, etc. (Levitzki 1999; Robertson et al. 2000).

L'auto-suffisance en terme de signaux de croissance, l'insensibilité aux signaux d'inhibition de croissance, la perte d'apoptose, le potentiel de réplication sans limite, l'induction d'une angiogenèse soutenue et l'invasion cellulaire et les métastases sont les principaux points qui caractérisent le cancer (Hanahan and Weinberg 2000).

Les RTK et leurs voies de signalisation sont impliqués dans au moins quatre de ces points. En effet, des mutations activatrices et/ou la surexpression des RTK et de leurs ligands peuvent conduire à un dérèglement de la croissance cellulaire. De plus, on peut citer en exemple le récepteur de l'IGF1 qui est un élément clé de la machinerie contrôlant la mort cellulaire (Furstenberger and Senn 2002) ou les récepteurs du VEGF, du PDGF et du FGF, impliqués dans la signalisation initiant l'angiogenèse (Distler et al. 2003; Folkman 2002). Des mutations au sein de la tyrosine kinase Met qui est un récepteur à haute affinité pour l'HGF/SF (*Hepatocyte Growth Factor/scatter Factor*) ont été observées dans des carcinomes papillaires rénaux humains. Jeffers et collaborateurs (Jeffers et al. 1998) ont montré que ces mutations entraînaient un dérèglement de l'activité enzymatique de Met. De plus des cellules NIH 3T3 exprimant la protéine Met activée par mutation devenaient transformantes *in vitro* et tumorigènes *in vivo*. Ils ont également montré que ce mutant Met induit des métastases expérimentales *in vivo*.

Chacun de ces points constituent une cible pour une nouvelle génération de molécules à visée thérapeutique, ce qui positionnerait les RTK à un rang privilégié parmi les cibles potentielles pour lutter contre des pathologies liées à ces anomalies.

En effet, en 2002, Cohen affirmait que les protéines kinases seraient au cours du 21<sup>ème</sup> siècle les principales cibles de traitements dirigés contre le cancer mais aussi contre d'autres pathologies (Cohen 2002).

Il y a une vingtaine d'années déjà, Bishop et Varmus (Varmus and Bishop 1986) montraient que les gènes de rétrovirus animaux induisant le cancer comme v-src correspondent en fait à

des gènes de l'hôte ayant été mutés et recombinaés dans le génome viral. Cela a soulevé le problème suivant : est-ce que l'oncogène en tant que concept présente un lien avec l'apparition de cancers humains?. Ce premier lien établi entre le produit d'un gène humain qui régule la prolifération cellulaire normale et un oncogène viral suggérait que le développement de cancers humains pourrait impliquer des anomalies au niveau de l'expression et de la structure de gènes endogènes ayant des rôles importants dans la régulation de la prolifération cellulaire. Le clonage et l'analyse de séquences d'un ADNc codant pour une protéine de surface cellulaire qu'est l'EGFR humain, a apporté une réponse partielle à cette question en révélant un lien très proche entre ce récepteur et l'oncogène ErbB (Downward et al. 1984; Ullrich et al. 1984). Ces résultats furent d'autant plus intéressants que le clonage a été réalisé à partir d'une lignée cellulaire humaine appelée A431 obtenue à partir d'un sarcome de l'épiderme humain. La recherche d'aberrations génétiques similaires dans des tissus tumoraux en utilisant des sondes d'ADNc de l'EGFR et de son proche parent ErbB2/HER2/Neu a permis de révéler que le gène codant pour le récepteur HER2/Neu est 100 fois amplifié dans des cellules tumorales, et ce chez 30% des patientes atteintes de cancers invasifs du sein. Une corrélation clinique a d'ailleurs été établie entre l'amplification du gène codant pour HER2/Neu, la surexpression de ces récepteurs et un taux de survie et de rémissions largement diminué par rapport à des patientes atteintes de cancers qui présentent une quantité normale de ces récepteurs (Slamon et al. 1989). C'est pourquoi la famille d'oncogènes ErbB/HER est une cible prometteuse pour de nouvelles molécules à visée anticancéreuse, et elle a été la première pour le développement de médicaments anti-oncogènes. L'importance de la découverte de nouvelles cibles dans le traitement du cancer semble évidente.

En effet, ces récepteurs peuvent être inactivés soit par des molécules solubles dans l'eau comme des anticorps ou des antagonistes de liaison, soit par des substances perméables à la cellule comme des inhibiteurs de kinase.

Je traiterai dans un premier temps de l'altération des gènes ErbB avant de voir dans une seconde partie le lien entre les découvertes croissantes sur les fonctions de ces récepteurs ErbB et leurs ligands et leurs dysfonctionnements dans certains cancers chez l'homme. Dans une troisième partie, nous verrons le développement de nouveaux médicaments ciblant ces récepteurs.

## II Altération des gènes ErbB

Plusieurs altérations génétiques des récepteurs ErbB sont constatées dans les cancers humains (**tableau A.1.1**) (Normanno et al. 2006).

La surexpression du gène de ErbB1 a été trouvée dans plusieurs tumeurs et cela est souvent associé à une surexpression de la protéine ErbB1 (Normanno et al. 2003; Normanno et al. 2003; Salomon et al. 1995). Des mutations de l'EGFR se produisent fréquemment dans un certain nombre de cancers comme des gliomes, des cancers du poumon n'étant pas à petites cellules (NSCLC), des cancers de la prostate et des ovaires (Garcia de Palazzo et al. 1993; Normanno et al. 2006; Pedersen et al. 2001; Schwechheimer et al. 1995). Il a été constaté une amplification du gène ErbB1 dans environ 37% à 58% des glioblastomes multiformes (Wikstrand et al. 1998). Toutefois cette amplification est combinée à des mutations (délétion, duplication en tandem) du gène ErbB1 pouvant être importantes pour le pouvoir transformant de ce récepteur (Frederick et al. 2000).

Certaines de ces mutations sont associées à une activité de ErbB1 indépendante du ligand (Kuan et al. 2001). La mutation la plus commune et la mieux décrite de l'EGFR est la mutation de type III correspondant à une suppression d'une partie du domaine extracellulaire par délétion des exons 2-7 (**tableau A.1.1**). Celle-ci aboutit à une activation constitutive du récepteur mutant appelé EGFRvIII (ou d801 EGFR, de 2-7 EGFR) qui n'est pas endocyté (régulation en retour) et qui a un fort pouvoir transformant (Huang et al. 1997). Selon Shinojima et collaborateurs (Shinojima et al. 2003), la présence de cette forme de mutation est un indicateur d'un pronostic très réservé chez les patients atteints de glioblastomes multiformes. Différents groupes ont étudié l'auto dimérisation de l'EGFRvIII et bien que des résultats opposés existent, des données récentes indiquent que l'auto dimérisation de ce récepteur est comparable à la dimérisation de l'EGFR induite par la fixation du ligand (Chu et al. 1997; Fernandes et al. 2001). Ce mutant est fréquemment exprimé dans divers types de tumeurs incluant des tumeurs gliales, du sein, des ovaires et des NSCLC (Moscatello et al. 1998). De plus, il semblerait que les anticorps anti-EGFR dirigés contre les épitopes extracellulaires des récepteurs mutés par délétion n'aient pas d'activité inhibitrice, tandis que des petites molécules inhibitrices pourraient être actives comme il l'a été démontré par inhibition de l'autophosphorylation de l'EGFRvIII dans les cellules NR6M mises en culture (Ciardiello and Tortora 2003). Il existe d'autres formes de mutations de ErbB1 (voir **tableau**

**A.1.1)** qui sont impliquées dans les cancers c'est le cas par exemple des mutations du domaine TK.

La surexpression de ErbB2 est associée à une amplification du gène ErbB2, même si cette surexpression a été aussi observée en l'absence d'amplification du gène (Normanno et al. 2003; Salomon et al. 1995). La mutation du gène ErbB2 n'est pas fréquente dans les cancers même si des récentes études (Stephens et al. 2004) montrent une insertion ne changeant pas le cadre de lecture ou des substitutions non sens au niveau du domaine TK. Ces mutations sont plus fréquemment observées chez des patientes et des non fumeurs ayant un adénocarcinome avec le même phénotype que les patients portant des mutations EGFR.

Dans le cancer du sein, il a aussi été observé des mutations de ErbB2 qui sont très proches de celles de Neu chez les souris transgéniques (Normanno et al. 2006). Ces mutants pourraient avoir une activité constitutive comme ceux de Neu et jouer un rôle important dans le développement du cancer du sein (Siegel et al. 1999).

Les mutations de ErbB1 et ErbB2 sont observées chez des patients ne présentant pas de mutations ras. Ces protéines sont donc engagées dans des voies de transformations différentes. La surexpression de ErbB3 et ErbB4 conduirait à des cancers du sein de forme différente.

Récepteurs	Altérations génétiques	Dépendance du ligand
ErbB1	Amplification des gènes	+
	N-terminal tronqué (EGFRvI)	-
	Délétion des exons 14-15 (EGFRvII)	+
	Délétion des exons 2-7 (EGFRvIII)	
	Délétion des exons 25-27 (EGFRvIV)	+
	C-terminal tronqué (EGFRvV)	+
	Duplication en tandem des exons 2-7	+
	Duplication en tandem des exons 18-25	-
	Duplication en tandem des exons 18-26	-
	Petite délétion ne décalant pas le cadre de lecture ou mutations ponctuelles dans le domaine kinase (exons 18-21)	+
ErbB2	Amplification des gènes	Non évaluée
	Différents fragments ressemblant aux mutants Neu	Non évaluée
	Petites insertions ne décalant pas le cadre de lecture ou substitutions non sens dans le domaine kinase (exons 18-24)	Non évaluée

**Tableau A.1.1 :** Altérations génétiques des récepteurs ErbB dans les cancers humains (Normanno et al. 2006).

### III Les récepteurs ErbB et leurs ligands dans les cancers humains

Le dérèglement de la croissance cellulaire est la caractéristique principale de tout néoplasme. Les néoplasmes ont la capacité d'envahir des tissus normaux et de former des métastases pouvant se développer dans des sites éloignés du foyer tumoral primaire, ce qui constitue un autre critère essentiel de la définition du cancer (Blume-Jensen and Hunter 2001). Ce dérèglement de la croissance cellulaire résulte d'une perturbation de la transduction du signal pouvant être définie, au sens large, comme tout signal cellulaire modulant ou altérant le comportement de la cellule (Hunter 2000; Pawson and Nash 2000). Par conséquent, les cancers ne résultent pas nécessairement d'une simple augmentation du taux de prolifération cellulaire : c'est plutôt l'équilibre entre le taux de progression du cycle cellulaire (division cellulaire) et la croissance cellulaire (masse cellulaire) d'un côté, et la mort cellulaire programmée (apoptose) de l'autre qui est important (Reed 1999).

Quatre mécanismes principaux ont été définis pour être à l'origine de la transformation oncogénique des protéines à activité tyrosine kinase. Le premier consiste en une transduction d'un proto-oncogène par un rétrovirus (correspondant à une protéine à activité tyrosine kinase) concomitante avec des modifications de structure altérant les fonctions de la protéine : c'est un mécanisme de transformation commun chez le poulet et le rongeur. Le second mécanisme consiste en un ré-arrangement génomique (comme les translocations chromosomiques) qui peut aboutir à la formation de protéines de fusion oncogéniques impliquant au minimum un domaine catalytique de PTK et une protéine non apparentée qui possède une fonction de dimérisation. Le troisième mécanisme implique des mutations permettant un gain de fonctions (GOF, *gain of function*) où des petites délétions dans les RTK ou les PTKs cytoplasmiques sont associées à différentes tumeurs malignes. Le dernier mécanisme consiste en la surexpression des PTKs résultant d'une amplification de gènes associés à différents types de cancers humains. En général, l'effet transformant peut être attribué à une activité kinase constitutive (c'est-à-dire indépendante du ligand et augmentée) avec une signalisation en aval altérée à la fois sur le plan quantitatif et qualitatif (Blume-Jensen and Hunter 2001).

Plusieurs études confirment le rôle des récepteurs ErbB et de leurs ligands dans les cancers humains. Il a été montré depuis que les récepteurs ErbB sont impliqués dans plusieurs cancers

en particulier des tumeurs solides. Le **tableau A.1.2** montre quelques implications des récepteurs ErbB dans les carcinomes humains (Normanno et al. 2006).

Les altérations observées sont des mutations, des surexpressions ou une stimulation anormale par leurs ligands (Gschwind et al. 2004; Holbro et al. 2003). Ainsi, il a été observé dans 50% à 70% des cancers des poumons, du sein et du colon, une expression de ErbB1 ou ErbB3 (Abd El-Rehim et al. 2004; Normanno et al. 2003). ErbB1 est par contre observé dans seulement 15% des 807 cas de cancers du sein invasifs par DiGiovanna et collaborateurs (DiGiovanna et al. 2005). L'expression de ErbB2 est aussi plus faible (30% dans les cancers du sein). S'agissant de ErbB4, il est exprimé dans 50% des cancers du sein (Abd El-Rehim et al. 2004; Normanno et al. 2003), et environ dans 22% des cancers du colon (Lee et al. 2002).

