

**COMMUNAUTE FRANCAISE DE BELGIQUE  
ACADEMIE UNIVERSITAIRE WALLONIE-EUROPE  
FACULTE UNIVERSITAIRE DES SCIENCES AGRONOMIQUES DE  
GEMBLoux**

**Nouveaux outils *in silico* de modélisation de peptides :  
Design de peptides complémentaires et prédiction  
de structures peptidiques.**

Marc DECAFFMEYER

Essai présenté en vue de l'obtention du grade de docteur  
en sciences agronomiques et ingénierie biologique

Promoteurs : Prof. Robert Brasseur  
Dr Annick Thomas

- 2008 -

**Decaffmeyer Marc (2008).** Nouveaux outils *in silico* de modélisation de peptides : Design de peptides complémentaires et prédiction de structures peptidiques (thèse de doctorat). **Facultés des Sciences Agronomiques de Gembloux, 146 p.,1 tabl., 14 fig.**

**Résumé :** Les peptides, malgré qu'ils soient des molécules d'avenir pour le développement des biotechnologies sont à ce jour le parent pauvre au niveau des méthodes de modélisation disponibles. Pourtant, de plus en plus de sociétés pharmaceutiques tentent de contrecarrer la pénurie relative en nouvelles molécules médicamenteuses en utilisant des peptides comme agents thérapeutiques, voire comme outils de conception de nouvelles petites molécules. Il s'agit donc d'un domaine qui mérite l'attention des chercheurs et en particulier des « modélisateurs ».

Le premier aspect abordé dans ce travail se rapporte à la modélisation *in silico* de peptides complémentaires à une cible. A cet effet, la méthode PepDesign, qui s'appuie sur une série d'outils préexistants au CBMN, a été développée. Elle permet de contourner les désavantages des quelques stratégies existantes à ce jour, telles que le screening de banques de peptides qui est onéreux et nécessite un grand nombre d'essais biologiques ou la dynamique moléculaire qui est très consommatrice en temps et donc peu compatible avec les exigences économiques de l'industrie pharmaceutique. La méthode PepDesign a été appliquée avec succès à la conception d'un peptide anti A $\beta$ 29-42.

Le second aspect rapporté dans cette thèse est relatif à la prédiction *in silico* de la structure des peptides. On le sait, la fonction des peptides, comme celle des protéines, est intimement liée à leur structure. En dépit de cela seulement deux méthodes non expérimentales de détermination structurale des peptides étaient accessibles récemment. Notre laboratoire a développé une méthode innovante basée sur une procédure de type Boltzmann-stochastique qui a été utilisée pour deux applications. Dans la première nous avons vérifié l'hypothèse selon laquelle le polymorphisme structural des peptides vecteurs (cell penetrating peptides) est un critère nécessaire pour leur capacité d'internalisation. Dans la seconde application, nous avons proposé un modèle de topologie pour le segment hydrophobe de DGK $\epsilon$ , une protéine de structure inconnue à ce jour.

**Decaffmeyer Marc (2008).** New *in silico* tools for peptide modelling : Design of complementary peptides and peptidic structural prediction (PhD thesis in french). **Gembloux Agricultural Faculty, 146 p.,1 tabl., 14 fig.**

**Abstract:** Despite the promising future of peptides in the development of biotechnologies, today available modelling methods for the peptide design are not numerous. Nevertheless, more and more pharmaceutical companies try to overcome the relative scarceness of new medicinal molecules by using peptides as therapeutic agents,

or even as tool to conceive new small molecules. Thus peptide modelling deserves our attention.

The first point in this work relates to *in silico* modelling of a target complementary peptides. The PepDesign method has been developed based on a series of pre-existing tools at the CBMN. It enables to short-circuit the disadvantages of rare existing experimental and *in silico* strategies, such as the screening of peptides libraries which is expensive and requires a big number of biological tries or molecular dynamics which are time consuming and therefore poorly compatible with economic requirements of pharmaceutical industries. PepDesign has been applied successfully for the design of an anti A $\beta$ 29-42 peptide.

The second aspect reported in this thesis is related to the *in silico* prediction of peptide structures. As we know, the function of peptides, such as the function of proteins, is closely linked to their structures. Only two *in silico* methods for structure determination of peptides were available. Our laboratory has developed an innovative algorithm based on a Boltzmann-stochastic procedure which was used for two applications. In the first one we tested the structural polymorphism of cell penetrating peptides (CPP) and demonstrated that this polymorphism is mandatory for CPP internalisation capacity. In a second application, we proposed a 3D model of topology for the hydrophobic segment of DGK $\epsilon$ , a protein with an unknown structure.

**Copyright :** *Aux termes de la loi belge du 22 mars 1886 sur le droit d'auteur, seul l'auteur a le droit de reproduire cet ouvrage ou d'en autoriser la reproduction de quelque manière et sous quelque forme que ce soit. Toute photocopie ou reproduction sous une autre forme est donc faite en violation avec la loi.*

## **Remerciements.**

*Je tiens tout particulièrement à remercier Robert Brasseur dont l'esprit d'innovation et l'imagination sont sans limite ainsi que pour l'enthousiasme qu'il communique à toute son équipe depuis de nombreuses années et pour son amitié. Un grand merci aussi à Annick Thomas pour sa disponibilité, le partage sans restriction de ses compétences et de son expérience. Merci aussi à Laurence Lins et à Benoît Charlotheaux pour leur expertise en bioinformatique et en biochimie qu'ils sont toujours prêts à partager.*

*Mes remerciements vont aussi Marie-Hélène Van Eyck, Stéphanie Ravault, Sebastien Deshayes, Sébastien Santini, Jean-Marc Crowet, Nicolas Delsaux pour les discussions scientifiques et parfois moins sérieuses que nous avons eues ainsi qu'à Jean-marc Delseth qui a assuré jusqu'il y a peu l'intendance du labo avant une retraite bien méritée et à Colette Zeches qui le remplace avec efficacité ainsi qu'à Vivianne Godecharles. Merci aussi à Christelle pour le travail remarquable qu'elle réalise devant ses paillasses. Merci à Aurélien Lorin et Benoît Adam qui ont partagé le même bureau que moi pendant de longues années et à Gaëtan pour son assistance informatique.*

*Je remercie également les deux rapporteurs de cette thèse, le Professeur Paquot et le Dr Gilles Divita, ainsi que les membres du jury pour le temps consacré ainsi que les différents organismes qui ont financés les travaux relatés dans cette thèse : Le PAI, la DG TRE au travers des conventions  $\beta$ A4 et PepSein.*

*Thanks to Professor Richard Epanand from the Mc Master university in Hamilton (Canada) for his availability, his kindness and also the numerous discussions we had relative to the DGKe.*

*Et, last but not least, un grand merci aux cops de l'Ermitage : André, Fred, Joelle, et tous les autres ainsi qu'à Marianne, Allison et Tom qui m'ont soutenu durant ces dernières années et qu'à mes parents, Yvette et André pour leur soutien dans tous les domaines depuis plus de 40 ans.*



## Liste des tables et figures

<b>Figure 1.</b> Forme générale des acides $\alpha$ -aminés (A). Isomères optiques des acides aminés dus à l'asymétrie du C $\alpha$ (B). Les 20 acides aminés (C)	18
<b>Figure 2.</b> Définition des angles $\omega, \Phi$ et $\Psi$ .	20
<b>Figure 3.</b> Carte générale théorique de Ramachandran.	21
<b>Figure 4.</b> Feuillet $\beta$ parallèle (A), antiparallèle (B) et hélice $\alpha$ (C).	23
<b>Figure 5.</b> Angles $\chi_1$ et $\chi_2$ pour l'acide glutamique .	23
<b>Figure 6.</b> Insertion membranaires des peptides amphipatiques (A), transmembranaires (B) et obliques (C).	24
<b>Figure 7.</b> Dépenses en recherche pharmaceutique aux USA ( $10^9$ \$) par an. Nouvelles molécules déposées à la FDA par an (B).	28
<b>Figure 8.</b> Energie en fonction de la longueur de liaison.	33
<b>Figure 9.</b> Périodicité de l'énergie pour les angles dièdres.	34
<b>Figure 10.</b> Energie d'Interaction de van der Waals	35
<b>Figure 11.</b> Interactions électrostatiques.	35
<b>Figure 12.</b> Evolution linéaire (gauche) et sigmoïdale (droite) de la constante diélectrique ( $\epsilon$ ) en fonction de la distance interatomique.	36
<b>Table 1.</b> Energie de transfert des différents types atomiques.	37
<b>Figure 13.</b> Description schématique de la procédure multi-étapes PepDesign	47
<b>Figure 14.</b> Schéma de principe de la méthode PepLook.	67
<b>Figure 15.</b> Angles $\Phi\Psi$ utilisés pour la prédiction de structure.	81
<b>Figure 16.</b> Distribution de l'énergie des structures calculées.	83
<b>Figure 17.</b> Transformation des fréquences en énergie. Valeurs de référence des MFP des 20 résidus.	85
<b>Figure 18.</b> Score de polymorphisme le long de la séquence.	86
<b>Figure 19.</b> Capacité d'internalisation des peptides étudiés.	87
<b>Figure 20.</b> Structure des primes pour la Pénétratine et ses mutants en milieu hydrophobe et hydrophile.	88
<b>Figure 21.</b> Alignement structural des structures de la Pénétratine.	89
<b>Figure 22.</b> Distribution des structures calculées pour la Pénétratine.	90
<b>Figure 23.</b> RMSd[9] des structures représentatives des différentes populations pour la Pénétratine.	90
<b>Figure 24.</b> Structure des représentants des familles structurales de la Pénétratine.	91
<b>Figure 25.</b> Distribution des structures calculées pour la Pénétratine 48.	91
<b>Figure 26.</b> RMSd[9] des structures représentatives des différentes populations pour la Pénétratine 48.	92
<b>Figure 27.</b> Structure des représentants des familles structurales de la Pénétratine 48.	92
<b>Figure 28.</b> Distribution des structures calculées pour la Pénétratine 48-56.	93
<b>Figure 29.</b> RMSd[9] des structures représentatives des différentes populations pour la Pénétratine 48-56.	93
<b>Figure 30.</b> Structure des représentants des familles structurales de la Pénétratine 48.	94
<b>Figure 31.</b> Structure des primes pour le Transportan et ses variants.	95
<b>Figure 32.</b> Distribution des structures calculées pour le Transportan.	96