Le récepteur de l'EGF fut le premier récepteur transmembranaire mis en cause dans la cancérogenèse. Différents éléments ont mis en évidence l'importance de l'EGFR en tant que cible dans la thérapie anticancéreuse (Ciardiello 2000; Ciardiello and Tortora 2001). Il est exprimé ou surexprimé dans un certain nombre de tumeurs solides incluant des cancers du sein, de la prostate et dans des NSCLC (Salomon et al. 1995) : on peut citer en exemple une étude menée par Klijn et collaborateurs (Klijn et al. 1992) révélant que environ 45% des 5000 patientes atteintes de cancer du sein était de «positifs en EGFR ». Dans certaines tumeurs, l'EGFR est associé à une phase avancée du cancer (Salomon et al. 1995), mais aussi à un mauvais pronostic. De plus, il a été observé une résistance à la chimiothérapie, à la thérapie hormonale et à la radiothérapie (Brabender et al. 2001; Sartor 2000). En fait, une expression élevée de l'EGFR est considérée communément comme le principal mécanisme par lequel la signalisation de l'EGFR est augmentée dans les cellules cancéreuses.



Transgène	Tissu cible	Type de tumeur
EGFR	Sein Urothélium Gliomes Œsophage	Adénocarcinome mammaire chez les animaux allaitant Aucun Aucun Aucun
c-Neu activé	Sein	Cancer mammaire chez des femelles vierges et chez les mâles
c-Neu	Sein	Cancer mammaire après une longue phase de latence
ErbB2 du rat	Peau	Cancer de cellules de la peau
ErbB3	Sein	Cancer des poumons après une longue phase de latence
TGF $\alpha$	Multiple Sein Peau	Cancer du foie ou cancer mammaire chez les femelles multipareuses Cancer mammaire chez les femelles vierges et multipareuses Aucun
AR	Peau	Aucun
HRG	Sein	Cancer mammaire chez les femelles multipareuses

**Tableau A.1.2 : Résumé des études transgéniques des récepteurs ErbB et leurs ligands.**

Cependant, un certain nombre d'autres mécanismes semblent également avoir leur importance telles que des mutations activatrices de l'EGFR, une diminution du niveau de phosphatases, une augmentation de la co-expression des récepteurs et des ligands comme le TGF $\alpha$  et l'amphiréguline, une hétérodimérisation avec HER2 et/ou d'autres membres de la famille des récepteurs ErbB et enfin des interactions de l'EGFR avec d'autres systèmes de récepteurs hétérologues (Ciardiello and Tortora 2003).

L'augmentation de l'expression de ligands spécifiques de l'EGFR comme l'EGF, l'amphiréguline ou le TGF $\alpha$  a un effet sur la signalisation intracellulaire. L'hypothèse de l'existence d'une boucle autocrine s'explique par le fait que la plupart des cancers épithéliaux exprimant un haut niveau de TGF $\alpha$  co-exprime l'EGFR (Normanno and Ciardiello 1997; Salomon et al. 1995). Par exemple, une analyse immunohistochimique sur 131 adénocarcinomes primaires provenant de poumons humains a montré dans les cas «EGFR-positifs», un taux de survie à cinq ans significativement réduit pour les patients présentant des niveaux d'expression élevés de l'EGF et du TGF $\alpha$  par rapport aux patients présentant des faibles niveaux d'expression de ces deux ligands (Tateishi et al. 1990). Un autre exemple d'étude réalisée sur 91 patients atteints de cancers du cou et de la tête, traités par résection chirurgicale radicale, a montré que la surexpression de l'EGFR et du TGF $\alpha$  prédit de façon

significative une période asymptomatique plus courte et une survie totale (Grandis and Drenning 1998).

La signalisation de l'EGFR est également dépendante de la formation différentielle des dimères. Les combinaisons de ces récepteurs engendrent diverses activités dans les cellules en culture.

Plusieurs études présentent des résultats similaires. La co-expression des différents récepteurs ErbB est liée à un mauvais pronostic en comparaison à l'expression d'un seul récepteur dans les cancers du colon par exemple. Brabender et collaborateurs (Brabender et al. 2001) ont montré que les patients atteints du cancer des poumons avec une forte expression de ErbB1 et ErbB2 présentaient un pronostic réservé comparativement à ceux ayant des cancers où un seul récepteur est surexprimé.

Ainsi, les niveaux d'expression d'autres récepteurs de la famille ErbB, plus particulièrement HER2, affecteraient de façon significative la signalisation dans les cancers humains. En plus de l'activation par ses propres ligands, le réseau de signalisation de l'EGFR peut être également influencé par d'autres systèmes de signalisation cellulaire. Le plus étudié d'entre eux étant celui des récepteurs couplés aux protéines G (GPCR) (Carpenter 1999; Prenzel et al. 2000; Yarden and Sliwkowski 2001). Les GPCR peuvent avoir des effets positifs sur la signalisation de l'EGFR par des mécanismes impliquant les métalloprotéases de la matrice ou par une activation indirecte de Src qui phosphoryle les résidus tyrosine kinase intracellulaires du récepteur. La trans-activation de la signalisation de l'EGFR peut également se produire par les hormones stéroïdiennes activant la transcription de gènes codant pour les ligands de l'EGFR comme le TGF $\alpha$  (Bates et al. 1988).

Neu, isolé à partir d'un neuroglioblastome de rats, a été découvert comme un mutant oncogénique puissant présentant des homologies de séquences avec des récepteurs de la famille ErbB. Il a notamment été démontré que Neu était équivalent à ErbB2 (Schechter et al. 1984). ErbB2 sauvage a souvent été retrouvé amplifié au niveau génomique et/ou surexprimé au niveau protéique. Il semble que ce soit le mécanisme principal d'activation de l'oncogène ErbB2 chez l'homme. L'amplification du proto-oncogène HER2 a été pour la première fois mise en évidence par Slamon et collaborateurs (Slamon et al. 1987) dans une étude menée sur 189 carcinomes primaires du sein. Dans la majorité des cas, la surexpression est corrélée avec une chimiorésistance et un très mauvais pronostic quant à la survie du patient (Hynes and Stern 1994; Klapper et al. 2000). Ce sujet a surtout été traité dans le domaine de la biologie du cancer du sein où le récepteur ErbB2 est amplifié dans 20 à 30% des cas; d'autre part, les

thérapies altérant l'expression d'ErbB2 ont maintenant un impact clinique. La surexpression d'ErbB2 observée dans les tumeurs du sein mais également dans d'autres types de tumeurs telles que 10 à 30% des tumeurs ovariennes et 30 à 35% des adénocarcinomes pulmonaires (Tzahar and Yarden 1998), augmente la tumorigénicité (Di Fiore et al. 1987; Hudziak et al. 1987).

Le mécanisme expliquant l'action tumorigène des récepteurs ErbB2 surexprimés pourrait être en relation avec l'activité d'autophosphorylation basale relativement élevée de ce récepteur (Lonardo et al. 1990). Ainsi, leur surexpression pourrait conduire à une formation spontanée des homodimères analogues aux formes activées par des carcinogènes des récepteurs Neu/ErbB2 de rongeurs. Cette activation est causée par une mutation au sein du domaine transmembranaire provoquant une dimérisation et une activation constitutive du récepteur (Bargmann and Weinberg 1988; Weiner et al. 1989).

Parmi les membres de la famille ErbB, ErbB3 est inimitable du fait de la déficience de l'activité catalytique de son domaine TK (Guy et al. 1994; Prigent and Gullick 1994) et aussi du fait de la conformation du domaine extracellulaire de ses monomères (Citri and Yarden 2006). ErbB3, dont les ligands principaux sont les NDF (Neu Differentiation Factor) ou neuréguline, est largement exprimé dans les tissus normaux (Prigent et al. 1992). Ainsi, bien que les homodimères de ErbB3 soient inactifs, ces récepteurs sont capables d'induire une signalisation par formation d'hétérodimères en s'associant soit à l'EGFR, soit à ErbB4, mais surtout à ErbB2, décrit comme le partenaire privilégié. Les récepteurs ErbB3 ne semblent donc pas avoir d'activité transformante par eux-mêmes, mais ils sont capables d'augmenter le pouvoir transformant des récepteurs ErbB2 lorsqu'ils sont surexprimés dans les cellules NIH 3T3. La transformation de cellules NIH 3T3 par ces hétérodimères est corrélée avec la transphosphorylation de ces récepteurs induite par le ligand (Zhang et al. 1996). De plus, ErbB3 est surexprimé dans un certain nombre de tumeurs malignes comme des cancers gastriques, des ovaires, du pancréas, de la prostate, de la vessie, du poumon et de l'endomètre (Gullick 1996). La surexpression de ErbB3 est retrouvée dans 20 à 30% des formes invasives et dans presque 1/3 des cancers du sein (Badra et al. 2006). De plus, la forme la plus mitogénique des dimères ErbB est le couple ErbB2/ErbB3 (Citri et al. 2003). Cette forme présente une activité constitutive dans les cellules cancéreuses du sein avec une surexpression du gène ErbB2. Ainsi, l'implication des récepteurs ErbB3 dans des cancers humains est essentiellement due à leur capacité à potentialiser les effets d'autres récepteurs comme les récepteurs ErbB1 et 2.

Comme pour ErbB3, les ligands primaires de ErbB4 sont les NDF (ou héréguines) et ses ligands sont capables d'induire la phosphorylation du récepteur HER2 lorsque celui-ci est associé de façon hétérodimérique à ErbB3 ou ErbB4 (Alimandi et al. 1995; Carraway and Cantley 1994; Kita et al. 1994; Plowman et al. 1993). De plus, la formation d'hétérodimères associant ErbB4 avec soit ErbB1 soit ErbB2 est capable d'induire la transformation de cellules NIH 3T3. Cette transformation est dépendante du NDF (Zhang et al. 1996). Des études (Bacus et al. 1996; Bacus et al. 1994; Srinivasan et al. 2000) ont montré que ErbB4 n'est pas surexprimé dans des carcinomes mammaires comme ErbB1 et ErbB2. En effet, son niveau d'expression est plutôt modéré voire faible. Tandis que l'expression d'ErbB1 et d'ErbB2 est corrélée avec des tumeurs négatives pour les récepteurs des oestrogènes et à un très mauvais pronostic, l'expression d'ErbB4 est associée à des effets opposés. En effet, l'expression d'ErbB4 est corrélée avec la présence de récepteurs aux oestrogènes, à un type histologique mieux différencié, et à un pronostic plus favorable (Carpenter 2003).

Le récepteur ErbB4 joue un rôle important dans le développement et la différenciation des différents tissus, particulièrement les tissus cardiovasculaires, le système nerveux et les glandes mammaires et aussi dans différents processus pathologiques comme les tumeurs malignes (Karamouzis et al. 2007). Son expression a été observée dans d'autres tumeurs que les cancers mammaires (Bei et al. 1999; Gilmour et al. 2001; Merimsky et al. 2002; Srinivasan et al. 1999) notamment des cancers du colon, de la prostate, du poumon, des ovaires, du pancréas, de l'endomètre, des bronches, du col de l'utérus, de l'estomac et de la thyroïde. Dans le cas des médulloblastomes, une expression d'ErbB4 associée à celle d'ErbB2 constitue un marqueur d'agressivité tumorale et est corrélée à un mauvais pronostic de survie (Gilbertson et al. 1997). Ainsi, de nombreux RTK et notamment les récepteurs de la famille ErbB sont impliqués par différents mécanismes dans diverses pathologies dans lesquelles on trouve de façon redondante le cancer. Leur implication dans cette maladie fait d'eux une cible privilégiée dans la thérapie anticancéreuse.

Dans les cancers humains, il n'y a pas que l'expression des récepteurs ErbB. Des études montrent aussi une co-expression de différents ligands de type EGF dans plusieurs cancers humains. En 2001, Normanno et collaborateurs ont montré qu'il y avait co-expression de TGF $\alpha$ , AR et/ou NRG dans le cancer du colon, du sein, des poumons, des ovaires et dans le cancer gastrique (Normanno et al. 2001). De plus le ligand NRG1 par exemple est impliqué dans les maladies neurodégénératives telles l'Alzheimer (Go et al. 2005; Willem et al. 2006).

Il apparaît donc que le réseau d'interactions des récepteurs ErbB et leurs ligands est impliqué dans la progression des cancers humains. De ce fait, pour lutter plus efficacement contre ces cancers, il est nécessaire de connaître les différentes voies de signalisation impliquant les récepteurs ErbB et leurs ligands, pour pouvoir y exercer un contrôle plus rigoureux.

## **VI Les récepteurs ErbB et la thérapie**

Les récepteurs ErbB sont des cibles anticancéreuses. Il existe deux stratégies visant à inhiber leur action, soit l'utilisation des anticorps dirigés contre le domaine extracellulaire, soit l'utilisation des drogues inhibant l'activité TK.

Plusieurs traitements (voir **tableau A.1.3**) ont été développés concomitamment à l'augmentation de la détection précoce des cancers du sein. Le récepteur ErbB2 est considéré comme étant un bon régulateur de la prolifération des cellules cancéreuses du sein. Ainsi, plusieurs stratégies sont élaborées pour réduire le taux de ErbB2 ou son activité. Une surexpression de ErbB2 est présente chez environ 15 à 25% des cancers du sein et est associée à une agressivité tumorale.