<b>Figure 33.</b> RMSd[9] des structures représentatives des différentes populations pour le Transportan.	<b>97</b>
<b>Figure 34.</b> Structure des représentants des familles structurales du Transportan.	<b>97</b>
<b>Figure 35.</b> Distribution des structures calculées pour le TP07.	<b>98</b>
<b>Figure 36.</b> RMSd[9] des structures représentatives des différentes populations pour le TP07.	<b>98</b>
<b>Figure 37.</b> Structure des représentants des familles structurales de TP07.	<b>99</b>
<b>Figure 38.</b> Distribution des structures calculées pour le TP10.	<b>99</b>
<b>Figure 39.</b> RMSd[9] des structures représentatives des différentes populations pour le TP10.	<b>100</b>
<b>Figure 40.</b> Structure des représentants des familles structurales de TP10.	<b>100</b>
<b>Figure 41.</b> Distribution des structures calculées. RMSd[9] des structures représentatives des différentes populations. Structure des représentants des familles structurales.	<b>101</b>
<b>Figure 42.</b> Tableau récapitulatif des peptides étudiés, de leurs scores de polymorphisme et de leur capacité d'internalisation.	<b>102</b>
<b>Figure 43.</b> Implication de DGKε dans la conversion des DAGs en acides phosphatidiques.	<b>103</b>
<b>Figure 44.</b> Représentation du graphe $\langle \mu_i \rangle$ en fonction de $\langle H_i \rangle$ dans la méthode d'Eisenberg.	<b>104</b>
<b>Figure 45.</b> Représentation de la membrane (valeur de $C(z)$ en bleu).	<b>106</b>
<b>Figure 46.</b> Graphe d'Eisenberg pour DGKε.	<b>108</b>
<b>Figure 47.</b> Energie d'interaction des segments allongés de DGKε en fonction de Z.	<b>108</b>
<b>Figure 48.</b> Insertion du segment D18-Q42 dans le modèle Impala.	<b>109</b>
<b>Figure 49.</b> Les 99 structures de basse énergie calculées par PepLook. Structures représentatives des deux populations structurales.	<b>110</b>
<b>Figure 50.</b> Distribution des structures calculées en fonction de leur gRMSd. RMSd [9] le long de la séquence.	<b>110</b>
<b>Figure 51.</b> Les 99 modèles PepLook classés selon leur taux d'hélicité. gRMSd (en Å) par rapport au Prime. Delta d'énergie par rapport au Prime des 8 peptides les plus stables.	<b>111</b>
<b>Figure 52.</b> Structure secondaire du Prime et du modèle 89.	<b>112</b>
<b>Figure 53.</b> Score de stabilité du Prime et du modèle 89 .	<b>112</b>
<b>Figure 54.</b> Gain de stabilité pour le dimère sous forme Prime.	<b>114</b>
<b>Figure 55.</b> Résultats de microscopie de fluorescence après perméabilisation et sans perméabilisation.	<b>116</b>
<b>Figure 56.</b> Les 4 hypothèses concernant la topologie de DGKε.	<b>117</b>
<b>Figure 57.</b> Vue des 99 structures calculées par Peplook . Distribution des structures par gRMSd, RMSd[9] pour les deux populations. Résultats de microscopie de fluorescence sur cellules non perméabilisées.	<b>118</b>

## Liste des abréviations.

Å : Angström ( $10^{-10}$ m).

A $\beta$ 29-42 : Peptide amyloïde de fusion.

AchR : Acetylcholine receptor.

ADMET : Absorption, distribution, métabolisme, excrétion, toxicité.

ADN : Acide désoxyribonucléique

ADNP : Activity dependent neuroprotective protein.

AMBER : Assisted model building with energy refinement.

AMP : Antimicrobial peptide.

Apo: Apolipoprotéine.

APP : Amyloid precursor protein.

ARN : Acide ribonucléique

$\alpha$ -BTX : Alpha Bungarotoxin.

CASP : Critical assessment of techniques for protein structure prediction.

CBMN: Centre de biophysique moléculaire numérique.

CD : Dichroïsme circulaire.

DAG: Diacylglycérol.

DGK $\epsilon$  : Diacylglycerol kinase  $\epsilon$ .

DM : Dynamique moléculaire.

DPPC : Dipalmitoyl phosphatidyl choline.

CPP : Cell penetrating peptide.

FT-IR : Spectroscopie infrarouge à transformée de Fourier.

HA2 : Peptide de l'Hémagglutinine humaine.

hCT : Calcitonine humaine.

HIV : Human immunodeficiency virus.

kD : kiloDalton.

MC : Monte carlo.

MFP : Mean forces potential.

NDV : Newcastle Disease Virus.

NRPS : Nonribosomal peptide synthetase.

PDB : Protein Data Bank.

RMN : Résonance magnétique nucléaire.

RMSd : Root mean square deviation.

RX : Rayons X

SIV : Simian immunodeficiency virus.

SRR : Systematic residue replacement.

# Table des matières.

<b>1. INTRODUCTION.</b>	<b>10</b>
<b>1.1. Les acides aminés.</b>	<b>11</b>
<b>1.2. Les peptides.</b>	<b>14</b>
1.2.1. Structure des peptides.	14
1.2.2. Interaction des peptides avec les membranes.	18
1.2.3. Les peptides naturels.	19
1.2.4. Production industrielle de peptides.	20
1.2.5. Utilisation thérapeutique des peptides.	21
a) Contexte économique.	21
b) Avantages des peptides.	22
c) Inconvénients des peptides	23
d) Conception de peptides thérapeutiques.	25
<b>1.3. Les champs de forces.</b>	<b>26</b>
1.3.1. Atomes liés.	26
1.3.2. Atomes non liés.	28
<b>2. OBJECTIF.</b>	<b>32</b>
<b>3. RESULTATS.</b>	<b>35</b>
<b>3.1. Aspect design.</b>	<b>36</b>
3.1.1. Modélisation rationnelle de peptides complémentaires au peptide de fusion A $\beta$ 29-42: Une application de PepDesign.	40
3.1.2. Peptides complémentaires au peptide $\beta$ -Amyloïde 29-42 (Brevet conjoint FSAGx-Ulg).	41
<b>3.2. Aspect prédiction de la structure.</b>	<b>48</b>
3.2.1. Prédiction de la structure peptidique : Où en sommes nous ?	48
3.2.2. Le polymorphisme structural, un paramètre nécessaire pour la pénétration cellulaire des peptides : hypothèse.	53

a) Introduction.	53
b) Méthode.	54
c) Résultats et discussion.	60
3.2.3 Détermination de la topologie du segment hydrophobe de la diacylglycerol kinase $\epsilon$ de mammifère dans une membrane cellulaire et sa relation avec les prédictions en modélisation.	76
a) Introduction.	76
b) Méthode.	77
c) Résultats et discussion.	80
<b>4. CONCLUSION GENERALE ET PERSPECTIVES.</b>	<b>93</b>
<b>ARTICLES ET BREVETS PUBLIES DURANT CETTE THESE.</b>	<b>97</b>
<b>ANNEXE (Article de synthèse publié dans la revue BASE). <i>Les médicaments peptidiques : Mythe ou réalité</i></b>	<b>98</b>
<b>Références bibliographiques.</b>	<b>107</b>

# **1. INTRODUCTION**

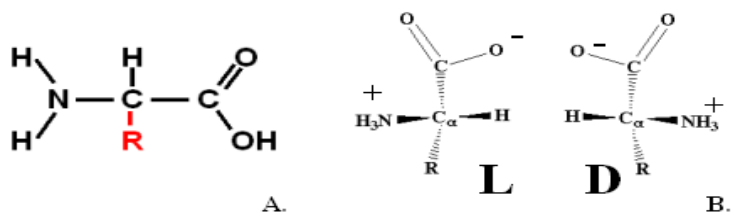
## 1.1. Les acides aminés.

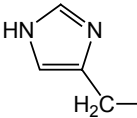
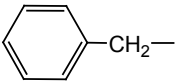
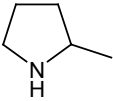
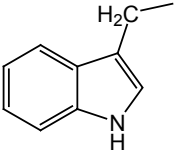
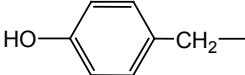
Les acides aminés, sont les éléments constitutifs des peptides et des protéines, les anglo-saxons parlent de « building blocks ». Ils sont liés entre eux par des liaisons peptidiques (ou liens amides) pour former ces biopolymères que sont les peptides et les protéines (**Creighton, 1993; Branden et Tooze, 1999**). Les acides aminés dont on parle le plus couramment sont des acides  $\alpha$ -aminés, indiquant que la fonction amine et la fonction carboxylique sont toutes deux directement voisines du carbone  $\alpha$  porteur de la chaîne latérale. Leur structure générale est donnée à la figure 1A. Ce carbone  $\alpha$ , excepté pour la glycine est asymétrique, conférant aux acides aminés une activité optique. Deux stéréoisomères (L ou D) sont donc possibles (figure 1B).

Il existe une centaine d'acides  $\alpha$ -aminés, mais vingt d'entre eux sont quasi exclusivement impliqués dans la construction des peptides et protéines naturels (figure 1C), Deux autres acides aminés naturels ont toutefois été identifiés récemment : la sélénocystéine (**Gladyshev et Hatfield, 1999**) et la pyrrolysine (**Hao et al., 2002**). Ces vingt acides aminés naturels sont tous lévogyres. Notons toutefois que des acides aminés dextrogyres ont été identifiés chez certains animaux (crustacés, amphibiens,...) (**Kreil, 1997**) et dans les parois cellulaires bactériennes (**Van Heijenoort J, 2001**).

Tous les acides  $\alpha$ -aminés ont un squelette commun et ne sont différenciés que par la chaîne latérale qu'ils portent. Ce sont ces chaînes latérales qui confèrent leurs propriétés spécifiques. Ils sont le plus souvent classés en 4 catégories qui se rapportent à leur hydrophobicité (ou lipophilicité) et leur capacité à s'ioniser : les hydrophobes (ou apolaires) telles l'alanine, la valine, le tryptophane,..., les hydrophiles (ou polaires) telles l'asparagine la sérine,..., les acides (aspartate et glutamate) et les bases (lysine, arginine et histidine). On peut également les classer selon leur pK, c'est à dire leur capacité à s'ioniser en captant ou en perdant un proton, ou encore en fonction de leur importance relative dans les banques de séquences telles que Swiss-Prot ou du volume qu'ils occupent (figure 1c).

La fonction amine des acides aminés peut capter un proton pour se trouver sous forme cationique à pH acide. A pH basique, c'est la fonction carboxylique qui est déprotonée pour adopter la forme anionique. Au pH isoélectrique, les deux fonctions sont sous forme ionisée et l'acide aminé est zwitterionique. Dans les peptides et les protéines, seules la première fonction amine (en N terminal) et la dernière fonction acide (en C terminal) sont ionisables.



1L	3L	Nom	Chaîne latérale R	MM	pI	pK	%	Vol
<b>A</b>	Ala	Alanine	CH <sub>3</sub> - H <sub>2</sub> N-C(NH)-NH-	89.09	6.01		8.3	67
<b>R</b>	Arg	Arginine	CH <sub>2</sub> CH <sub>2</sub> CH <sub>2</sub> -	174.20	10.76	12.48	5.7	148
<b>N</b>	Asn	Asparagine	H <sub>2</sub> NCO-CH <sub>2</sub> -	132.12	5.41		4.4	96
<b>D</b>	Asp	Ac. aspartique	HO <sub>2</sub> C-CH <sub>2</sub> -	133.10	2.85	3.90	5.3	91
<b>C</b>	Cys	Cysteine	HS-CH <sub>2</sub> -	121.16	5.05	8.18	1.7	86
<b>E</b>	Glu	Ac. glutamique	HO <sub>2</sub> C-CH <sub>2</sub> CH <sub>2</sub> -	147.13	3.15	4.07	4	114
<b>Q</b>	Gln	Glutamine	H <sub>2</sub> NCO-CH <sub>2</sub> CH <sub>2</sub> -	146.15	5.65		6.2	109
<b>G</b>	Gly	Glycine	H-	75.07	6.06		7.2	48
<b>H</b>	His	Histidine		155.16	7.60	6.04	2.2	118
<b>I</b>	Ile	Isoleucine	CH <sub>3</sub> CH <sub>2</sub> (CH <sub>3</sub> )CH-	131.17	6.05		5.2	124
<b>L</b>	Leu	Leucine	(CH <sub>3</sub> ) <sub>2</sub> CH-CH <sub>2</sub> -	131.17	6.01		9	124
<b>K</b>	Lys	Lysine	H <sub>2</sub> NCH <sub>2</sub> CH <sub>2</sub> CH <sub>2</sub> CH <sub>2</sub> -	146.19	9.60	10.54	5.7	135
<b>M</b>	Met	Methionine	H <sub>3</sub> C-S-CH <sub>2</sub> CH <sub>2</sub> -	149.21	5.74		2.4	124
<b>F</b>	Phe	Phenylalanine		165.19	5.49		3.9	135
<b>P</b>	Pro	Proline		115.13	6.30		5.1	90
<b>S</b>	Ser	Serine	HO-CH <sub>2</sub> -	105.09	5.68		6.9	73
<b>T</b>	Thr	Threonine	CH <sub>3</sub> CH(OH)-	119.12	5.60		5.8	93
<b>W</b>	Trp	Tryptophane		204.23	5.89		1.3	163
<b>Y</b>	Tyr	Tyrosine		181.19	5.64	10.46	3.2	141
<b>V</b>	Val	Valine	(CH <sub>3</sub> ) <sub>2</sub> CH-	117.15	6.00		6.6	105

C.

Figure 1. Forme générale des acides  $\alpha$ -aminés (A). Isomères optiques des acides aminés dus à l'asymétrie du C $\alpha$  (B).. Les 20 acides aminés (C). Avec le code 1 lettre, le code 3 lettres, le nom, la formule, la masse molaire (Da), le pH iso, le pK de la chaîne latérale, l'abondance dans la banque Swiss-Prot (%) et le volume de Van der Waals (Å<sup>3</sup>). Les hydrophobes en orange, les hydrophiles en vert, les basiques en rouge et les acides en bleu.



Tous ces acides aminés ont des propriétés chimiques particulières en raison de la nature de leur chaîne latérale. La cystéine, par exemple, est capable de former, par oxydation, des ponts disulfures qui sont largement impliqués dans la stabilisation de la structure tridimensionnelle de certaines protéines. Les acides aminés acides et basiques quant à eux peuvent se protoner/déprotoner en fonction du pH du milieu. Par exemple, l'acide glutamique porte une charge négative au dessus de son pK ( $\text{pH} > 4$ ), tandis que la lysine porte une charge positive en dessous de son pK ( $\text{pH} < 10.5$ ). Quoiqu'il en soit, tous les résidus acides et basiques, se présentent sous forme ionique dans la plupart des conditions physiologiques ( $\text{pH} \approx 7.4$ ).

## 1.2. Les peptides.

### 1.2.1. Structure des peptides.

Les peptides, comme les protéines sont constitués d'acides aminés, liés entre eux par des liaisons peptidiques. On considère généralement, même si cette distinction est subjective que l'on a à faire à un peptide en dessous de 50 résidus ou 7kD, au delà de cette limite on parlera de protéine.

En raison d'une délocalisation partielle du doublet d'électrons de l'azote, le lien peptidique est une liaison double partielle. En effet, la distance atomique mesurée dans les protéines foldées entre le C du carbonyle et le N de l'amine est de 1,33 Å, distance qui est intermédiaire entre celle d'une double liaison NC (1,25Å) et celle d'une simple liaison (1,45Å). Cette délocalisation implique que le lien peptidique est plan puisque la rotation autour de l'axe NH-CO n'est plus possible en raison du caractère double de la liaison. Dès lors, l'angle de rotation autour de la liaison peptidique, angle  $\omega$  ne peut adopter que les valeurs de 0 ou de 180° (forme cis ou trans). La structure de la chaîne principale (backbone) des peptides n'est donc conditionnée que par l'angle que forment les deux plans de deux liens peptidiques successifs et donc par la rotation autour de l'axe N-C $\alpha$  (angle  $\Phi$ ) et de l'axe C $\alpha$ -C (angle  $\Psi$ ) (figure2).

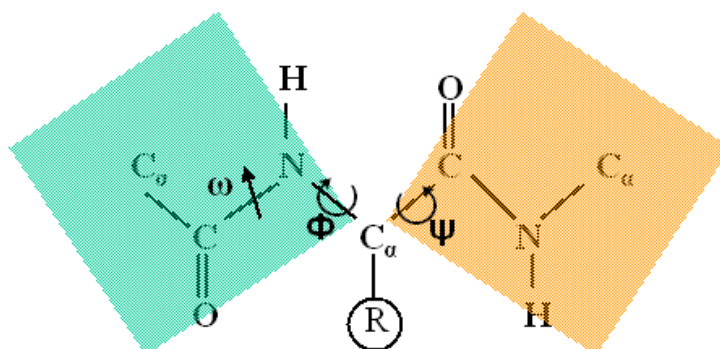
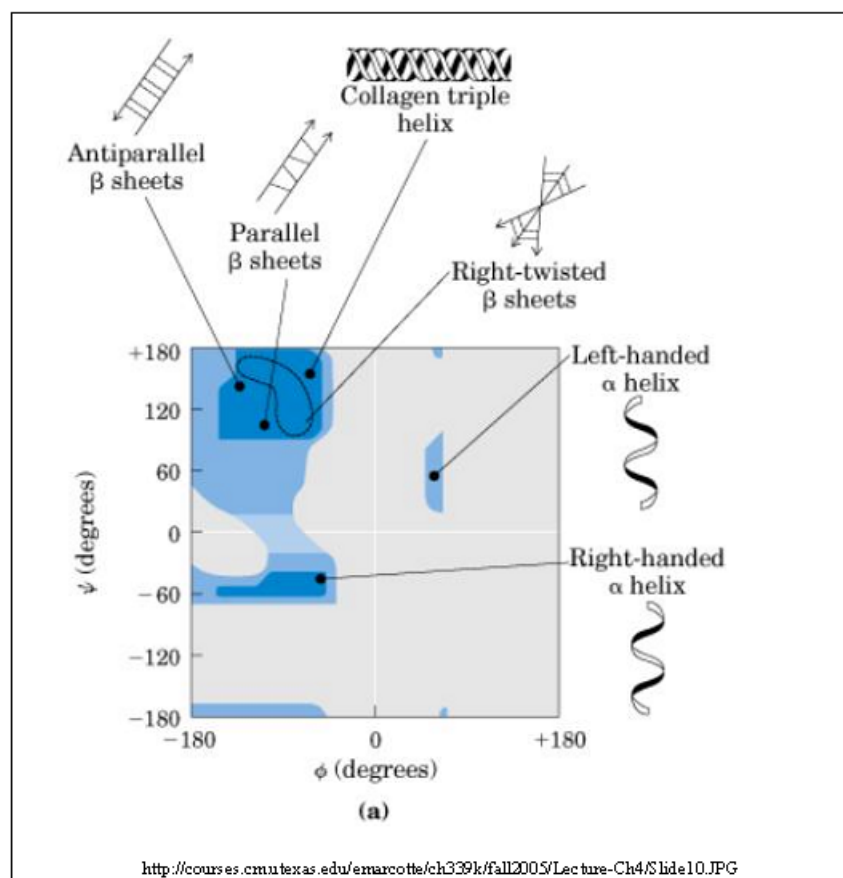


Figure 2. Définition des angles  $\omega$ ,  $\Phi$  et  $\Psi$ .

Si les peptides n'étaient constitués que d'une chaîne principale sans chaîne latérale, presque toutes les conformations (tous les angles  $\Phi$  et  $\Psi$ ) seraient permises, ce qui est le cas pour la glycine dont la chaîne latérale se limite à un hydrogène. Par contre, en raison de l'encombrement stérique plus ou moins important des chaînes latérales, certains couples d'angles  $\Phi\Psi$  sont favorisés et d'autres sont clairement interdits. C'est ce qu'a mis en

évidence Ramachandran au travers des cartes qui portent son nom (**Ramachandran et al., 1963**). Les conformations permises qui composent ces cartes sont calculées pour des tripeptides en considérant les atomes comme des sphères solides (figure 3).

L'espace conformationnel n'est en réalité pas complètement accessible même en l'absence de chaîne latérale; par exemple lorsque  $\Phi$  et  $\psi$  valent tous deux  $0^\circ$ , les deux liens peptidiques (les deux plans) qui entourent le  $C\alpha$  sont dans le même plan, ce qui conduit à une impossibilité stérique entre l'hydrogène de l'amine amont et l'oxygène du carbonyle aval. La carte générale de Ramachandran est donnée à la figure 3. Trois zones sont clairement définies, plus ou moins larges, en fonction de la distance d'approche interatomique autorisée lors du calcul théorique.