L'herceptine (Slamon et al. 1987) est un anticorps monoclonal humanisé recombinant de classe IgG1, dirigé contre le récepteur ErbB2, qui se fixe à sa partie extracellulaire. Ce site est identifié comme une cible potentielle pour une thérapie anticancéreuse. Il est actuellement utilisé dans le traitement du cancer du sein où ErbB2 est surexprimé. En effet, plusieurs études ont montré, chez des patientes présentant un cancer du sein métastasé ErbB2-positif, une efficacité du trastuzumab administré en monothérapie ou en association avec une chimiothérapie. L'association de ce médicament à une chimiothérapie classique augmente les bénéfices cliniques (Slamon et al. 2001). L'herceptine est l'un des meilleurs agents connus pour cela. Cependant, dans 70% des cas, les patients font une résistance à ce traitement (Kute et al. 2004). Son efficacité est limitée chez certains patients, cela laisse penser que d'autres facteurs, autres que ErbB2, sont nécessaires pour le développement des cellules cancéreuses du sein (Karamouzis et al. 2007). Un des mécanismes de l'herceptine est de bloquer l'agrégation des récepteurs et d'augmenter leur endocytose, diminuant ainsi le nombre de récepteurs à la surface de la cellule.

Le pertuzumab, à la différence du trastuzumab, est actuellement évalué en phase II sur des patients atteints de tumeurs aux seins, poumons, prostate et ovaires, où le récepteur ErbB2 n'est pas surexprimé.

Contrairement au pertuzumab, l'anticorps anti-ErbB2 est très prometteur même dans les cas où ErbB2 n'est pas surexprimé (Jackson et al. 2004).

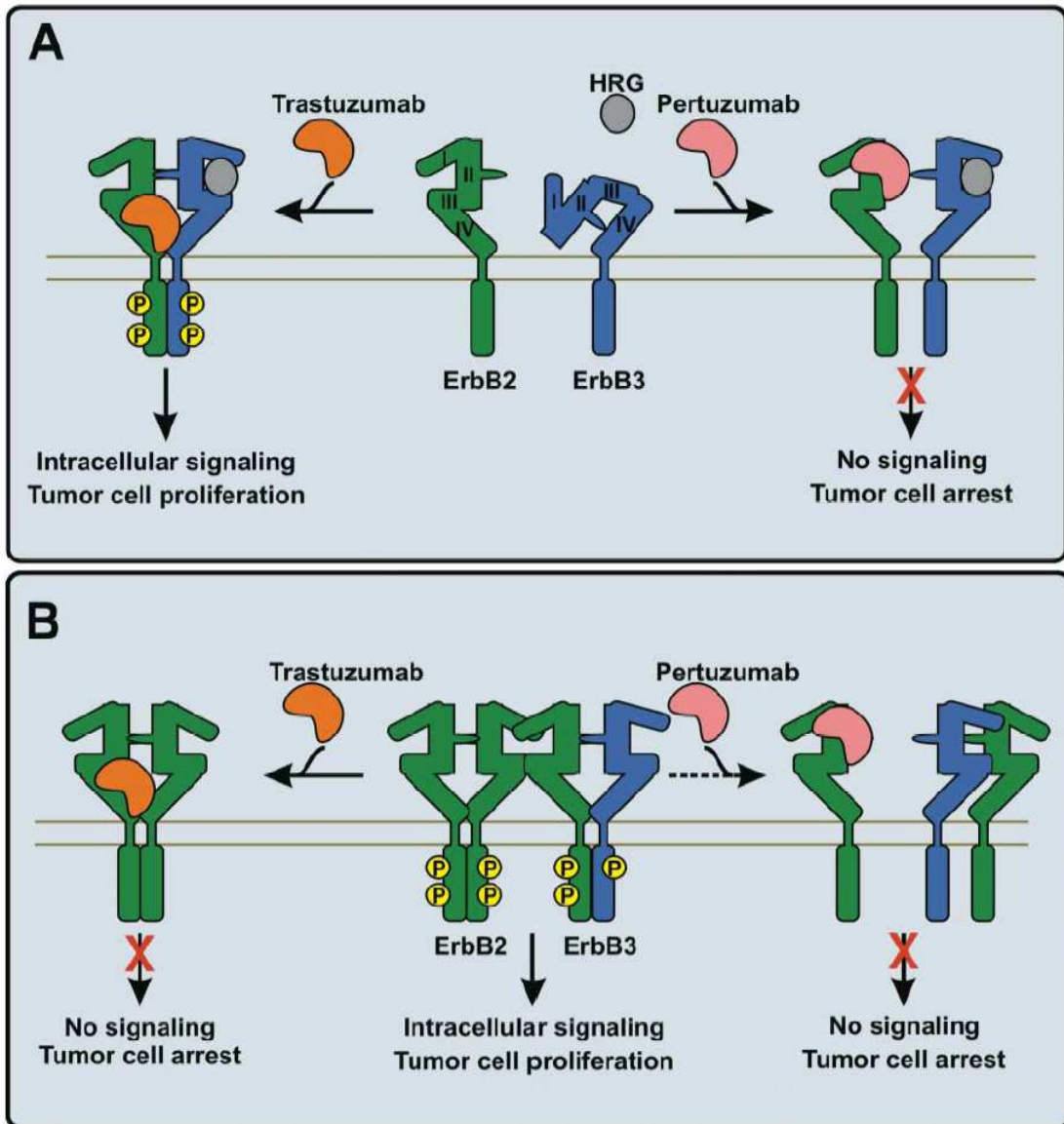
Un exemple du mode de fonctionnement de l'herceptine (ou trastuzumab) et du pertuzumab sur le récepteur ErbB2 est indiqué sur la **figure A.1.1** (Badache and Goncalves 2006; Badache and Hynes 2004).

Le récepteur ErbB4 est considéré comme une cible anti-proliférative nécessaire et suffisante dans la lutte contre les cancers du sein (Barnes et al. 2005). De plus, la co-expression des récepteurs ErbB serait responsable de l'augmentation du potentiel transformant et du mauvais pronostic. Ainsi, une stratégie plus efficace serait d'utiliser des inhibiteurs bispécifiques pour bloquer les récepteurs ErbB1, ErbB1 et ErbB2 et/ou toute la famille ErbB (Karamouzis et al. 2007), par exemple, le lapatinib (ou EKB-569) pour agir sur ErbB1 et ErbB2 (Citri and Yarden 2006). Actuellement, certains inhibiteurs de TKs (TKIs) sont considérés comme des bons inhibiteurs des récepteurs ErbB dans les différentes formes du cancer du sein (Xu et al. 2005). Les premiers essais cliniques utilisant les TKIs sur des patientes atteintes de cancers du sein ayant des métastases réfractaires aux traitements anciens ont été peu concluants (Roy and Perez 2006).

Le ZD1839 (gefitinib, ou Iressa) est actuellement utilisé dans le traitement du cancer des NSCLC aux USA et au Japon. C'est un traitement en monothérapie (qui se fixe sur la partie extracellulaire de ErbB1) chez des patients atteints du cancer du poumon après échecs des chimiothérapies standards (Roskoski 2004). Ce médicament est actuellement en test en combinaison avec d'autres produits dans le traitement des tumeurs solides comme le cancer colorectal. Bien qu'il soit considéré comme un inhibiteur spécifique de l'activité kinase de ErbB1, il y a actuellement très peu d'études montrant cette sélectivité (Cohen et al. 2004).

Erlotinib (OSI-774, ou Terceva) est un autre inhibiteur de ErbB1 approuvé pour la lutte contre le cancer des NSCLC.

Des ligands comme NRG seraient aussi impliqués dans les cancers du sein. Ainsi, le blocage de NRG inhiberait, in vitro, les tumeurs dans les cancers du sein (Tsai et al. 2003) alors que son expression est souvent associée à un pronostic de survie faible (Menendez et al. 2006).



**Figure A.1.1 : Plus d'une voie pour inhiber le récepteur ErbB2. Action de l'herceptine (trastuzumab) et du pertuzumab. A :** Dans les cellules ayant un faible taux de ErbB2, la fixation de HRG (héréguline) au récepteur ErbB3 induit un passage de la conformation fermée à l'état ouvert, exposant le bras de dimérisation du domaine II et conduisant à la formation de l'hétérodimère ErbB2/ErbB3. La fixation du pertuzumab sur le domaine II de ErbB2 empêche l'hétérodimérisation induite par le ligand et la signalisation intracellulaire (à droite). De l'autre côté, la fixation du trastuzumab au domaine IV de ErbB2 ne joue pas sur l'association ErbB2-ErbB3 induite par la fixation du ligand HRG et donc n'interrompt pas l'activité de signalisation et la prolifération des cellules cancéreuses (à gauche). **B :** Les cellules surexprimant le récepteur ErbB2 montrent une signalisation constitutive de ErbB2 et ErbB3. In vivo, les deux anticorps ont une activité antitumorale qui pourrait impliquer plusieurs mécanismes. Le trastuzumab empêche le relargage de ErbB2 (Molina et al. 2001), diminue la signalisation de ErbB2 et ErbB3 et inhibe la signalisation intracellulaire (à gauche). Alors que le pertuzumab n'affecte pas le relargage de ErbB2, d'autres activités antitumorales in vivo demeurent encore non élucidées (chemin indiqué par des pointillés sur l'image de droite). Il est possible que le pertuzumab bloque la formation des homo- et/ou hétéro-dimères contenant le récepteur ErbB2 et conduise à l'inhibition des cellules tumorales. Dans les cellules surexprimant le récepteur ErbB2, le rôle du bras de dimérisation du domaine II de ErbB2 dans le formation des dimères n'est pas encore élucidé. D'après (Badache and Goncalves 2006; Badache and Hynes 2004).

Médicament	Type		Statuts	Remarques
<b><u>Anticorps Monoclonaux</u></b>				
Cetuximab (Erbix)	Chimérique	ErbB1	Approuvé pour les cancers du CRC, de la tête et du cou en combinaison avec la chimiothérapie et la radiothérapie respectivement. Essais en cours en combinaison avec d'autres thérapies sur les cancers de NSCLC, des cellules de la peau, de la tête et du cou, du pancréas, du sein et cervical.	
Trastuzumab (Herceptine)	Humanisé	ErbB2	Approuvé pour les cancers du sein en combinaison avec la chimiothérapie. Essais en cours en combinaison avec d'autres thérapies sur les cancers des glandes salivaires, de l'endomètre, des ovaires, de la vessie.	
Pertuzumab	Humanisé	ErbB2	Essais en cours en combinaison avec d'autres thérapies sur les tumeurs exprimant un faible taux de ErbB2.	Un inhibiteur de dimérisation.
Panitumumab (Vectibix, ABX-EGF)	Totalement humain	ErbB1	Approuvé pour les cancers du CRC en combinaison avec la chimiothérapie. Essais en cours en combinaison avec d'autres thérapies sur les cancers du colon et du rectum. Essais en cours pour des combinaisons avec d'autres thérapies pour les NSCLC.	1 <sup>er</sup> anticorps totalement humain approuvé pour les cancers
Matuzumab (EMD-7200)	Humanisé	ErbB1	Essais en cours pour des combinaisons avec d'autres thérapies pour les NSCLC, les cancers gastrique et de l'œsophage.	
Nimotuzumab (TheraCim)	Humanisé	ErbB1	Essais en cours pour des combinaisons avec d'autres thérapies pour les NSCLC.	
MDX214	Totalement humain	ErbB1	Essais pour les NSCLC.	
<b><u>Inhibiteurs de TK</u></b>				
Gefitinib (Iressa, ZD1839)	TKI	ErbB1	Approuvé pour les NSCLC. Essais en cours pour des combinaisons avec d'autres thérapies pour les cancers gastro-intestinal, du sein, de la vessie et de la tête et du cou.	
Erlotinib (Tarceva)	TKI	ErbB1	Approuvé pour les NSCLC et les cancers du pancréas. Essais en cours pour des combinaisons avec d'autres thérapies pour les cancers CRC et les gliomes.	
Lapatinib	TKI	ErbB1 et ErbB2	Essais pour les cancers du sein	
CI-1033	TKI	Pan-ErbB	Essais pour les NSCLC	Irreversible
EKB-569	TKI	ErbB1 et ErbB2	Essais pour les NSCLC	Irreversible
AEE-788	TKI	ErbB1, ErbB2 et VEGFR	Essais en cours pour des combinaisons avec d'autres thérapies pour les glioblastomes.	1 <sup>er</sup> inhibiteur multifonction

**Tableau 1.3 : Résumé des différents médicaments approuvés et/ou en essais ciblant les récepteurs de la famille ErbB dans les cancers.** (Bubril and Yarden 2007).

Les RTK sont des protéines transmembranaires à multiple fonctions qui sont exprimées à la surface cellulaire : ces caractéristiques font d'eux des cibles présentant des sites d'attaque multiples et accessibles.



En effet, l'identification d'un (ou de plusieurs) inhibiteurs ciblant un ou plusieurs membres de la famille ErbB, soit au niveau extracellulaire, soit au niveau de la régulation des cascades de signalisation, pourrait augmenter l'efficacité thérapeutique. Bien que la recherche de nouveaux traitements contre le cancer vise essentiellement la sous-famille ErbB, il est raisonnable de penser que la plupart des RTK oncogéniques pourront constituer une cible privilégiée dans les années à venir.

## **BIBLIOGRAPHIE ANNEXE 1**

- Abd El-Rehim DM, Pinder SE, Paish CE, Bell JA, Rampaul RS, Blamey RW, Robertson JF, Nicholson RI, Ellis IO (2004) Expression and co-expression of the members of the epidermal growth factor receptor (EGFR) family in invasive breast carcinoma. *Br J Cancer* 91:1532-42
- Alimandi M, Romano A, Curia MC, Muraro R, Fedi P, Aaronson SA, Di Fiore PP, Kraus MH (1995) Cooperative signaling of ErbB3 and ErbB2 in neoplastic transformation and human mammary carcinomas. *Oncogene* 10:1813-21
- Bacus SS, Chin D, Yarden Y, Zelnick CR, Stern DF (1996) Type 1 receptor tyrosine kinases are differentially phosphorylated in mammary carcinoma and differentially associated with steroid receptors. *Am J Pathol* 148:549-58
- Bacus SS, Zelnick CR, Chin DM, Yarden Y, Kaminsky DB, Bennington J, Wen D, Marcus JN, Page DL (1994) Medullary carcinoma is associated with expression of intercellular adhesion molecule-1. Implication to its morphology and its clinical behavior. *Am J Pathol* 145:1337-48
- Badache A, Goncalves A (2006) The ErbB2 signaling network as a target for breast cancer therapy. *J Mammary Gland Biol Neoplasia* 11:13-25
- Badache A, Hynes NE (2004) A new therapeutic antibody masks ErbB2 to its partners. *Cancer Cell* 5:299-301
- Badra FA, Karamouzis MV, Ravazoula P, Likaki-Karatza E, Tzorakoleftherakis E, Koukouras D, Iconomou G, Xiros N, Siablis D, Papavassiliou AG, Kalofonos HP (2006) Non-palpable breast carcinomas: correlation of mammographically detected malignant-appearing microcalcifications and epidermal growth factor receptor (EGFR) family expression. *Cancer Lett* 244:34-41
- Bargmann CI, Weinberg RA (1988) Oncogenic activation of the neu-encoded receptor protein by point mutation and deletion. *EMBO J* 7:2043-2052
- Barnes NL, Khavari S, Boland GP, Cramer A, Knox WF, Bundred NJ (2005) Absence of HER4 expression predicts recurrence of ductal carcinoma in situ of the breast. *Clin Cancer Res* 11:2163-8
- Bates SE, Davidson NE, Valverius EM, Freter CE, Dickson RB, Tam JP, Kudlow JE, Lippman ME, Salomon DS (1988) Expression of transforming growth factor alpha

- and its messenger ribonucleic acid in human breast cancer: its regulation by estrogen and its possible functional significance. *Mol Endocrinol* 2:543-55
- Bei R, Masuelli L, Moriconi E, Visco V, Moretti A, Kraus MH, Muraro R (1999) Immune responses to all ErbB family receptors detectable in serum of cancer patients. *Oncogene* 18:1267-75
- Blume-Jensen P, Hunter T (2001) Oncogenic kinase signalling. *Nature* 411:355-65
- Brabender J, Danenberg KD, Metzger R, Schneider PM, Park J, Salonga D, Holscher AH, Danenberg PV (2001) Epidermal growth factor receptor and HER2-neu mRNA expression in non-small cell lung cancer Is correlated with survival. *Clin Cancer Res* 7:1850-5
- Bublil EM, Yarden Y (2007) The EGF receptor family: spearheading a merger of signaling and therapeutics. *Curr Opin Cell Biol* 19:124-34
- Carpenter G (1999) Employment of the epidermal growth factor receptor in growth factor-independent signaling pathways. *J Cell Biol* 146:697-702
- Carpenter G (2003) ErbB-4: mechanism of action and biology. *Exp Cell Res* 284:66-77
- Carraway KL, Cantley LC (1994) A neu acquaintance for erbB3 and erbB4: a role for receptor heterodimerization in growth signaling. *Cell* 78:5-8
- Chu CT, Everiss KD, Wikstrand CJ, Batra SK, Kung HJ, Bigner DD (1997) Receptor dimerization is not a factor in the signalling activity of a transforming variant epidermal growth factor receptor (EGFRvIII). *Biochem J* 324 (Pt 3):855-61
- Ciardello F (2000) Epidermal growth factor receptor tyrosine kinase inhibitors as anticancer agents. *Drugs* 60 Suppl 1:25-32; discussion 41-2
- Ciardello F, Tortora G (2001) A novel approach in the treatment of cancer: targeting the epidermal growth factor receptor. *Clin Cancer Res* 7:2958-70
- Ciardello F, Tortora G (2003) Epidermal growth factor receptor (EGFR) as a target in cancer therapy: understanding the role of receptor expression and other molecular determinants that could influence the response to anti-EGFR drugs. *Eur J Cancer* 39:1348-54
- Citri A, Skaria KB, Yarden Y (2003) The deaf and the dumb: the biology of ErbB-2 and ErbB-3. *Exp Cell Res* 284:54-65
- Citri A, Yarden Y (2006) EGF-ERBB signalling: towards the systems level. *Nat Rev Mol Cell Biol* 7:505-16
- Cohen MH, Williams GA, Sridhara R, Chen G, McGuinn WD, Jr., Morse D, Abraham S, Rahman A, Liang C, Lostritto R, Baird A, Pazdur R (2004) United States Food and

- Drug Administration Drug Approval summary: Gefitinib (ZD1839; Iressa) tablets. Clin Cancer Res 10:1212-8
- Cohen P (2002) Protein kinases--the major drug targets of the twenty-first century? Nat Rev Drug Discov 1:309-15
- Di Fiore PP, Pierce JH, Kraus MH, Segatto O, King CR, Aaronson SA (1987) erbB-2 is a potent oncogene when overexpressed in NIH/3T3 cells. Science 237:178-82
- DiGiovanna MP, Stern DF, Edgerton SM, Whalen SG, Moore D, 2nd, Thor AD (2005) Relationship of epidermal growth factor receptor expression to ErbB-2 signaling activity and prognosis in breast cancer patients. J Clin Oncol 23:1152-60
- Distler JH, Hirth A, Kurowska-Stolarska M, Gay RE, Gay S, Distler O (2003) Angiogenic and angiostatic factors in the molecular control of angiogenesis. Q J Nucl Med 47:149-61
- Downward J, Yarden Y, Mayes E, Scrace G, Totty N, Stockwell P, Ullrich A, Schlessinger J, Waterfield MD (1984) Close similarity of epidermal growth factor receptor and v-erbB oncogene protein sequences. Nature 307:521-7
- Fernandes H, Cohen S, Bishayee S (2001) Glycosylation-induced conformational modification positively regulates receptor-receptor association: a study with an aberrant epidermal growth factor receptor (EGFRvIII/DeltaEGFR) expressed in cancer cells. J Biol Chem 276:5375-83
- Folkman J (2002) Role of angiogenesis in tumor growth and metastasis. Semin Oncol 29:15-8
- Frederick L, Wang XY, Eley G, James CD (2000) Diversity and frequency of epidermal growth factor receptor mutations in human glioblastomas. Cancer Res 60:1383-7
- Furstenberger G, Senn HJ (2002) Insulin-like growth factors and cancer. Lancet Oncol 3:298-302
- Garcia de Palazzo IE, Adams GP, Sundareshan P, Wong AJ, Testa JR, Bigner DD, Weiner LM (1993) Expression of mutated epidermal growth factor receptor by non-small cell lung carcinomas. Cancer Res 53:3217-20
- Gilbertson RJ, Perry RH, Kelly PJ, Pearson AD, Lunec J (1997) Prognostic significance of HER2 and HER4 coexpression in childhood medulloblastoma. Cancer Res 57:3272-80
- Gilmour LM, Macleod KG, McCaig A, Gullick WJ, Smyth JF, Langdon SP (2001) Expression of erbB-4/HER-4 growth factor receptor isoforms in ovarian cancer. Cancer Res 61:2169-76

- Go RC, Perry RT, Wiener H, Bassett SS, Blacker D, Devlin B, Sweet RA (2005) Neuregulin-1 polymorphism in late onset Alzheimer's disease families with psychoses. *Am J Med Genet B Neuropsychiatr Genet* 139:28-32
- Grandis JR, Melhem, M. F., Gooding, W. E., Day, R., Holst, V. A., Wagener, M. M., Drenning SDeT, D. J. (1998) Levels of TGF-alpha and EGFR protein in head and neck squamous cell carcinoma and patient survival. *J Natl Cancer Inst* 90:824-32
- Gschwind A, Fischer OM, Ullrich A (2004) The discovery of receptor tyrosine kinases: targets for cancer therapy. *Nat Rev Cancer* 4:361-70
- Gullick WJ (1996) The c-erbB3/HER3 receptor in human cancer. *Cancer Surv* 27:339-49
- Guy PM, Platko JV, Cantley LC, Cerione RA, Carraway KL, 3rd (1994) Insect cell-expressed p180erbB3 possesses an impaired tyrosine kinase activity. *Proc Natl Acad Sci U S A* 91:8132-6
- Hanahan D, Weinberg RA (2000) The hallmarks of cancer. *Cell* 100:57-70
- Holbro T, Civenni G, Hynes NE (2003) The ErbB receptors and their role in cancer progression. *Exp Cell Res* 284:99-110
- Huang HS, Nagane M, Klingbeil CK, Lin H, Nishikawa R, Ji XD, Huang CM, Gill GN, Wiley HS, Cavenee WK (1997) The enhanced tumorigenic activity of a mutant epidermal growth factor receptor common in human cancers is mediated by threshold levels of constitutive tyrosine phosphorylation and unattenuated signaling. *J Biol Chem* 272:2927-35
- Hudziak RM, Schlessinger J, Ullrich A (1987) Increased expression of the putative growth factor receptor p185HER2 causes transformation and tumorigenesis of NIH 3T3 cells. *Proc Natl Acad Sci U S A* 84:7159-63
- Hunter T (2000) Signaling-2000 and beyond. *Cell* 100:113-127
- Hynes NE, Stern DF (1994) The biology of erbB-2/neu/HER-2 and its role in cancer. *Biochim. Biophys. Acta* 1198:165-184
- Jackson JG, St Clair P, Sliwkowski MX, Brattain MG (2004) Blockade of epidermal growth factor- or heregulin-dependent ErbB2 activation with the anti-ErbB2 monoclonal antibody 2C4 has divergent downstream signaling and growth effects. *Cancer Res* 64:2601-9
- Jeffers M, Fiscella M, Webb CP, Anver M, Koochekpour S, Vande Woude GF (1998) The mutationally activated Met receptor mediates motility and metastasis. *Proc Natl Acad Sci U S A* 95:14417-22

- Karamouzis MV, Badra FA, Papavassiliou AG (2007) Breast cancer: the upgraded role of HER-3 and HER-4. *Int J Biochem Cell Biol* 39:851-6
- Kita YA, Barff J, Luo Y, Wen D, Brankow D, Hu S, Liu N, Prigent SA, Gullick WJ, Nicolson M (1994) NDF/hereregulin stimulates the phosphorylation of Her3/erbB3. *FEBS Lett* 349:139-43
- Klapper LN, Kirschbaum MH, Sela M, Yarden Y (2000) Biochemical and clinical implications of the ErbB/HER signaling network of growth factor receptors. *Adv Cancer Res* 77:25-79
- Klijn JG, Berns PM, Schmitz PI, Foekens JA (1992) The clinical significance of epidermal growth factor receptor (EGF-R) in human breast cancer: a review on 5232 patients. *Endocr Rev* 13:3-17
- Kuan CT, Wikstrand CJ, Bigner DD (2001) EGF mutant receptor vIII as a molecular target in cancer therapy. *Endocr Relat Cancer* 8:83-96
- Kute T, Lack CM, Willingham M, Bishwokama B, Williams H, Barrett K, Mitchell T, Vaughn JP (2004) Development of Herceptin resistance in breast cancer cells. *Cytometry A* 57:86-93
- Lee JC, Wang ST, Chow NH, Yang HB (2002) Investigation of the prognostic value of coexpressed erbB family members for the survival of colorectal cancer patients after curative surgery. *Eur J Cancer* 38:1065-71
- Levitcki A (1999) Protein tyrosine kinase inhibitors as novel therapeutic agents. *Pharmacol Ther* 82:231-9
- Lonardo F, Di Marco E, King CR, Pierce JH, Segatto O, Aaronson SA, Di Fiore PP (1990) The normal erbB-2 product is an atypical receptor-like tyrosine kinase with constitutive activity in the absence of ligand. *New Biol* 2:992-1003
- Menendez JA, Mehmi I, Lupu R (2006) Trastuzumab in combination with heregulin-activated Her-2 (erbB-2) triggers a receptor-enhanced chemosensitivity effect in the absence of Her-2 overexpression. *J Clin Oncol* 24:3735-46
- Merimsky O, Issakov J, Bickels J, Kollender Y, Flusser G, Soyfer V, Schwartz I, Inbar M, Meller I (2002) ErbB-4 expression in limb soft-tissue sarcoma: correlation with the results of neoadjuvant chemotherapy. *Eur J Cancer* 38:1335-42
- Molina MA, Codony-Servat J, Albanell J, Rojo F, Arribas J, Baselga J (2001) Trastuzumab (herceptin), a humanized anti-Her2 receptor monoclonal antibody, inhibits basal and activated Her2 ectodomain cleavage in breast cancer cells. *Cancer Res* 61:4744-9