**Figure 3. Carte générale théorique de Ramachandran.**

Les cartes de Ramachandran calculées sur base des protéines foldées donnent des résultats différents selon le type de chaîne latérale. La glycine, qui est un résidu particulier parce que sa chaîne latérale est constituée uniquement d'un hydrogène peut occuper quasi n'importe quelle position sur la carte (sauf celles exclues en raison de la géométrie du backbone lui-même). C'est ce qui confère à la glycine ce rôle de « plasticité » dans les structures peptidiques ou protéiques. D'autre part, la proline, on le comprendra aisément, ne

peut occuper que des zones plus restreintes sur la carte. En effet, en raison de son caractère cyclique englobant le C $\alpha$  et le N, la proline n'a pas de possibilité de rotation autour de cette liaison. La proline induit donc plutôt une rigidité dans les structures ainsi qu'une déformation des hélices  $\alpha$ .

Ces combinaisons d'angles  $\Phi$  et  $\psi$ , donnent naissance à différents types de structures secondaires (Creighton, 1993; Branden et Tooze, 1999). Parmi celles-ci on trouve les hélices. Elles sont au nombre de cinq. Premièrement l'hélice  $\alpha$  pour laquelle les angles  $\Phi$  et  $\psi$  valent respectivement  $-57^\circ$  et  $-47^\circ$ . Dans cette conformation, 3,6 résidus sont nécessaires pour faire un tour d'hélice et on observe une translation de  $1,5\text{\AA}$  par résidu. La structure est stabilisée par la formation d'un pont hydrogène entre l'oxygène du carbonyle de l'acide aminé  $i$  et l'azote de l'acide aminé  $i+4$ . Dans le cas de l'hélice  $\alpha$ , la structure est très compacte, favorisant la stabilisation par des interactions intramoléculaires de type van der Waals. Deuxièmement l'hélice  $3_{10}$ , avec des angles  $\Phi$  et  $\psi$  de respectivement  $-49^\circ$  et  $-26^\circ$ , 3 résidus par tour, une translation de  $2\text{\AA}$  par résidu et une stabilisation par formation d'un pont hydrogène entre l'oxygène du carbonyle de l'acide aminé  $i$  et l'azote de l'acide aminé  $i+3$ . Troisièmement l'hélice  $\pi$ , avec des angles  $\Phi$  et  $\psi$  de respectivement  $-57^\circ$  et  $-70^\circ$ , 4,4 résidus par tour, une translation de  $1,15\text{\AA}$  par résidu et une stabilisation par formation d'un pont hydrogène entre l'oxygène du carbonyle de l'acide aminé  $i$  et l'azote de l'acide aminé  $i+5$ . Enfin des hélices polyprolines et polyglycines existent également mais sont plus anecdotiques.

Le deuxième type de structure secondaire est le brin  $\beta$ , qui donne au peptide une conformation quasi totalement allongée, qui s'associe en feuillets parallèles ou antiparallèles. Les brins  $\beta$  sont instables et ce n'est qu'au travers de l'interaction entre brins que la stabilisation par la formation de ponts hydrogènes est possible. Pour le brin  $\beta$  on a une translation de  $3,4\text{\AA}$  par résidu et des angles  $\Phi$  et  $\psi$  de respectivement  $-119^\circ$  et  $+113^\circ$  ou  $-139^\circ$  et  $+135^\circ$  selon que l'on parle de feuillets parallèles ou antiparallèles.

Environ 50% des acides aminés des protéines de la pdb (protein data bank) sont structurés en hélices  $\alpha$  ou en brins  $\beta$  (respectivement 27% et 23%).

Ces structures secondaires sont généralement connectées entre elles par des boucles ou des coudes, habituellement stabilisés par un pont hydrogène entre le CO du résidu  $i$  et le NH du résidu  $i+3$ . Actuellement une bonne dizaine de types de coudes sont clairement identifiés dans les protéines foldées.

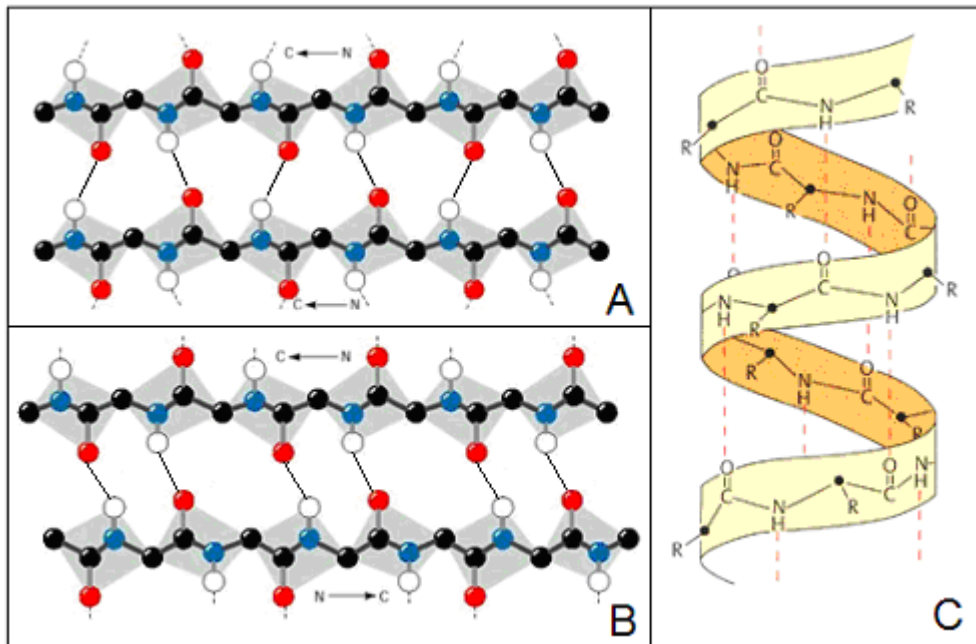


Figure 4. Feuillet  $\beta$  parallèle (A), antiparallèle (B) et hélice  $\alpha$  (C).

Les chaînes latérales bénéficient également d'un certain degré de liberté conformationnelle et sont quant à elles caractérisées par leurs angles Chi ( $\chi$ ) qui prennent habituellement les valeurs de  $+60^\circ$ ,  $-60^\circ$  et  $180^\circ$ . Outre la conformation de la chaîne principale d'un peptide ou d'une protéine, les angles  $\chi$  déterminent donc la position que le groupe fonctionnel occupera dans l'espace. La figure 5 illustre cela pour l'acide glutamique.

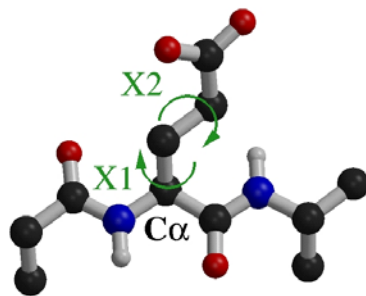


Figure 5. Angles  $\chi$  1 et  $\chi$  2 pour l'acide glutamique .

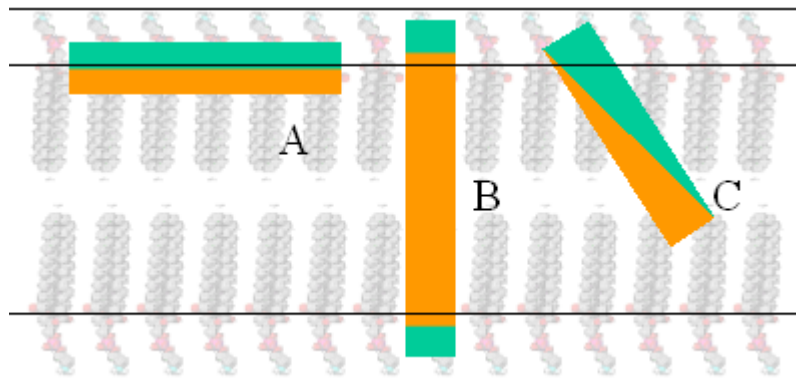
Dès lors, en fonction de son nombre d'angles  $\chi$  qui sont autant de degré de liberté, chaque type de chaîne latérale aura plus ou moins de possibilités conformationnelles. Par exemple, l'alanine avec un angle  $\chi$  unique aura peu de possibilités conformationnelles alors que l'acide glutamique ou la lysine, avec respectivement 3 et 4 angles  $\chi$ , auront plus de possibilités.

### 1.2.2. Interaction des peptides avec les membranes.

En fonction de leur composition en acides aminés, les peptides auront tendance soit à se situer en dehors des membranes, soit à interagir avec elles en adoptant une position particulière dans la membrane.

Les peptides dont la surface accessible au solvant est hydrophile resteront hors de la membrane et préféreront interagir avec l'eau, tandis que les peptides dont la surface accessible au solvant est hydrophobe auront tendance à s'enfuir dans le cœur hydrophobe des membranes constitué par les chaînes acylées des phospholipides. Les peptides amphiphiles quant à eux s'orienteront à l'interface hydrophile/hydrophobe de la membrane de manière à faire pointer les résidus polaires vers les têtes polaires des phospholipides et les résidus apolaires vers le cœur hydrophobe de la membrane (figure 6).

Jusqu'il y a 20 ans, seuls deux types d'interactions peptide/membrane étaient décrits (amphipatiques et adsorbés à la surface membranaire, complètement hydrophobes et enchâssés dans le cœur de la membrane). En 1988, une étude des propriétés hydrophobes de la partie N terminale de la protéine d'enveloppe de NDV (Newcastle Disease Virus), a révélé l'existence d'un troisième type d'interaction membrane/peptide (**Brasseur et al., 1988**).



**Figure 6. Insertion membranaires des peptides amphipatiques (A), transmembranaires (B) et obliques (C).**

Ces peptides particuliers sont amphipatiques et présentent une distribution asymétrique de leur hydrophobicité lorsqu'ils sont sous forme hélicoïdale. Ce gradient d'hydrophobicité explique l'insertion particulière, oblique par rapport à la normale à la membrane, de ces peptides dit « obliques ». Ils ont en général une longueur de 11 à 20 résidus, et ont une hydrophobicité moyenne de 0,2 à 0,9 sur l'échelle consensus d'Eisenberg (**Eisenberg et al., 1982**). L'orientation particulière des peptides obliques dans les membranes (entre 30° et 60°) induit une déstabilisation de l'organisation des phospholipides qui mène

généralement au phénomène de fusion membranaire, raison pour laquelle ces peptides ont également reçu le vocable de "peptides de fusion".