- Moscattello DK, Holgado-Madruga M, Emlet DR, Montgomery RB, Wong AJ (1998) Constitutive activation of phosphatidylinositol 3-kinase by a naturally occurring mutant epidermal growth factor receptor. *J Biol Chem* 273:200-6
- Normanno N, Bianco C, De Luca A, Maiello MR, Salomon DS (2003) Target-based agents against ErbB receptors and their ligands: a novel approach to cancer treatment. *Endocr Relat Cancer* 10:1-21
- Normanno N, Bianco C, De Luca A, Salomon DS (2001) The role of EGF-related peptides in tumor growth. *Front Biosci* 6:D685-707
- Normanno N, Ciardiello F (1997) EGF-related peptides in the pathophysiology of the mammary gland. *J Mammary Gland Biol Neoplasia* 2:143-51
- Normanno N, De Luca A, Bianco C, Strizzi L, Mancino M, Maiello MR, Carotenuto A, De Feo G, Caponigro F, Salomon DS (2006) Epidermal growth factor receptor (EGFR) signaling in cancer. *Gene* 366:2-16
- Normanno N, Maiello MR, De Luca A (2003) Epidermal growth factor receptor tyrosine kinase inhibitors (EGFR-TKIs): simple drugs with a complex mechanism of action? *J Cell Physiol* 194:13-9
- Pawson T, Nash P (2000) Protein-protein interactions define specificity in signal transduction. *Genes Dev* 14:1027-47
- Pedersen MW, Meltorn M, Damstrup L, Poulsen HS (2001) The type III epidermal growth factor receptor mutation. Biological significance and potential target for anti-cancer therapy. *Ann Oncol* 12:745-60
- Plowman GD, Culouscou JM, Whitney GS, Green JM, Carlton GW, Foy L, Neubauer MG, Shoyab M (1993) Ligand-specific activation of HER4/p180erbB4, a fourth member of the epidermal growth factor receptor family. *Proc. Natl. Acad. Sci. U S A* 90:1746-50
- Prenzel N, Zwick E, Leserer M, Ullrich A (2000) Tyrosine kinase signalling in breast cancer. Epidermal growth factor receptor: convergence point for signal integration and diversification. *Breast Cancer Res* 2:184-90
- Prigent SA, Gullick WJ (1994) Identification of c-erbB-3 binding sites for phosphatidylinositol 3'-kinase and SHC using an EGF receptor/c-erbB-3 chimera. *Embo J* 13:2831-41
- Prigent SA, Lemoine NR, Hughes CM, Plowman GD, Selden C, Gullick WJ (1992) Expression of the c-erbB-3 protein in normal human adult and fetal tissues. *Oncogene* 7:1273-8
- Reed JC (1999) Dysregulation of apoptosis in cancer. *J Clin Oncol* 17:2941-53

- Robertson SC, Tynan J, Donoghue DJ (2000) RTK mutations and human syndromes: when good receptors turn bad. *Trends Genet* 16:368
- Roskoski R, Jr. (2004) The ErbB/HER receptor protein-tyrosine kinases and cancer. *Biochem Biophys Res Commun* 319:1-11
- Roy V, Perez EA (2006) New therapies in the treatment of breast cancer. *Semin Oncol* 33:S3-8
- Salomon DS, Brandt R, Ciardiello F, Normanno N (1995) Epidermal growth factor-related peptides and their receptors in human malignancies. *Crit Rev Oncol Hematol* 19:183-232
- Sartor CI (2000) Biological modifiers as potential radiosensitizers: targeting the epidermal growth factor receptor family. *Semin Oncol* 27:15-20; discussion 92-100
- Schechter AL, Stern DF, Vaidyanathan L, Decker SJ, Drebin JA, Greene MI, Weinberg RA (1984) The neu oncogene: an erb-B-related gene encoding a 185,000-Mr tumour antigen. *Nature* 312:513-6
- Schwechheimer K, Huang S, Cavenee WK (1995) EGFR gene amplification--rearrangement in human glioblastomas. *Int J Cancer* 62:145-8
- Shinojima N, Tada K, Shiraishi S, Kamiryo T, Kochi M, Nakamura H, Makino K, Saya H, Hirano H, Kuratsu J, Oka K, Ishimaru Y, Ushio Y (2003) Prognostic value of epidermal growth factor receptor in patients with glioblastoma multiforme. *Cancer Res* 63:6962-70
- Siegel PM, Ryan ED, Cardiff RD, Muller WJ (1999) Elevated expression of activated forms of Neu/ErbB-2 and ErbB-3 are involved in the induction of mammary tumors in transgenic mice: implications for human breast cancer. *Embo J* 18:2149-64
- Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL (1987) Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* 235:177-82
- Slamon DJ, Godolphin W, Jones LA, Holt JA, Wong SG, Keith DE, Levin WJ, Stuart SG, Udove J, Ullrich A, et al. (1989) Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. *Science* 244:707-12
- Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, Fleming T, Eiermann W, Wolter J, Pegram M, Baselga J, Norton L (2001) Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N Engl J Med* 344:783-92



- Srinivasan R, Benton E, McCormick F, Thomas H, Gullick WJ (1999) Expression of the c-erbB-3/HER-3 and c-erbB-4/HER-4 growth factor receptors and their ligands, neuregulin-1 alpha, neuregulin-1 beta, and betacellulin, in normal endometrium and endometrial cancer. *Clin Cancer Res* 5:2877-83
- Srinivasan R, Gillett CE, Barnes DM, Gullick WJ (2000) Nuclear expression of the c-erbB-4/HER-4 growth factor receptor in invasive breast cancers. *Cancer Res* 60:1483-7
- Stephens P, Hunter C, Bignell G, Edkins S, Davies H, Teague J, Stevens C, O'Meara S, Smith R, Parker A, Barthorpe A, Blow M, Brackenbury L, Butler A, Clarke O, Cole J, Dicks E, Dike A, Drozd A, Edwards K, Forbes S, Foster R, Gray K, Greenman C, Halliday K, Hills K, Kosmidou V, Lugg R, Menzies A, Perry J, Petty R, Raine K, Ratford L, Shepherd R, Small A, Stephens Y, Tofts C, Varian J, West S, Widaa S, Yates A, Brasseur F, Cooper CS, Flanagan AM, Knowles M, Leung SY, Louis DN, Looijenga LH, Malkowicz B, Pierotti MA, Teh B, Chenevix-Trench G, Weber BL, Yuen ST, Harris G, Goldstraw P, Nicholson AG, Futreal PA, Wooster R, Stratton MR (2004) Lung cancer: intragenic ERBB2 kinase mutations in tumours. *Nature* 431:525-6
- Tateishi M, Ishida T, Mitsudomi T, Kaneko S, Sugimachi K (1990) Immunohistochemical evidence of autocrine growth factors in adenocarcinoma of the human lung. *Cancer Res* 50:7077-80
- Tsai MS, Shamon-Taylor LA, Mehmi I, Tang CK, Lupu R (2003) Blockage of heregulin expression inhibits tumorigenicity and metastasis of breast cancer. *Oncogene* 22:761-8
- Tzahar E, Yarden Y (1998) The ErbB-2/HER2 oncogenic receptor of adenocarcinomas: from orphanhood to multiple stromal ligands. *Biochim Biophys Acta* 1377:M25-37
- Ullrich A, Coussens L, Hayflick JS, Dull TJ, Gray A, Tam AW, Lee J, Yarden Y, Libermann TA, Schlessinger J, et al (1984) Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. *Nature* 309:418-25
- Varmus H, Bishop JM (1986) Biochemical mechanisms of oncogene activity: proteins encoded by oncogenes. Introduction. *Cancer Surv* 5:153-8
- Weiner DB, Liu J, Cohen JA, Williams WV, Greene MI (1989) A point mutation in the neu oncogene mimics ligand induction of receptor aggregation. *Nature* 339:230-231
- Wikstrand CJ, Reist CJ, Archer GE, Zalutsky MR, Bigner DD (1998) The class III variant of the epidermal growth factor receptor (EGFRvIII): characterization and utilization as an immunotherapeutic target. *J Neurovirol* 4:148-58

- Willem M, Garratt AN, Novak B, Citron M, Kaufmann S, Rittger A, DeStrooper B, Saftig P, Birchmeier C, Haass C (2006) Control of peripheral nerve myelination by the beta-secretase BACE1. *Science* 314:664-6
- Xu H, Yu Y, Marciniak D, Rishi AK, Sarkar FH, Kucuk O, Majumdar AP (2005) Epidermal growth factor receptor (EGFR)-related protein inhibits multiple members of the EGFR family in colon and breast cancer cells. *Mol Cancer Ther* 4:435-42
- Yarden Y, Sliwkowski MX (2001) Untangling the ErbB signalling network. *Nat Rev Mol Cell Biol* 2:127-37
- Zhang K, Sun J, Liu N, Wen D, Chang D, Thomason A, Yoshinaga SK (1996) Transformation of NIH 3T3 cells by HER3 or HER4 receptors requires the presence of HER1 or HER2. *J Biol Chem* 271:3884-90

**ANNEXE 2 : ETUDE DU DOMAINE TM DU RECEPTEUR NEU PROTO-  
ONCOGENIQUE ET ONCOGENIQUE (*article JBSD 2005*).**

## Introduction

Les données expérimentales obtenues sur les récepteurs Neu (récepteur homologue de ErbB2 chez le rat) sous ses formes sauvage et mutée (Val664/Glu) ne sont pas suffisantes pour décrire la structure du domaine transmembranaire dans le récepteur actif.

Ainsi, afin de mettre en évidence d'éventuelles modifications structurales et / ou dynamiques entre la forme sauvage Neu<sup>V</sup><sub>TM</sub> et la forme mutée Neu<sup>E</sup><sub>TM</sub> du peptide transmembranaire nous avons effectué des simulations de dynamique moléculaire sur le domaine TM du récepteur Neu, dans un modèle explicite de membrane composé de molécules de DMPC.

Les résultats obtenus ont donné lieu à une publication dans le Journal of Biomolecular Structure and Dynamics (JBSD) en 2005.

## **Transmembrane Peptides from Tyrosine Kinase Receptor. Mutation-related Behavior in a Lipid Bilayer Investigated by Molecular Dynamics Simulations**

**Oumarou Samna Soumana<sup>§</sup>  
Pierre Aller<sup>§</sup>  
Norbert Garnier  
Monique Genest<sup>\*</sup>**

<http://www.jbsdonline.com>

Centre de Biophysique Moléculaire  
UPR 4301, CNRS

rue Charles Sadron  
45071 Orléans Cedex 02

Affiliated to the University of Orléans  
France.

<sup>§</sup>Participate equally to this work.

### **Abstract**

Polar mutations in transmembrane  $\alpha$  helices may alter the structural details of the hydrophobic sequences and control intermolecular contacts. We have performed molecular dynamics simulations on the transmembrane domain of the proto-oncogenic and the oncogenic forms of the Neu receptor in a fluid DMPC bilayer to test whether the Glu mutation which replaces the Val residue at position 664 may alter the helical structure and its insertion in the membrane. The simulations show that the wild and the mutant forms of the transmembrane domain have a different behavior in the bilayer. The native transmembrane sequence is found to be more flexible than in the presence of the Glu mutation, characterized by a tendency to  $\pi$  deformation to accommodate the helix length to the membrane thickness. The mutant form of this domain does not evidence helical deformation in the present simulation. Hydrophobic matching is achieved both by a larger helix tilt and a vertical shift of the helix towards the membrane interface, favoring the accessibility of the Glu side chain to the membrane environment. A rapid exchange of hydrogen bond interactions with the surrounding water molecules and the lipid headgroups is observed. The difference in the behavior between the two peptides in a membrane environment was also observed experimentally. Both simulation and experimental results agree with the hypothesis that water may act as an intermediate for the formation of cross links between the facing Glu side chains stabilizing the dimer.

Key words: Tyrosine kinase receptor; Transmembrane helix; MD simulations; Membrane environment; and Helix insertion.

### **Introduction**

Understanding the interactions that govern oligomerization of transmembrane domains of membrane proteins is of great importance for elucidating biological structure relationships. Peptide-lipid interactions may affect both the folding of membrane protein and the bilayer structure, and can serve to promote peptide-peptide association and protein activity (1-3). Membrane protein folding is described as the insertion of individual helices into a membrane followed by side-to-side helix interactions to form stable functional structures (4). These direct homodimer or heterodimer interactions are the central model for receptor tyrosine kinase (RTK) signaling (5, 6). Transmembrane helix association is essentially governed by van der Waals forces between complementary helix faces but polar residues in transmembrane  $\alpha$  helices may strongly influence their association (7-9). The formation of strong inter helical hydrogen bonds in membranes may play an essential role in the stability of association but could increase unregulated packing with dramatic biological consequences (10). A famous example is the replacement of a valine residue by a glutamic acid residue in the transmembrane domain of Neu which enhances receptor dimerization, leading to the constitutive activation of the

\*Phone: 33 (0)2 38 25 55 68  
Fax : 33 (0)2 38 63 15 17  
Email: m.genest@cnrs-orleans.fr

protein and to a lethal phenotype (11-13). The Glu mutation is believed to control intermolecular contacts between the transmembrane helices and alter the structural characteristics of the native hydrophobic sequence. These structural details may also be modulated by the membrane environment (14, 15). Hence, a detailed examination of the behavior of the transmembrane monomer in a bilayer environment could provide a mechanism for modulating monomer association.

Because of the biological importance of the transmembrane domain of the Neu/ErbB2 receptor, we and others have focused on this system in the hope of understanding its regulatory role in receptor dimerization and activation. Significant background information exists on this domain. The picture that emerges from the combination of theoretical and experimental techniques is that the structure of the transmembrane segment remains intrinsically helical in organic solvent and in a membrane-like environment (16-22) but that helicity can be modulated by  $\pi$  local defects (21, 23-27).

Studies of the Glu mutation effect on the structure and the dynamics of the transmembrane domain of the rodent Neu receptor and its human homologous ErbB2 using NMR spectroscopy have shown that the two mutant peptides are more conformationally constrained in a fluid lipid bilayer than the wild peptides suggesting a greater tendency to form oligomers (17, 19). Differences in flexibility were also found from molecular dynamics (MD) simulations (22).

More recently, amide exchange experiments on transmembrane Neu peptides in DPC (dodecylphosphocholine) micelles (28) have demonstrated that the proto-oncogenic and the oncogenic forms have a different influence on the membrane environment and that this influence is more pronounced in the N-terminal region of the peptide. The major effects observed at Ala661 and Gly665 residues close to the mutation are thought to be due to a different level of solvent exposure or to helical destabilization at these sites. However, as data acquired in detergent micelles are not sufficient to distinguish between these effects, no clear conclusion can be drawn. Atomic details on the transmembrane peptides are needed in order to interpret experimental results and computer simulations may contribute to their understanding.

In this report, we present a study of the transmembrane domain of the Neu receptor in the proto-oncogenic and the oncogenic forms embedded within a fully hydrated DMPC (dimyristoylphosphatidylcholine) bilayer. Simulation results reveal that, for the two peptides, the hydrophobic region well inserted in the hydrophobic core of the bilayer remains close to the helical structure. However, the native form of the transmembrane domain appears more flexible than the oncogenic form and shows a strong tendency for  $\pi$  deformation. Differences in helix insertion into the bilayer between the two peptides are evidenced, due to the presence of the mutation and its influence on the membrane environment. Simulation data fully support the experimental results and suggest that a water environment at the N-terminus might have a role in the dimerization process.

### ***Material and Methods***

#### *Simulation Conditions*

Calculations were performed using CHARMM Version 26 (29) and the all hydrogen PARAM22 force field (30, 31), which includes the phospholipid library (32) and the TIP3P water potential (33). The SHAKE algorithm (34) was applied to fix all hydrogen atom bond distances and the time step for integration was 2 fs. The dimensions of the simulation box are 48 Å, 48 Å, and 72 Å along the x, y, and z axes respectively. The z axis is oriented along the normal to the membrane plane (xy plane) and the center of the bilayer is at z = 0. Periodic boundary conditions in the three dimensions were used and the temperature was set to 330K, well above the gel/liquid crys-

tal transition temperature of pure DMPC (35) to ensure that the bilayer was in the liquid crystalline phase. The same temperature was used in other simulations of proteins inserted into DMPC bilayers with convincing results (36-38). A high temperature will also increase atomic fluctuations (39), leading to a better exploration of dynamical events occurring over a short simulation time. The cross-sectional area per lipid is  $64 \text{ \AA}^2$  (40), each being hydrated by 40 water molecules (41).

### *Microscopic System*

The microscopic system consists of one canonical  $\alpha$  helix, 68 lipids (34 in each layer), and about 2780 water molecules. The helix comprises 35 residues including the putative hydrophobic transmembrane region Val656-Ile680 flanked by 6 juxtamembrane residues at the N-terminus and 4 basic residues at the C-terminus serving as signal sequence for membrane anchoring at the cytoplasmic side. The peptide sequence was capped by a methyl acetamide group and a methyl amide group at the N- and the C-terminal ends respectively. The sequence of the transmembrane domain used in the simulations is the following:



Glu and Arg at the N-terminal extremity and Lys and Arg at the C-terminus are charged residues and 6 counter ions (1  $\text{Na}^+$ , 5  $\text{Cl}^-$ ) were added to ensure the neutrality of the transmembrane peptide. In the low dielectric membrane environment a Glu residue most likely has a protonated carboxylic group (42) and this was confirmed from experimental data (19, 20). Glu664 within the hydrophobic sequence was then modeled in a protonated state.

Incorporation of the lipids into the simulation box follows the protocol used in previous simulations (36, 43), and is based on that developed by Woolf and Roux (44-46). Succinctly, the lipid headgroups represented by Lennard Jones spheres were placed at  $\pm 17 \text{ \AA}$  along the z direction to respect the membrane thickness (47). After optimization of their position in the membrane plane they were replaced by the all atom model of the DMPC molecules. Random configurations of pre-hydrated lipids were chosen in the library available in CHARMM (48). Lipid rotation and translation were successively applied during energy minimization steps to relax bad contacts. Then pre-equilibrated cubes of water were placed at the top and the bottom of the lipid box containing the peptide to obtain the desired level of hydration of the lipids. Successive minimizations and MD runs were applied to give the initial configuration of the peptide/DMPC system used as starting point for the simulation.

Two peptide/DMPC systems were built separately, one including the native Neu transmembrane domain and the other including the oncogenic form corresponding to the Val664Glu mutation. These two systems will be referred to as  $\text{Neu}_{\text{TM}}^{\text{V}}/\text{DMPC}$  and  $\text{Neu}_{\text{TM}}^{\text{E}}/\text{DMPC}$  respectively.

### *MD Simulations*

The  $\text{Neu}_{\text{TM}}^{\text{V}}/\text{DMPC}$  and  $\text{Neu}_{\text{TM}}^{\text{E}}/\text{DMPC}$  systems have been simulated over a time scale of 4.6 ns and 4.0 ns respectively after the construction. First, each system was equilibrated during 1.0 ns in the microcanonical ensemble (NVE). Non bonded interactions were truncated at  $11 \text{ \AA}$  and electrostatic interactions were switched off over a distance of  $3 \text{ \AA}$ . To prevent global drift of the lipids, harmonic restraints were applied to the center of mass of the DMPC molecules in the z direction using a force constant of  $5.0 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^{-2}$ . These restraints are sufficiently weak to allow small changes in the bilayer thickness. Also, soft cylindrical restraints (force constant of  $5 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^{-2}$ ) were applied to the center of mass of the peptide in order to prevent too large a lateral drift while still allowing reasonable helix motions. Major helix distortions were observed at the beginning of the simulation of the mutant pep-

tide. To avoid the complete break of the helical structure harmonic restraints were used to maintain the  $\alpha$  helical hydrogen bonding pattern. The force constant value was progressively reduced from 50 to 0 kcal·mol<sup>-1</sup>·Å<sup>-2</sup> at the end of the NVE period. Helical distortions were not observed for the native Neu transmembrane peptide and the NVE period was pursued without any hydrogen bond constraint.

After the equilibration phase, 3.6 ns and 3.0 ns of simulation were carried out at constant normal pressure ( $P_n$ ) and constant surface area of the central box (xy plane) (A) ( $NP_nAT$  ensemble) for the Neu<sup>V</sup><sub>TM</sub>/DMPC and Neu<sup>E</sup><sub>TM</sub>/DMPC systems respectively. The extended system method described by Hoover (49) and implemented in CHARMM was used. Normal pressure  $P_n$ , pressure piston mass, surface area A and the Langevin piston collision frequency were set to 1 atm, 750 amu., (48 × 48) Å<sup>2</sup> and 5 ps<sup>-1</sup> respectively. The thermal piston mass was equal to 3000 kcal·ps<sup>2</sup>. During the whole production trajectory the soft restraints applied to the DMPC molecules and to the center of mass of the transmembrane helix were maintained. No other restraint was applied during this period. Electrostatic interactions were calculated using the Particle Mesh Ewald method (50) implemented in CHARMM. The control of the geometrical parameters of the helix was performed over different time ranges of the  $NP_nAT$  simulation, giving rise to very similar results. We have chosen to present the results obtained over the last ns of the simulation.

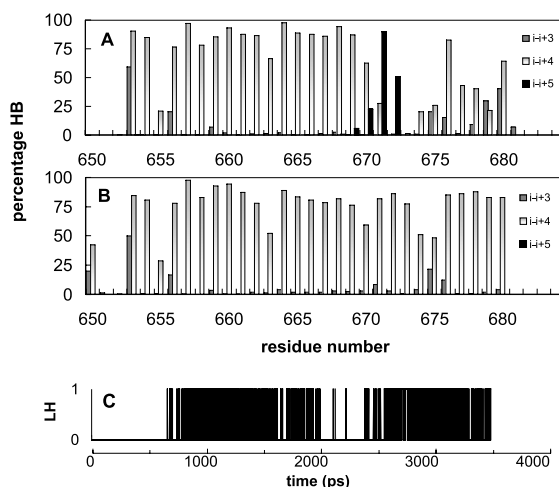
The simulations and analyses were carried out on a multi-processor platform (Linux Beowulf clusters composed of four nodes each having dual 1 GHz Pentium III processors and composed of six nodes each having dual Athlon MP2200+ processors).

## Results and Discussion

### Helical Structure

The root mean square displacement (RMSD) calculated on the C $\alpha$  atoms of the transmembrane sequence from the initial structure is 2.2 Å and 1.3 Å on average for the wild and the mutant forms respectively, attesting that the helical character of the two peptides is conserved over the course of the simulation. However, differences in the hydrogen bond network exist between the two peptides as shown in Figure 1.

**Figure 1:** Percentage of intra helical hydrogen bonds CO<sub>i</sub>-NH<sub>j</sub> (j=i+3, i+4 or i+5) formed in the proto-oncogenic Neu<sup>V</sup><sub>TM</sub> peptide (A) and the oncogenic Neu<sup>E</sup><sub>TM</sub> peptide (B) over the last ns of the simulation. An example of the dynamics of  $\pi$  hydrogen bond is given in (C). The time series of the  $\pi$  hydrogen bond formed between Leu672 and Gly677 shows its appearance at 1.6 ns. The Hbond is very transient and disappears around 3ns and is formed again 0.5 ns later. The geometric criteria for the formation of a hydrogen bond are: O(CO)-N(NH) distance < 2.4 Å and N-H-O angle > 130°. LH = 1 when these criteria are respected; LH = 0 when one of these two criteria is not respected.



The native Neu sequence shows a well defined  $\alpha$  helical portion between Val656 and Phe669. Then local  $\pi$  hydrogen bonds are formed, beginning at Leu670 with a relatively low rate of presence, much higher at the level of the two residues Ile671 and Leu672. The 2  $\pi$  hydrogen bonds involving the carbonyl group of these residues and their corresponding amide groups pertaining to Val676 and Gly677 are relatively strong. However, they are very transient, suggesting rapid fluctuations of the helical backbone at this level of the sequence. As a consequence,  $\pi$  hydrogen bonds may disappear for several hundred ps and reappear



later as illustrated in Figure 1. The formation of the  $\pi$  bulge is followed by the breakage of the hydrogen bonding pattern, leading to the exclusion of the Val673 carbonyl group from the hydrogen bond network. This  $\alpha$  network reappears at the C-terminal side but is strongly weakened.  $^3_10$  hydrogen bonds are detected at a rate reaching 30-40% of the time. The helical behavior of the wild type peptide contrasts with that of the mutant peptide which exhibits a strong  $\alpha$  hydrogen bonding pattern all along the sequence without any sign of  $\pi$  tendency. It should be noted that the very low percentage of  $\alpha$  hydrogen bonds observed at the N-terminus of both peptides is directly related to the presence of the Pro residue. Similar full atom MD simulation results have been very recently reported (51). In a DMPC environment the authors observe the formation of  $\pi$  hydrogen bonds for the native transmembrane domain. The  $\pi$  deformation which involves the carbonyl groups of Leu672 and Val673 was observed relatively late during the course of the simulation but remained formed for the rest of the time, the result being a kink formed near Gly677 which is not observed in our simulations. No  $\pi$  hydrogen bond was detected for the mutant form. Interestingly, these simulation results are comparable to experimental observations. NMR studies of the Neu<sup>V</sup><sub>TM</sub> peptide dissolved in TFE (trifluoroethanol) have revealed that the  $\alpha$  helix exhibits a local distortion identified as  $\pi$  hydrogen bonds between Ile671 and Gly677 (23) (no results were reported for the mutant form). Helix distortion was confirmed by similar NMR experiments using the same solvent where it is shown that the wild peptide deviates from a straight canonical  $\alpha$  helix one helical turn downstream the mutation site, spanning residues Ile671 to Val676 (27). No  $\pi$  hydrogen bond could be confirmed in the mutant form of the peptide from these experiments.