A ce jour des peptides obliques ont été détectés dans nombre de protéines impliquées dans des processus pathogènes tant viraux tels que le SIV (**Horth et al., 1991**), le HIV (**Rahman et al., 1991**), que liés à des maladies neurodégénératives telles que la maladie d'Alzheimer (**Pillot et al., 1997**), mais aussi dans des processus physiologiques normaux tels que la fécondation ou l'action des lipases (**Brasseur, 2000**).

### 1.2.3. Les peptides naturels.

La plupart des organismes vivants font usage de peptides endogènes pour remplir diverses fonctions biologiques. De manière générale, les peptides sont synthétisés, comme les protéines, au travers du cycle classique de la transcription de l'information génétique portée par l'ADN en ARN, la transformation en ARN messenger et la traduction au niveau des ribosomes (**Karp, 1996**). On parlera dans ce cas de peptides ribosomiaux. Notons toutefois que certaines bactéries et certains champignons produisent des peptides par voie non ribosomiale en faisant appel à de gros complexes multi-enzymes de type nonribosomal peptide synthétases (NRPS) (**Finking et Marahiel, 1996 ; Sieber et Marahiel, 2003**).

Les peptides humains sont quant à eux exclusivement produits par les ribosomes bien qu'en réalité, la forme fonctionnelle du peptide résulte le plus souvent du clivage d'un polypeptide plus long ou d'une protéine. C'est le cas par exemple de l'APP (Amyloïd Precursor Protein), dont la coupure enzymatique peut conduire à la libération de peptides p3 par l'action de la sécrétase  $\alpha$  ou à la libération, suite à l'action de la sécrétase  $\beta$ , de peptides amyloïdes A $\beta$ 1-40 et A $\beta$ 1-42 impliqués dans la formation des plaques séniles dans la maladie d'Alzheimer.

Chez l'homme, de nombreux peptides fonctionnels endogènes ont été identifiés. Il s'agit le plus souvent d'hormones. Au nombre de ceux-ci on compte l'insuline qui est probablement le peptide hormone qui a été le plus et depuis le plus longtemps étudié. L'insuline dont la découverte en 1923 valu à Banting et Macleod le prix Nobel de médecine (**Lindahl, 2003**) est utilisée depuis plusieurs décennies dans le traitement du diabète. Il s'agit d'un peptide assez particulier qui prend la forme de deux chaînes peptidiques de 21 et de 30 résidus stabilisées par 3 ponts disulfures. Elle est produite dans les îlots de Langerhans du pancréas par les cellules  $\beta$  sous la forme d'une pro-insuline mono-chaîne de 9kD lorsque le taux de sucre sanguin (la glycémie) est trop élevé. Cette chaîne sera clivée par une enzyme protéolytique pour former le dimère actif. A contrario, le glucagon, autre peptide hormonal

de 29 résidus, est produit par les cellules  $\alpha$  dans les mêmes îlots de Langerhans lorsque le taux de glucose dans le sang est insuffisant. Le maintien de la glycémie à 0,8 et 1,2 g/l est contrôlé par l'interaction de ces deux peptides hormones antagonistes du système endocrinien (**Bauduin, 2001**).

D'autres peptides hormonaux peuvent être cités tels que la vasopressin, hormone anti-diurétique de 9 résidus (**Moss GP, 1974**) ou la substance P de la famille des tachykinin, peptide de 11 acides aminés, impliquée dans nombre de phénomènes biologiques parmi lesquels la vasodilatation et la douleur (**Birney et al, 1995**), ou encore des peptides intestinaux tels que ghreline, gastrine, secretine... (**Konturek PC et Konturek SJ, 2003**)

D'autre part, de nombreux peptides synthétisés par d'autres organismes et ingérés au travers de l'alimentation ont été mis en évidence pour leur bioactivité sur l'homme. C'est le cas par exemple du peptide ovokinine issu de l'ovalbumine des œufs qui présente des propriétés hypotensives pour l'homme. Un autre exemple est celui de l'oryzatensine issu de l'albumine de riz et qui peut jouer un rôle d'immuno-modulateur (**Hartmann et Meisel, 2007**). Parmi les sources exogènes de peptides bioactifs pour l'homme, celle qui a la plus été étudiée à ce jour est le lait. Plusieurs dizaines de peptides bioactifs issus du lait ont été identifiés, tels que le casoplatelin qui est un antithrombotique, la lactoferrine ou la caseidine qui sont des peptides antimicrobiens. (**Clare et Swaisgood, 2000**).

Enfin, de plus en plus d'entreprises pharmaceutiques, comme nous le verrons plus loin, concentrent leurs efforts sur le développement de peptides (ou pseudo-peptides) thérapeutiques basés sur des peptides non humains et montrant une activité intéressante, c'est notamment le cas pour une série de peptides antimicrobiens dont on espère pouvoir se servir à l'avenir pour contourner le problème de la résistance bactérienne aux antibiotiques traditionnels (**Gordon et al, 2005**). D'autres groupes espèrent exploiter des peptides endogènes anti-infectieux. C'est le cas de la société Innova proteomics qui a déposé un brevet relatif à des peptides antibactériens, antimicotiques et antiviraux issus des testicules (**Pineau et al, 2006**) qui présentent en outre l'avantage d'avoir un large spectre d'activité, d'être naturellement résistants et faiblement immunogènes.

#### **1.2.4. Production industrielle de peptides.**

Comme on le verra plus loin, l'utilisation de peptides à des fins thérapeutiques est une voie de plus en plus envisagée par les sociétés pharmaceutiques. Toutefois ce n'est que depuis le début des années 60 que cette possibilité d'exploitation est envisageable. En effet, jusqu'à cette époque, la synthèse peptidique se faisait par voie chimique classique qui nécessitait plusieurs mois pour la synthèse d'un peptide. Ce n'est qu'en 1963, avec la



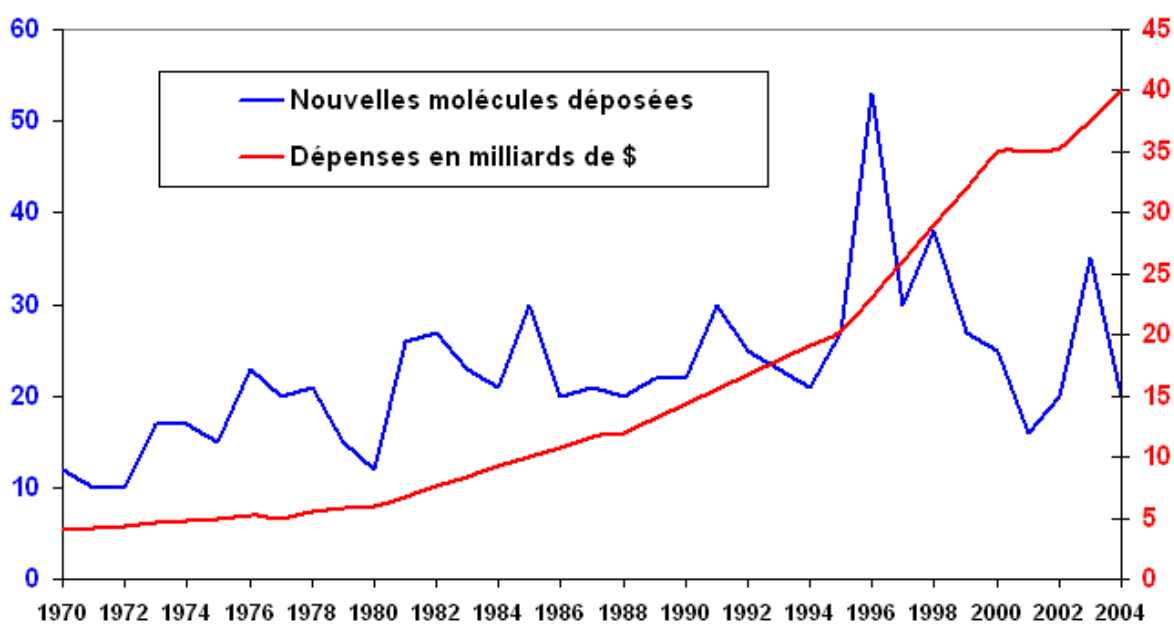
méthode de synthèse par voie solide imaginée par Merrifield et qui lui valut le prix Nobel de chimie en 1984 (**Merrifield, 1963**), que l'utilisation industrielle de peptides synthétiques a pu être envisagée. Ceci d'autant que des techniques de séparation/purification performantes telles que l'HPLC (chromatographie liquide haute performance) (**Chen et Horvath, 1995**) ou la SPE (extraction de phase solide) (**Kamysz et al., 2004**) prenaient elles aussi leur essor. Ces méthodes de synthèse par voie chimique en phase liquide ou solide sont souvent privilégiées par les sociétés pharmaceutiques car ce sont les seules qui permettent l'utilisation efficace d'acides aminés non naturels. Toutefois d'autres méthodes faisant appel aux techniques recombinantes sont parfois envisagées car elles sont moins chères pour des peptides d'une longueur supérieure à 25 acides aminés (**Mehta, 2004**). Leur mise au point à l'échelle industrielle reste toutefois longue et difficile (**Latham, 1999**).

### **1.2.5. Utilisation thérapeutique des peptides.**

Depuis la nuit des temps, l'homme cherche à se soigner. Depuis toujours et encore jusqu'à ce jour, les médecins et chercheurs ont développé une chimie du médicament essentiellement basée sur l'exploitation de principes actifs issus d'organismes naturels tels que les plantes. Cette source d'inspiration pour les scientifiques est en voie d'épuisement.

#### **a) Contexte économique.**

Les rapports des analystes financiers, qui parlent de "pipeline problem" montrent clairement que la créativité des entreprises pharmaceutiques s'essouffle et qu'un grand nombre de brevets d'exploitation sont en fin de vie alors qu'un nombre de plus en plus réduit de nouvelles molécules thérapeutiques sont présentées aux autorités compétentes. Aux Etats Unis, le nombre de nouvelles substances présentées à la FDA (Food and Drug Administration) a diminué de 60% sur les 10 dernières années. (**Sorowiecki, 2004**). Un rapport de 2006 du Congrès américain montre que depuis une vingtaine d'années les budgets R&D des compagnies pharmaceutiques augmentent pour atteindre plus de 20% de leurs revenus (pour seulement 5% en moyenne pour les autres secteurs industriels) alors que dans le même temps le nombre de nouveaux médicaments diminue (**figure 7 A et B**) (**Congress of the United States, 2006**). L'analyse de ces données indique clairement que l'innovation dans l'industrie pharmaceutique est de plus en plus compliquée, risquée, longue (de l'ordre de 10 ans) et coûteuse.



**Figure 7. Dépenses en recherche pharmaceutique aux USA (10<sup>9</sup>\$) par an (A) Nouvelles molécules déposées à la FDA par an (B).**

En effet selon le même rapport, les coûts de recherche ont véritablement explosé ces dernières années, passant de 100 millions d’euros par molécule en 1975 à 600 millions d’euros en 2000, même si selon certains auteurs ces chiffres sont surévalués et devraient être divisés par deux (**Goozner, 2004, Prescrire, 2003**).

Pour contourner ces problèmes, de nombreuses sociétés pharmaceutiques se sont tournées vers les peptides ou les dérivés des peptides (peptidomimétiques) comme agents thérapeutiques (**Lathan et al, 1999**). L’objectif étant le médicament parfait, ce qui s’apparente à la recherche du Graal. En effet, cette molécule idéale doit présenter un faible poids moléculaire, être soluble, capable de mimer parfaitement tant chimiquement que structuralement un ligand naturel, ne pas induire d’effets secondaires et être éliminée aussi rapidement que possible de l’organisme. En outre sa production doit être facile, peu onéreuse et son administration au patient doit être aisée (**Edwards et al, 1999**). Les peptides peuvent, sous certaines conditions, répondre à ces critères et l’analyse de la situation et des moyens de recherches consacrés montre que les sociétés pharmaceutiques croient en l’avenir de cette filière. En effet en 2004 plus de 20% des médicaments appartenant au top 200 des ventes étaient à base de protéines ou de peptides, avec des chiffres d’affaires atteignant les 30 milliards d’euros (**Med Ad News, 2005**), soit environ 10% du chiffre global de l’industrie pharmaceutique. A la même période, entre 600 et 700 peptides étaient en stade de développement et plus de 150 étaient à différents stades de phases cliniques (**McGee, 2005** ;

**Ayoub et Scheidegger, 2006**). Les peptides thérapeutiques sont présents dans différents domaines tels que le traitement de certaines formes de cancer, du SIDA, de l'ostéoporose, ou de maladies neurodégénératives (**Lien and Lowman, 2003**). Le chiffre d'affaire annuel généré par les peptides thérapeutiques est de 4,8 milliards d'euros. Les peptides offrent à la fois des avantages et des inconvénients par rapport aux médicaments traditionnels ( petites molécules).

#### **b) Avantages des peptides.**

Parmi ces avantages, on peut citer le fait que les peptides thérapeutiques sont souvent la plus petite partie fonctionnelle d'une protéine et offrent donc une efficacité, une sélectivité et une spécificité qu'il est difficile d'atteindre avec des petites molécules (**Hummel et al., 2006**). On peut citer le cas de NAP qui est un peptide de 8 acides aminés actuellement en phase d'essais cliniques qui constitue le plus petit fragment actif de ADNP (Activity Dependent Neuroprotective Protein) et qui montre une activité neuroprotectrice (**Gozes and Spivak-Pohis, 2006**). Deuxièmement les produits de dégradation des peptides sont des acides aminés, ce qui diminue grandement les risques de toxicité. Troisièmement les peptides s'accumulent peu dans les tissus, diminuant d'autant les risques de toxicité. De plus ils offrent une diversité structurale supérieure aux petites molécules (**Loffet, 2002 ; Ayoub et Scheidegger, 2006**). Enfin, et ce n'est pas anecdotique pour le secteur du médicament, les peptides médicaments sont innovants et il y a donc moins de risques qu'ils soient déjà protégés par un brevet.

#### **c) Inconvénients des peptides.**

Si les peptides présentent des avantages, leur conception pour une utilisation thérapeutique se heurte à de nombreux écueils. Un des problèmes est le coût de production à l'échelle industrielle que ce soit par des méthodes chimiques ou biologiques, coût qui est plus élevé que celui des petites molécules. Par ailleurs, la quantité nécessaire est plus importante, notamment lorsque l'administration orale au patient n'est pas possible (**Latham, 1999**).

Ensuite, la fameuse règle des cinq (**Lipinski et al., 2001**), qui reste une référence pour les modélisateurs de molécules thérapeutiques, n'est pas respectée pour les peptides, notamment pour ce qui concerne le poids moléculaire. A ce titre on considère que les peptides sont peu administrables par voie orale car ils seraient peu enclins à passer du système digestif vers le système sanguin. C'est la raison pour laquelle l'administration des

peptides est à ce jour principalement de type intraveineuse, ce qui n'est pas particulièrement confortable pour le patient. Toutefois plusieurs stratégies alternatives de délivrance ont été développées ces dernières années (**Lien et Lowman, 2003**). C'est le cas par exemple de l'administration par les muqueuses que ce soit par voie nasale (**Maggio, 2005**) ou pulmonaire (**Skyler et al., 2001**). D'autres groupes ou sociétés telles que Roche avec sa formulation pegylée d'interféron (Pegasys<sup>TM</sup>), ont mis au point des enrobages permettant le relargage retardé du peptide ou encore la société Emisphere qui a développé la technologie eligen<sup>R</sup> qui permet d'enrober le peptide pour permettre un transfert du tube digestif vers le sang, autorisant une prise orale (**Malkov et al., 2005**). D'autres techniques telles que celles mettant en œuvre des injecteurs à haute pression permettant au peptide de traverser la peau sans utilisation d'aiguilles ont également été développées (**Stout et al., 2007**). Enfin les stratégies faisant appel à des vecteurs (cell penetrating peptides) associés de manière covalente ou non à des peptides thérapeutiques sont en plein développement.

Un second aspect négatif des peptides thérapeutiques réside dans le fait qu'ils sont très rapidement dégradés et que leur durée de vie dans l'organisme humain est de l'ordre de quelques minutes ce qui est largement insuffisant pour atteindre la cible et agir de façon optimale (**Latham, 1999**). La plupart du temps cette dégradation est due aux protéases ou peptidases principalement localisées dans le sang, le foie ou encore les reins (**Werle et Bernkop-Schnürch, 2006**). Différents moyens ont été imaginés pour contrecarrer cette action enzymatique. Au nombre de ceux-ci on retrouve les C-amidations et les N-acétylations pour empêcher l'action des exopeptidases ainsi que l'utilisation d'acides aminés D,  $\beta$  ou non naturels, moins sensibles à la dégradation enzymatique (**Fauchère et Thuriau, 1992 ; Lathan, 1999**). Les cyclisations au travers de ligations Nter-Cter, d'anneaux lactames ou de ponts disulfures sont également proposées (**Werle et Bernkop-Schnürch, 2006**). Enfin, d'autres modifications chimiques des peptides telles que l'attachement d'acides gras ou la pegylation (attachement d'unités de polyéthylène glycol) ont pour but d'en augmenter le poids moléculaire et donc le temps de demi-vie (**Caliceti et Veronese, 2003, Harris and Chess, 2003, Veronese et Passut, 2005**). Ces molécules ont des propriétés intéressantes pour des applications pharmaceutiques car elles sont solubles, non toxiques, peu immunogènes et sont facilement éliminées par l'organisme. Elles conservent ces propriétés lorsqu'elles sont associées au peptide d'intérêt. Finalement la co-administration d'un peptide avec des inhibiteurs d'enzyme est également une voie qui a été explorée (**Werle et Bernkop-Schnürch, 2006**).

Par ailleurs il a été montré qu'un peptide peut conserver sa fonction lorsqu'il est partiellement composé d'acides aminés non naturels (**Benkirane et al., 1993**) et qu'il est

possible de modéliser des peptides non naturels présentant une fonction ou un comportement donné (**Lins et al., 2006**).

En conclusion, au vu des moyens investis dans la recherche peptidique par les entreprises pharmaceutiques, il va sans dire que cette nouvelle voie thérapeutique est considérée comme hautement prometteuse.

#### **d) Conception des peptides thérapeutiques.**

La plupart du temps les nouveaux médicaments traditionnels, ou petites molécules, sont conçus selon une stratégie basée sur la structure de la cible, on parlera de "structure based design". Toutefois, lorsque la structure de la cible est inaccessible, on se basera sur la structure du ligand ("ligand based design"). A priori il n'en va pas autrement pour la conception d'un peptide ou d'un dérivé peptidique thérapeutique.

Certains peptides naturels tels que l'insuline sont utilisés à des fins thérapeutiques. Ces peptides actifs, appelés "leads", peuvent être adaptés et modifiés ultérieurement afin d'accroître l'efficacité ou les propriétés ADME-T (Absorption, Distribution, Métabolisme, Excrétion, Toxicité).

Dans d'autres cas, le peptide n'est pas utilisé directement comme agent thérapeutique mais comme un outil ou un modèle pour la conception de molécules thérapeutiques hybrides, à mi-chemin entre le peptide et la petite molécule. Ces molécules sont généralement appelées "peptidomimétiques" ou "pseudo-peptides".

La stratégie de recherche se décompose généralement en plusieurs étapes (**Hruby, 2002**). La première consiste à rechercher et à identifier la cible, qui en général est une protéine. Notons que si le développement fait appel à des méthodes de modélisation "in silico", la connaissance de la structure de la cible est indispensable, qu'elle soit donnée par méthode expérimentale ou par modélisation, l'idéal étant évidemment de pouvoir procéder à la co-cristallisation de la cible et de son ligand naturel).

La deuxième étape consiste à identifier l'effecteur (ligand, peptide bioactif ou protéine partenaire). Troisièmement, on s'attachera à la recherche du segment bioactif le plus court. Ensuite on utilisera des analogues peptidiques de ce segment court qui seront structurellement contraints, par exemple en méthylant le C<sub>α</sub> (**Sagan et al., 2001**) ou le C<sub>β</sub> (**Birney et al., 1995**), ou encore en le cyclisant afin d'établir la topologie d'un pharmacophore. Il faudra donc identifier les fonctions chimiques importantes et les situer dans l'espace par rapport au site de liaison de la cible. Finalement on fera un criblage de banques de petites molécules respectant les propriétés chimico-spatiales du pharmacophore

(chimiotèques). L'affinité de ces molécules sera alors évaluée par des méthodes de "docking" qui tiendront compte de la complémentarité stérique, électrostatique et hydrophobe entre le site actif de la cible et la molécule candidate.