The peptide environment seems to have a control on helix flexibility. Indeed, when the complete transmembrane domain of Neu is dispersed in DPC micelles its  $\alpha$  helical structure is maintained and  $\pi$  hydrogen bonds are not characterized. However, it is shown that the structure formed by the mutant peptide exhibits more flexibility at the N-terminal side upstream the mutation than that adopted by the proto-oncogenic peptide (28). Helical distortions are not rare in membrane proteins (52, 53) and are often implicated in functional roles (54-56).  $\pi$  hydrogen bonds observed experimentally and characterized from MD simulations at the Ile671-Gly677 site in the Neu<sub>TM</sub> are in full agreement with the high propensity of Ile, Leu, and Val residues for  $\pi$  helix formation.

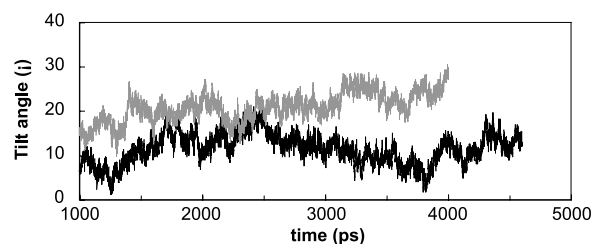
#### *Peptide-environment Interactions*

The influence of the membrane environment and its role in modulating the helicity of the transmembrane domain can be examined in terms of interactions between the helix and the membrane components. The decomposition into two separate contributions, from the lipids and water, reveals their respective importance. Averaged lipid-helix interactions are found to be very similar for the two peptides, on the order of  $-310 \text{ kcal}\cdot\text{mol}^{-1}$ , slightly stronger in the case of the native sequence. The difference in water-helix interactions is more significant. These are stronger for the native sequence than for the mutant by about  $30 \text{ kcal}\cdot\text{mol}^{-1}$  on average. The most favorable interactions of the native transmembrane domain with the membrane environment result from a better insertion in the hydrophobic core of the bilayer mainly due to the  $\pi$  deformation which reduces the helix length (57). However, the variations of other structural parameters induced by the Val substitution, such as peptide tilt may exert other conjugated effects within the bilayer.

#### *Helix Insertion*

Simulation results show that the two peptides adopt a different orientation within the bilayer. It is apparent from the time series presented in Figure 2 that the tilt of the native transmembrane helix does not deviate significantly from the bilayer normal. The tilt angle is around  $10^\circ \pm 2^\circ$  whereas it is around  $24^\circ \pm 2^\circ$  for the mutant

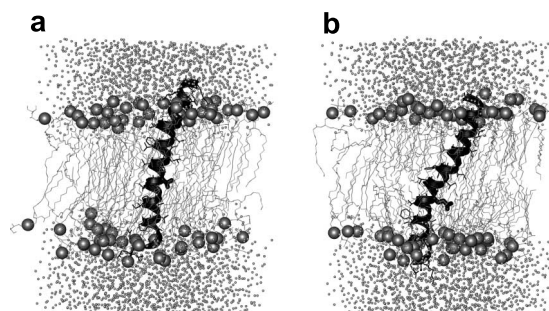
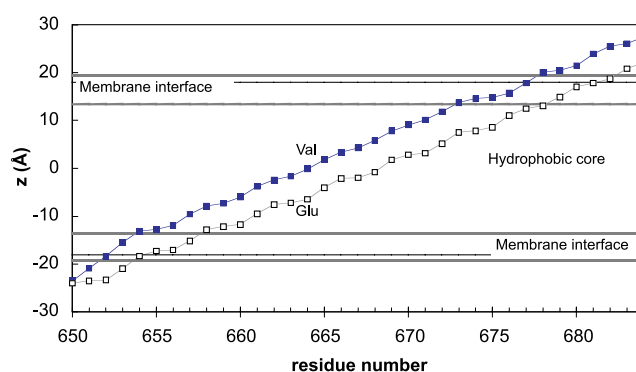
**Figure 2:** Time series of helix tilt relative to the membrane normal ( $z$  axis) from the Neu<sub>TM</sub>/DMPC trajectories throughout the NP<sub>n</sub>AT production period. The black line is for the wild transmembrane sequence; the gray line is for the mutant peptide. RMS fluctuations are around  $3^\circ - 4^\circ$  for the two peptides and the plots suggest a reasonable stability of the helices within the DMPC bilayer.



The tilt angle of the mutant peptide is  $10^\circ$  less than that deduced from polarized FTIR spectra obtained in a DMPC environment (20). The discrepancy between these data might be due to experimental conditions using a gel phase bilayer at low hydration. It should be noted that incorporation of DMPS (dimyristoylphosphatidylserine) in the bilayer changes the peptide orientation, reducing the tilt angle to around  $20-25^\circ$ . A similar helix orientation is observed in a POPC/POPS (palmitoleoylphosphatidylcholine/palmitoleoylphosphatidylserine) membrane (17, 58). Both these results and those recently reported (51) demonstrate that the mutant form of the transmembrane domain is more tilted than the native form. One may note that the simulations performed by Van der Ende *et al.* suggest that helices are not well equilibrated, even after 10 ns of simulation. This could be due to the use of truncated electrostatic interactions and a small number of lipids. Indeed, the use of cut off distances leads to profound artifacts in the properties of the lipid bilayer which may influence the behavior of the inserted peptide (59).

The insertion of the peptides into the DMPC bilayer is represented in Figure 3. It is observed that the hydrophobic domain of the wild peptide matches the hydrophobic thickness of the bilayer better than in the presence of the Glu mutation. The Ser654-Val673 sequence is completely embedded in the bilayer core and the C-terminal region is well anchored in the membrane interface. All the basic residues are within the bulk resulting in the strong interactions discussed above. The  $\pi$  helical distortion contributes to the good matching between the hydrophobic length of the transmembrane peptide and the bilayer thickness. The mutated peptide is shifted towards the N-terminal side leading the four residues Glu650 to Ala653 to be com-

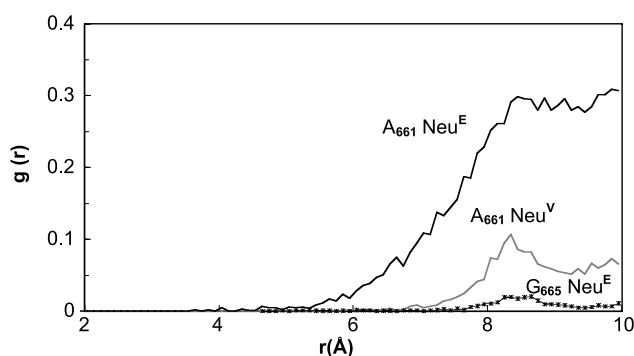
**Figure 3:** Helix insertion into the DMPC bilayer. Average  $z$  position of each residue ( $z$  coordinate of the center of mass of backbone atoms) of the wild transmembrane sequence (Neu<sup>V</sup><sub>TM</sub>) (black square) and of the oncogenic sequence (Neu<sup>E</sup><sub>TM</sub>) (white square). The membrane interface is defined between the planes formed by the choline groups and the glycerol groups (average  $z$  position of their center of mass). Note that these positions are practically identical for the two peptide/DMPC systems. The average  $z$  position of the phosphorus atoms are shown in thin line (close to the choline groups). Panels **a** and **b** show snapshots of the Neu<sup>V</sup><sub>TM</sub>/DMPC and Neu<sup>E</sup><sub>TM</sub>/DMPC systems respectively (helical backbone in a ribbon representation; top: C-terminus; bottom: N-terminus). The Val and Glu residues at position 664 are in thick lines and the phosphorus atoms are represented by large spheres.



pletely immersed in the bulk. The length of the transmembrane sequence embedded in the bilayer core extends from Phe658 to Ile678 and counts 20 residues as in the case of the native sequence. Two C-terminal residues are at the membrane interface, the two others are in the bulk.

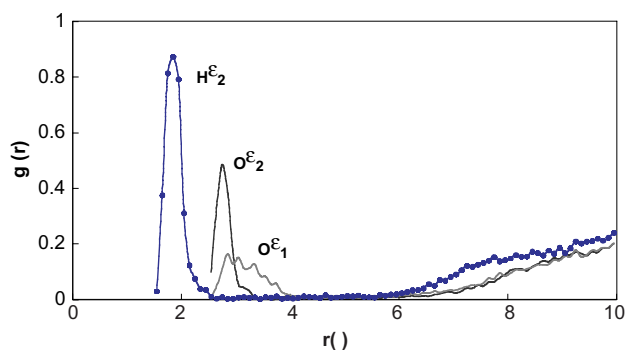
### Water Penetration

NMR experiments show significant differences in the amide exchange behavior between proto-oncogenic and oncogenic forms of the peptide when they are embedded in DPC micelles (28). The exchange rate at Gly665, which is directly adjacent to the mutation site, is found to be faster in the mutant form and it appears that Glu affects the amide signal from Ala661 but not that from Phe669, also close to the mutation. Phe669 exchanges very slowly in both forms of the peptide. The difference in exchange rates at Ala661 and Gly665 is thought to be due to helical destabilization or to a different solvent exposure. The hypothesis is that the N-terminal region of the mutant peptide is close to the surface of the micelle allowing a favorable hydration of the polar Glu side chain.



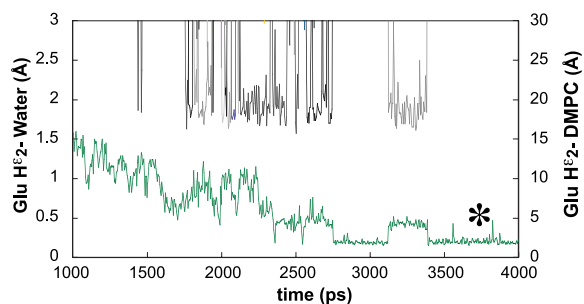
**Figure 4:** Radial distribution functions of water oxygen about the amide proton of Ala661 and Gly665 in the oncogenic form of the peptide. Only the radial distribution function about the amide proton of Ala661 of the proto-oncogenic form of the peptide is shown.  $g(r)$  is zero for Gly665.

Our simulation results confirm that the Glu residue is closer to the membrane interface than the Val residue, suggesting a large hydration of the side chain as well as of the residues upstream the mutation. This has been checked by calculating the radial distribution functions of the water oxygen about the amide proton. The plots in Figure 4 show that the hydration of Ala661 near the mutation is larger than in the wild peptide, which is in good agreement with a much faster amide signal decay than in the proto-oncogenic peptide. Gly665, adjacent to Glu, is poorly hydrated, leading to the conclusion that most of the water molecules that penetrate within the bilayer are captured by the Glu side chain extremity. In the wild peptide water molecules do not reach the Gly residue.



**Figure 5:** Radial distribution functions of water oxygen about the Glu side chain extremity. The primary hydration shell around  $H_{\epsilon 2}$  is constituted of 3.5 water molecules on average. That around  $O_{\epsilon 2}$  is constituted of 1.3 water molecules.

The radial distribution functions of the water oxygen about the Glu side chain extremity are shown in Figure 5. The  $H_{\epsilon 2}$ -O (wat)  $g(r)$  function has a maximum at 1.8 Å and the primary hydration shell (measured as the integral of  $g(r)$  up to 2.3 Å) is constituted of 3.5 water molecules. The  $O_{\epsilon 2}$ -O (wat)  $g(r)$  function shows a sharp peak at 2.7 Å and the primary hydration shell around  $O_{\epsilon 2}$  is constituted of 1.3 water molecules. The hydration of the  $O_{\epsilon 1}$  is more diffuse as suggested by the large peak of the radial distribution function.



**Figure 6:** Hydration of Glu illustrated by the formation of hydrogen bonds between the  $H_{\epsilon 2}$  atom and water (oxygen atom) or a polar group of a lipid during the course of the simulation. In this graph, 26 distinct water molecules exchange hydrogen bonded interactions with the Glu side chain (distance  $H_{\epsilon 2}$ - $O_W < 2.4$  Å and  $N-H-O > 130^\circ$ ). When water molecules are not linked to the carboxylic group, a lipid\* in close proximity can form hydrogen bond (carbonyl oxygen atom of the glycerol group of the Sn1 chain). The decrease in the  $H_{\epsilon 2}$ -O(CO) distance from 15 Å at the beginning of the  $NP_nAT$  period to 2 Å results from rotational and translational motions of one of the lipids closest to the helix.

The residence time of the hydrogen bonds between water and the Glu side chain ranges between a few ps and several hundred ps as shown in Figure 6. The dynamics of water is very rapid as deduced from numerous molecules that exchange hydrogen bond interactions. When water molecules are not hydrogen bonded to the Glu side chain, the closest lipids may form such interactions as shown in the figure.

It is interesting to note that our results lead to a less hydrated Glu side chain than that observed in another simulation study (51). The difference in Glu hydration could be related to the truncation of electrostatic interactions used to simulate the bilayer system which is charged and constituted by a reduced number of lipids. According to Rog *et al.* (60) reasonable simulation results of a bilayer system containing charged molecules can only be obtained when long range electrostatic interactions are evaluated using the PME summation technique.