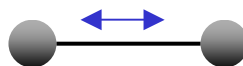
Notons également que grâce à la technique du « phage display » (Smith, 1985), le criblage de bibliothèques de peptides est également possible pour la recherche d'un peptide bioactif (Cortese et al, 1996). Ceci permet d'identifier des peptides fonctionnels à une phase précoce du développement du médicament potentiel et d'obtenir la définition d'un pharmacophore.

### 1.3. Les champs de forces.

On peut définir un champ de forces comme l'ensemble des forces qui s'appliquent à un atome à un endroit donné de l'espace. Les champs de forces sont utilisés dans toutes les approches de modélisation moléculaire, en particulier dans les méthodes de dynamiques moléculaires telles que CHARMM (Chemistry at HARvard Macromolecular Mechanics) développé par le groupe de Karplus (Brooks et al., 1983) ou GROMACS (Berendsen et al., 1995) ou d'optimisation de structures telle que HYPERCHEM (Froimowitz, 1993) qui utilisent généralement des champs de forces de type AMBER (Assisted Model Building with Energy Refinement). Le champ de forces développé par le Professeur Brasseur au CBMN fait appel à des équations du même type mais tient compte en sus de la contribution hydrophobe. L'énergie totale d'un système est la somme des énergies d'interaction entre atomes non liés et des énergies d'interaction entre atomes liés.

#### 1.3.1. Atomes liés.

- L'énergie liée à la compression/élongation (stretching) d'une liaison est approximée par une fonction de type harmonique et calculée en utilisant la loi de Hooke (équation 1, figure 8).



$$\text{(Eq. 1)} \quad E = \frac{1}{2} K_r (r - r_0)^2$$

où  $r$  est la distance entre les atomes,  $r_0$  la distance à l'équilibre et  $K$  une constante dépendant de la nature du couple atomique.

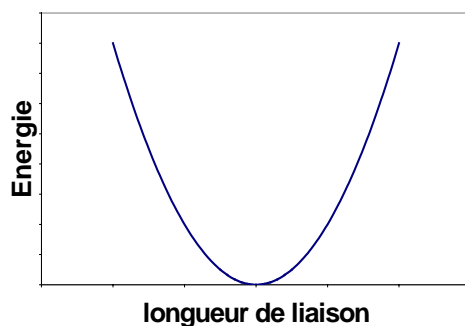
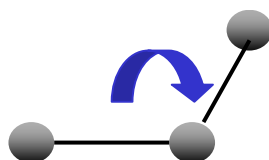


Figure 8. Energie en fonction de la longueur de liaison.

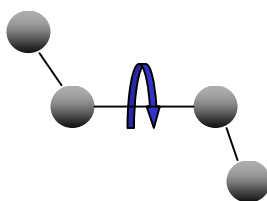
- L'énergie liée à la déformation des angles (bending) est exprimée par une équation du même type que pour le stretching.



$$(Eq. 2) E = \frac{1}{2} K_{\theta} (\theta - \theta_0)^2$$

où  $\theta$  est l'angle de la liaison,  $\theta_0$  l'angle de liaison à l'équilibre et  $K_{\theta}$  une constante liée à la nature de la liaison.

- L'énergie liée à la torsion des angles dièdres est donnée par une fonction de type périodique (équation 3, figure 9).



$$(Eq.3) E = \frac{V_n}{2} [1 + \cos(n\phi - \phi_0)]$$

où  $V_n$  est une constante de force dièdre,  $n$  est la périodicité,  $\Phi$  est la valeur de l'angle dièdre et  $\Phi_0$  est la phase.

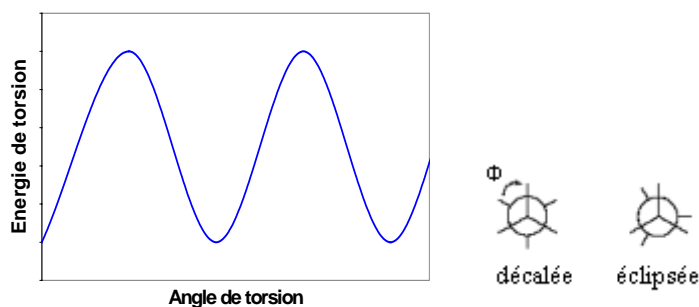


Figure 9. Périodicité de l'énergie pour les angles dièdres.

### 1.3.2. Atomes non liés.

- Les interactions de van der Waals interviennent entre dipôles et expliquent qu'en fonction de leur distance, deux atomes s'attirent lorsqu'ils sont trop éloignés ou se repoussent lorsqu'ils sont trop proches (figure 9) (**Dill et Bromberg**). L'énergie d'interaction de van der Waals est généralement exprimée sous la forme d'un potentiel de Lennard-Jones qui comporte un terme attractif et un terme répulsif. L'équation 4 est utilisée pour le calcul de la contribution de van der Waals à l'interaction entre deux entités moléculaires (**équation 4, figure 10**).

$$(eq.4) \quad E_{vdW} = \sum_{ij} \left[ A_{ij} \left( \frac{r_i^0 + r_j^0}{d_{ij}} \right)^{12} - B_{ij} \left( \frac{r_i^0 + r_j^0}{d_{ij}} \right)^6 \right]$$

où  $A_{ij}$  et  $B_{ij}$  sont des coefficients assignés au paires d'atomes,  $r_i^0$  et  $r_j^0$  sont les rayons de van der Waals des atomes  $i$  et  $j$ , et  $d_{ij}$  est la distance entre les atomes  $i$  et  $j$ .

Les interactions de van de Waals sont le reflet de l'ajustement stérique entre atomes non liés. Elles interviennent donc à très courte distance.



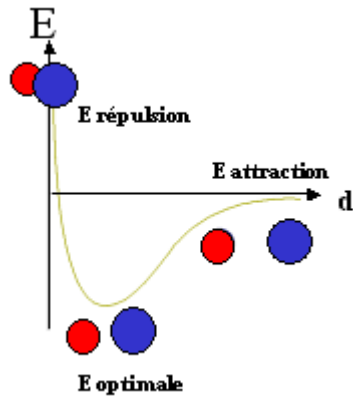


Figure 10. Energie d'Interaction de van der Waals

- L'énergie électrostatique est relative à l'interaction entre atomes chargés. Elle est calculée au moyen de la relation de Coulomb (équation 5, figure 11).

$$(eq.5) \quad E_{elec} = \lambda \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{q_i q_j}{\epsilon_{ij}(z) d_{ij}}$$

où  $\lambda$  est le facteur de conversion d'unité,  $d_{ij}$  est la distance entre les atomes  $i$  and  $j$ .  $\epsilon_{ij}(Z)$  la constante diélectrique qui varie de 1 à 80 suivant une fonction sigmoïdale de  $d_{ij}$  entre 2 and 10 Å,  $q_i$  et  $q_j$  sont les charges des atomes  $i$  et  $j$ .

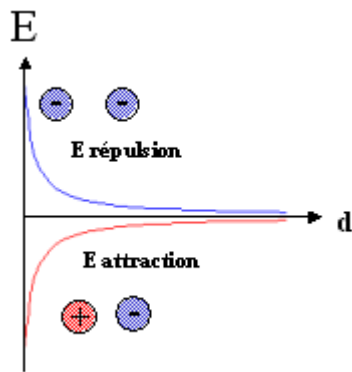
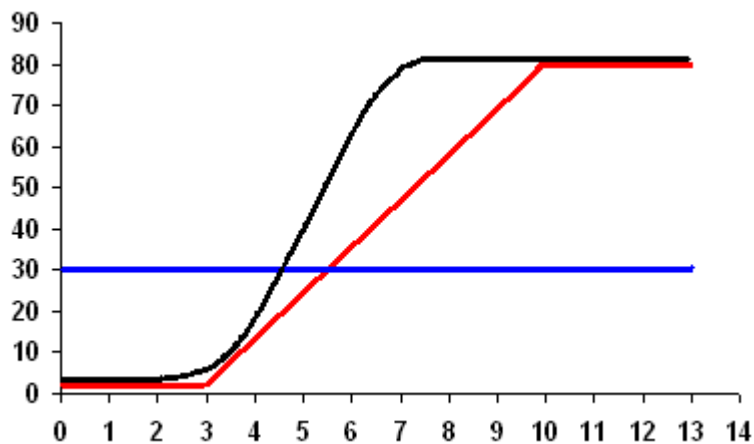


Figure 11. Interactions électrostatiques.

L'énergie d'interaction électrostatique varie avec la distance entre atomes chargés, de telle sorte que deux charges de même signe se repoussent et deux charges de signes opposés s'attirent. D'autre part, cette contribution énergétique varie avec la constante diélectrique ( $\epsilon$ ). Cette "constante" dont la valeur dépend du milieu peut varier de 1 en milieu hydrophobe à 80 en milieu hydrophile. En général, dans les champs de forces utilisés en modélisation,  $\epsilon$  est maintenue constante ou varie en fonction de la distance. La manière dont on simulera la variation de la constante diélectrique influencera la valeur de l'énergie des

interactions électrostatiques. Les fonctions les plus couramment utilisées sont de type constant à une valeur fixe, linéaire ou sigmoïdal (**figure 12**)



**Figure 12. Evolution linéaire (rouge) et sigmoïdale (noir) et valeur fixe (bleu) de la constante diélectrique ( $\epsilon$ ) en fonction de la distance interatomique.**

- Enfin, les ponts hydrogènes sont le troisième type d'interaction entre atomes non-liés. Ce type de lien non covalent peut s'établir entre un donneur d'hydrogène et un accepteur. Le donneur est un atome électronégatif portant un hydrogène (-NH, -CO, -SH) et l'accepteur dispose d'au moins un doublet libre (N, O, F). Par ailleurs, l'aspect directionnel entre l'hydrogène et le doublet libre est importante pour la formation des ponts hydrogènes, ce qui n'est pas le cas pour les autres types d'interactions. Les ponts hydrogènes sont des éléments stabilisateurs de la structure des protéines et des peptides. Les fonctions énergétiques se rapportant aux ponts hydrogènes sont du même type que celles relatives aux interactions de van der Waals ( $d^{10}$  au lieu de  $d^6$ ).

- Une des spécificités du champs de forces du CBMN est qu'il tient compte des interactions hydrophobes. Les molécules hydrophobes (apolaires) ont tendance à s'agréger de manière à limiter au mieux leur surface de contact avec l'eau. Ce phénomène est généralement simulé par deux types de fonctions, celle proposée par Fauchère et celle proposée par Brasseur (**Brasseur, 1995**) (équation 6). Les deux sont utilisées au C.B.M.N.. La valeur de la fonction développée par Brasseur décroît en fonction de la distance et prend en compte l'énergie de transfert des atomes mis en jeu.