### **Conclusion**

In this study MD simulations were used to explore the behavior of the proto-oncogenic and oncogenic forms of the transmembrane domain of the Neu receptor in a fluid DMPC bilayer. Analysis of the simulation data reveals the difference in intrinsic flexibility between the two helices and the difference in their insertion within the bilayer. The response to hydrophobic mismatch used by the wild transmembrane peptide is dictated by its intrinsic properties and its propensity for  $\pi$  helix deformation the role of which is to reduce the helix length (24, 52, 57). The reproducibility of helix deformation at the Val motif at the C-terminus (57) demonstrates that the intrinsic properties of this sequence are not affected by the lipid environment. The behavior of the mutated form of the transmembrane domain differs markedly from that of the wild form. The hydrogen bonding pattern of the peptide is largely retained during the trajectory. Then, a larger helix tilt than in the wild peptide was observed for a good matching between the peptide hydrophobic length and the bilayer thickness. In addition, the presence of the Glu mutation induces a vertical shift of the helix towards the membrane interface. Both these effects provide a greater accessibility of the side chain to the membrane environment. The helix face containing Glu is much more exposed to the water molecules and the lipid polar groups than the opposite face. Water molecules were found to form transient hydrogen bonds with the Glu side chain carboxylic extremity. The current simulation therefore suggests an important role for water molecules which may serve as an intermediate state for side chain-side chain or side chain-backbone interactions between the two monomers. Water bridges between Glu were observed in a hydrated DMPC bilayer both in ErbB<sub>2</sub><sub>TM</sub> dimer and Neu<sub>TM</sub> dimer (36, 43). These simulation data agree with experimental results (20, 58).  $\pi$  helix turn formed near the Val motif at the C-terminus of the wild peptide is consistent with NMR observations (23). This higher flexibility of the wild peptide was confirmed from other simulation studies (51). Our computer simulations complement experimental studies (27, 28) by elucidating the influence of the mutation on the properties of the membrane environment.

We are confident that the behavior of the Neu<sub>TM</sub> is reasonably characterized, despite the fact that interpretation of simulation results is limited compared to the complex mechanisms that occur in membranes. Many factors exert counterbalanced effects. The duration of the simulations is, with no doubt, insufficient to fully explore the motions of the transmembrane peptide for membrane insertion. Therefore, short simulations may capture average structural parameters similar to those deduced from longer trajectories (61). It is worth placing these results in the context of the studies of the transmembrane domain of the Neu/ErbB2 receptors. There have been a number of experimental and theoretical studies on the dimers, fewer on the monomers. The more recent experimental studies (27, 28) provide timely confirmation of our simulation results until now unpublished and inspired the current manuscript. These results may serve for the understanding of the mechanisms involved in the formation of the functional structure in the Neu receptor.

## References and Footnotes

1. A. Arora, L. K. Tamm. *Curr. Opin. Struct. Biol.* 11, 540-547 (2001).
2. S. H. White, W. C. Wimley. *Annu. Rev. Biophys. Biomol. Struct.* 28, 319-365 (1999).
3. P. C. Dave, E. K. Tiburu, K. Damodaran, G. A. Lorigan. *Biophys. J.* 86, 1564-1573 (2004).
4. J. L. Popot, D. M. Engelman. *Biochemistry* 29, 4031-4037 (1990).
5. M. A. Olayioye, R. M. Neve, H. A. Lane, N. E. Hynes. *Embo. J.* 19, 3159-3167 (2000).
6. M. D. Marmor, K. B. Skaria, Y. Yarden. *Int. J. Radiat. Oncol. Biol. Phys.* 58, 903-913 (2004).
7. J. L. Popot, D. M. Engelman. *Annu. Rev. Biochem.* 69, 881-922 (2000).
8. F. X. Zhou, M. J. Cocco, W. P. Russ, A. T. Brunger, D. M. Engelman. *Nat. Struct. Biol.* 7, 154-160 (2000).
9. C. Choma, H. Gratkowski, J. D. Lear, W. F. DeGrado. *Nat. Struct. Biol.* 7, 161-166 (2000).
10. A. W. Partridge, A. G. Therien, C. M. Deber. *Proteins* 54, 648-656 (2004).
11. C. I. Bargmann, M. C. Hung, R. A. Weinberg. *Nature* 319, 226-230 (1986).
12. D. B. Weiner, J. Liu, J. A. Cohen, W. V. Williams, M. I. Greene. *Nature* 339, 230-231 (1989).
13. N. E. Hynes, D. F. Stern. *Biochim. Biophys. Acta.* 1198, 165-184 (1994).
14. M. R. de Planque, J. W. Boots, D. T. Rijkers, R. M. Liskamp, D. V. Greathouse, J. A. Killian. *Biochemistry* 41, 8396-8404 (2002).
15. J. A. Killian, G. von Heijne. *T.I.B.S.* 25, 429-434 (2000).
16. W. J. Gullick, A. C. Bottomley, F. J. Lofts, D. G. Doak, D. Mulvey, R. Newman, M. J. Crumpton, M. J. Sternberg, I. D. Campbell. *EMBO J* 11, 43-48 (1992).
17. D. Jones, K. Barber, C. Grant. *Biochem Biophys Acta* 1371, 199-212 (1998).
18. D. H. Jones, E. H. Ball, S. Sharpe, K. R. Barber, C. W. Grant. *Biochemistry* 39, 1870-1878 (2000).
19. S. Sharpe, K. R. Barber, C. W. Grant. *Biochemistry* 39, 6572-6580 (2000).
20. S. O. Smith, C. S. Smith, B. J. Bormann. *Nat Struct Biol* 3, 252-258 (1996).
21. J. P. Duneau, S. Crouzy, N. Garnier, Y. Chapron, M. Genest. *Biophys. Chem.* 76, 35-53 (1999).
22. N. Garnier, D. Genest, E. Hebert, M. Genest. *J. Biomol. Struct. Dyn.* 11, 983-1002 (1994).
23. M. Goetz, C. Carlotti, F. Bontems, E. Dufourc. *Biochemistry* 40, 6534-6540 (2001).
24. J. P. Duneau, N. Garnier, M. Genest. *J. Biomol. Struct. Dyn.* 15, 555-572 (1997).
25. N. Sajot, M. Genest. *Eur. Biophys. J.* 28, 648-662 (2000).
26. N. Sajot, M. Genest. *J. Biomol. Struct. Dyn.* 19, 15-31 (2001).
27. R. S. Houliston, R. S. Hodges, F. J. Sharom, J. H. Davis. *FEBS Lett.* 535, 39-43 (2003).
28. R. S. Houliston, R. S. Hodges, F. J. Sharom, J. H. Davis. *J. Biol. Chem.* 279, 24073-24080 (2004).
29. B. R. Brooks, R. E. Brucoleri, B. D. Olafson, D. J. States, S. Swaminathan, M. Karplus. *J. Comp. Chem.* 4, 187-217 (1983).
30. J. A. D. MacKerell, D. Bashford, M. Bellot, R. L. Dunbrack, J. D. Evansek, M. J. Field, S. Fisher, J. Gao, H. Guo, S. Ha, D. J. McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, B. Schlenkrich, J. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, M. Karplus. *Biophys J* 61, A143 (1992).
31. J. A. D. MacKerell, D. Bashford, M. Bellot, R. L. Dunbrack, J. D. Evansek, M. J. Field, S. Fisher, J. Gao, H. Guo, S. Ha, D. J. McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, B. Schlenkrich, J. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, M. Karplus. *J. Phys. Chem. B.* 102, 3585-3616 (1998).
32. M. Schlenkrich, J. Brickmann, A. D. J. MacKerell, M. Karplus. In *Biological Membranes, A Molecular Perspective From Computation and Experiment*, pp. 31-81. Eds., K. J. Merz, B. Roux. Birkhäuser, Boston (1996).
33. W. L. Jorgensen, R. W. Impey, J. Chandrasekhar, J. D. Madura, M. L. Klein. *J Chem Phys* 79, 926-935 (1983).
34. J. P. Ryckaert, G. Ciccotti, B. H. J. C. J. Comput. Phys. 23, 327 (1977).
35. R. B. Gennis. *Biomembranes Molecular Structure and Function*. Springer-Verlag, New York (1989).
36. N. Garnier, S. Crouzy, M. Genest. *J. Biomol. Struct. Dyn.* 21, 179-200 (2003).
37. K. Belohorcova, J. Qian, J. H. Davis. *Biophys. J.* 79, 3201-3216 (2000).
38. P. Lague, M. J. Zuckermann, B. Roux. *Biophys. J.* 81, 276-284 (2001).
39. H. I. Petrache, S. W. Dodd, M. F. Brown. *Biophys. J.* 79, 3172-3192 (2000).
40. J. F. Nagle. *Biophys. J.* 64, 1476-1481 (1993).
41. J. F. Nagle, S. Tristram-Nagle. *Biochem. Biophys. Acta.* 1469, 159-195 (2000).
42. N. Sal-Man, D. Gerber, Y. Shai. *Biochemistry* 43, 2309-2313 (2004).
43. P. Aller, L. Voiry, N. Garnier, M. Genest. *Biopolymers* 77, 184-197 (2005).
44. T. B. Woolf, B. Roux. *Protein Struct. Funct. Genet.* 24, 92-114 (1996).
45. T. B. Woolf, B. Roux. *Proc. Natl. Acad. Sci. USA* 91, 11631-11635 (1994).
46. B. Roux, T. Woolf. In *Biological Membranes, A Molecular Perspective From Computation and Experiment*, pp. 555-587. Eds., K. J. Merz, B. Roux. Birkhäuser, Boston (1996).
47. B. A. Lewis, D. M. Engelman. *J. Mol. Biol.* 166, 211-217 (1983).
48. R. M. Venable, Y. Zhang, B. J. Hardy, R. W. Pastor. *Science* 262, 223-226 (1993).
49. W. G. Hoover. *Phys Rev A* 31, 1695-1697 (1985).
50. T. A. Darden, D. York, L. Pederson. *J Chem Phys* 98, 10089-10092 (1993).

51. B. M. Van Der Ende, F. J. Sharom, J. H. Davis. *Eur. Biophys. J.* 33, 596-610 (2004).
52. J. P. Cartailier, H. Luecke. *Structure (Camb)* 12, 133-144 (2004).
53. M. N. Fodje, S. Al-Karadaghi. *Protein Eng.* 15, 353-358 (2002).
54. H. Luecke, B. Schobert, H. T. Richter, J. P. Cartailier, J. K. Lanyi. *J. Mol. Biol.* 291, 899-911 (1999).
55. H. Luecke, B. Schobert, H. T. Richter, J. P. Cartailier, J. K. Lanyi. *Science* 286, 255-261 (1999).
56. H. Luecke, B. Schobert, J. K. Lanyi, E. N. Spudich, J. L. Spudich. *Science* 293, 1499-1503 (2001).
57. J. P. Duneau, D. Genest, M. Genest. *J. Biomol. Struct. Dyn.* 13, 753-769 (1996).
58. S. O. Smith, C. Smith, S. Shekar, O. Peersen, M. Ziliox, S. Aimoto. *Biochemistry* 41, 9321-9332 (2002).
59. M. Patra, M. Karttunen, M. T. Hyvonen, E. Falck, P. Lindqvist, I. Vattulainen. *Biophys J* 84, 3636-3645 (2003).
60. T. Rog, K. Murzyn, M. Pasenkiewicz-Gierula. *Acta. Biochim. Polon.* 50, 789-798 (2003).
61. H. I. Petrache, D. M. Zuckerman, J. N. Sachs, J. A. Killain, R. E. Koeppe, T. B. Woolf. *Langmuir* 18, 1340-1351 (2002).

*Date Received: Needed*

**Communicated by the Editor Ramaswamy H Sarma**

## Conclusion

Les résultats obtenus au cours de ces études de dynamique moléculaire montrent d'une part que le peptide sauvage Neu<sup>V</sup><sub>TM</sub> adopte une structure hélicoïdale avec une hernie  $\pi$  et le peptide muté Neu<sup>E</sup><sub>TM</sub> est entièrement sous forme d'une hélice  $\alpha$ . Ces résultats sont en accord avec les données expérimentales (Goetz et al. 2001).

D'autre part, la différence d'insertion dans la membrane du côté N-terminal de Neu<sup>V</sup><sub>TM</sub>  $\pi$  et muté Neu<sup>E</sup><sub>TM</sub>  $\alpha$  mise en évidence par dynamique moléculaire est aussi en accord avec des études récentes par RMN d'échange de proton amide (Goetz et al. 2001; Houliston et al. 2004).

Le peptide sauvage Neu<sup>V</sup><sub>TM</sub>  $\pi$  est le plus stable dans la membrane lipidique. Cette déformation pourrait donc avoir d'importantes conséquences dans les interactions des hélices transmembranaires.

L'étude des quatre monomères des récepteurs de la famille ErbB dans un environnement membranaire permettrait de comprendre le mode d'insertion du domaine TM. Cette étude dans des modèles de membrane plus proches d'une membrane biologique c'est-à-dire des modèles de membranes composés de lipides présentant des insaturations et/ou de membranes mixtes (POPC, DMPC/cholestérol, etc.) permettra d'avoir des informations encore plus proche de la réalité biologique.