(eq.6)

$$E_{pho} = \sum_{i=1}^{N-1} \sum_{j=i+1}^N \delta_{ij} \left[ (Etr_i f_{ij}) + (Etr_j f_{ji}) \right] \exp\left( \frac{r_i^0 + r_j^0 - d_{ij}}{2 r_{sol}} \right)$$

où  $\delta_{ij} = -1$  si les deux atomes en interaction sont hydrophobes ou hydrophiles et  $\delta_{ij} = +1$  si les atomes  $i$  and  $j$  sont de type opposé.  $Etr_i$  et  $Etr_j$  sont les énergies nécessaires pour le transfert des atomes  $i$  et  $j$  d'une phase hydrophile vers une phase hydrophobe;  $f_{ij}$  et  $f_{ji}$  sont les portions de surface atomique couverte par le partenaire dans l'interaction;  $r_i^0$  et  $r_j^0$  sont les rayons de van der Waals des atomes  $i$  et  $j$ ,  $d_{ij}$  est la distance entre  $i$  et  $j$  et  $r_{sol}$  est le rayon d'une molécule d'eau.

Les énergies de transfert sont données à la table 1. Elles dépendent de la nature des atomes concernés. Sept types atomiques ont été définis, correspondant aux différents types d'atomes des protéines.

<b>Types atomiques</b>	<b>Etr (kCal/mol)</b>
<b>Csp3</b>	<b>-2.436</b>
<b>Csp2</b>	<b>-1.513</b>
<b>H (q=0)</b>	<b>- 0.537</b>
<b>H (q)</b>	<b>1.030</b>
<b>O</b>	<b>2.833</b>
<b>N</b>	<b>3.035</b>
<b>S</b>	<b>-2.751</b>

**Table 1. Energie de transfert des différents types atomiques**

## **2. OBJECTIF**

Les peptides sont utilisés par les chercheurs dans des domaines variés de la biotechnologie, notamment pour la production d'anticorps ou comme marqueurs à l'usage des laboratoires de biochimie.

Plus récemment, dans l'espoir de réalimenter leur pipe-line en voie de tarissement avec de nouvelles molécules innovantes, des sociétés pharmaceutiques se sont intéressées aux peptides pour leurs potentialités thérapeutiques. On sait que des peptides peuvent fonctionner comme agonistes ou comme inhibiteurs mais aussi que certains d'entre eux, les AMPs (pour Antimicrobial peptides) ont des propriétés antibiotiques tandis que d'autres, les CPPs (pour Cell Penetrating Peptides), sont d'excellents vecteurs de molécules thérapeutiques.

Jusqu'il y a peu les médicaments étaient développés par une méthodologie de type essais/erreurs à partir de substances naturelles issues d'organismes tels que les plantes ou en se basant sur des méthodes de screening réel ou virtuel de banques de molécules organiques. Quoiqu'il en soit le développement des médicaments traditionnels, les « petites molécules », a fait l'objet de toutes les attentions et de nombreuses méthodes « *in silico* » du type docking ont été développées. Elles permettent de sélectionner le meilleur ligand pour une cible donnée.

Par contre, en dépit de l'avenir prometteur des peptides thérapeutiques, peu de méthodes de prédiction « *in silico* » ont vu le jour. Dans le cadre de cette thèse nous souhaitons nous intéresser à deux aspects primordiaux en rapport avec le design « *in silico* » de peptide.

Le premier aspect se rapporte à la conception rationnelle de peptides capables de se lier avec une bonne affinité à une cible déterminée. Les méthodes existantes à ce jour sont soit expérimentales et très coûteuses, c'est le cas de celles faisant appel à la production d'un grand nombre de peptide telles que le phage display, soit basées sur la dynamique moléculaire et donc très gourmandes en temps de calculs. Dans le cadre de cette thèse, nous souhaitons développer une méthodologie plus rapide et moins onéreuse, basée sur l'utilisation combinée de différentes méthodes déjà disponibles et éprouvées au sein du laboratoire. La stratégie que nous avons imaginée est développée à la section 3.1.1. A titre d'application nous voulions tenter de concevoir un peptide complémentaire affin pour le peptide de fusion A $\beta$ 29-42 en prenant pour modèle la 7<sup>ème</sup> hélice de l'ApoE.

Les résultats de ces travaux ont abouti à la rédaction d'un article et au dépôt d'un brevet de protection des peptides modélisés.

Le second aspect auquel nous souhaitons nous intéresser est relatif à la structure des peptides. On sait que la structure et la fonction des peptides sont intimement liées, comme elles le sont pour les protéines. De nombreuses méthodes expérimentales de détermination de la structure peptidique telles que la RMN, le CD, la RX, ... existent. Les résultats qu'elles donnent pour un même peptide sont souvent divergeants et clairement liés aux conditions expérimentales utilisées. D'autre part ces méthodes sont souvent lourdes à mettre en œuvre et les expérimentateurs sont confrontés à des problèmes tels que l'agrégation ou les difficultés de cristallisation. Ceci indique, qu'en raison de leur faible capacité à s'auto-stabiliser au travers d'interactions intramoléculaires, les peptides n'ont que rarement une structure unique, celle-ci variant avec le milieu et les possibilités d'interactions avec ce milieu. Dans ce domaine aussi, les méthodes de prédiction *in silico* de structure des peptides font défaut puisque seules deux méthodes étaient connues à l'entame de notre travail.

A priori, les temps de calcul des méthodes de prédiction de structure sont nécessairement longs puisque qu'une des caractéristiques des peptides est qu'ils peuvent, en théorie, adopter une quasi-infinité de conformations. Il est donc impossible de tester et de calculer l'énergie conformationnelle de toutes les combinaisons d'angle  $\Phi\Psi$  d'un peptide avec les moyens informatiques dont nous disposons actuellement. Dans le cadre de cette thèse nous avons donc cherché à développer une méthode de prédiction de structures peptidiques basée sur des outils innovants, permettant de limiter de manière rationnelle l'espace conformationnel à explorer. Cette procédure, que nous avons appelée PepLook, a été utilisée dans deux applications. Dans la première application nous souhaitons étoffer les informations quant aux propriétés nécessaires à la capacité d'internalisation des CPPs ou peptides vecteurs (section 3.2.2.), dans le but de fournir un nouvel outil d'aide à la décision pour la conception de nouveaux peptides vecteurs. Dans la seconde application, dans le cadre d'une collaboration avec l'équipe du Professeur Epanand à Hamilton (Canada), nous voulions identifier la structure du segment membranaire de la protéine DGK $\epsilon$  de structure inconnue afin de proposer une topologie cohérente et vérifiable de manière expérimentale (section 3.2.3.).

### **3. RESULTATS**

### **3.1. Aspect design.**

#### ***3.1.1. Modélisation rationnelle de peptides complémentaires au peptide de fusion A $\beta$ 29-42: Une application de PepDesign.***

*Decaffmeyer M., Lins L., Charloreaux B., VanEyck MH., Thomas A., Brasseur R. (2006). Rational design of complementary peptides to the  $\beta$ Amyloid 29-42 fusion peptide: an application of PepDesign. Biochim Biophys Acta. Mar;1758(3):320-7.*

Plusieurs programmes de modélisation *in silico*, principalement des méthodes de docking telles que DOCK, AutoDock, FlexX existent pour la conception de petites molécules (**Ghosh et al., 2006**). A contrario, et en dépit de l'avenir prometteur des peptides, leur modélisation *in silico* est restée à ce jour une voie fort peu explorée.

Une des voies actuellement utilisée pour le choix rationnel de peptides fait appel au criblage de banques peptidiques. Plusieurs groupes ont réussi à trouver des peptides très affins pour leur cible. Ils utilisent généralement des banques peptidiques obtenues par phage display ou par synthèse en phase solide et s'appuient sur une méthode appelée "systematic residue replacement" (SRR) qui consiste à identifier un lead dans la banque de peptides, puis à caractériser son interaction avec la cible par exemple par cristallographie RX et enfin à remplacer certains résidus de manière systématique à la recherche du peptide le plus efficace. C'est le cas de travaux qui ont mené à la conception de peptides de haute affinité liant l' $\alpha$ -Bungarotoxin ( $\alpha$ -BTX), une neurotoxine de serpent. Un des peptides conçus inhibe l'interaction de l' $\alpha$ -BTX avec le récepteur acétylcholine (AChR) à la jonction neuromusculaire (**Kasher et al., 2001 ; Katchalsky et al., 2003**).

L'autre procédure actuellement utilisée se base sur la dynamique moléculaire (DM). Yang et ses collaborateurs ont modélisé un peptide antagoniste de l'interleukine-6 humaine (hIL-6) en s'appuyant sur l'analyse de la trajectoire en DM (**Yang et al., 2005**)

Il semble que ces deux approches donnent d'excellents résultats. Toutefois, la première requiert un grand nombre de peptides et nécessite un grand nombre de tests biologiques, ce qui rend l'approche onéreuse. La seconde méthode implique de très longs temps de calculs et de très grandes puissances de calcul. Ce facteur temps limite le nombre de peptides testés dans des délais compatibles avec la réalité économique de l'industrie du médicament.



Afin de contourner ces problèmes nous nous sommes proposés de mettre au point une méthode *in silico* rapide et de modélisation de peptides capables de lier une cible. Cette méthode, que nous avons baptisée PepDesign, est basée sur le même principe de mutation de résidus que dans la méthode SSR, toutefois, il s'agit d'une approche entièrement *in silico*. PepDesign s'appuie sur la conjonction de méthodes déjà développées et utilisées par le CBMN telles que la méthode hypermatrice (**Brasseur et al., 1981**) pour la détermination du complexe peptide/cible, la méthode PEX (**Thomas et al., 2001**) pour l'analyse de l'interaction et l'identification des résidus à muter ainsi que la dynamique angulaire (DA) (**Brasseur, 1995**) et la procédure de Monte Carlo pour l'optimisation des complexes mutants/cible. Une vue schématique de la procédure est donnée à la figure 13 (**figure 13**).

Nous décrivons dans l'article qui suit la manière dont nous avons appliqué cette méthode pour concevoir des peptides complémentaires au peptide amyloïde 29-42 (A $\beta$  29-42). On sait que ce peptide est impliqué dans la formation des plaques séniles caractéristiques de la maladie d'Alzheimer (**Shoji, 2002**). Il présente par ailleurs une activité fusogène sur les liposomes (**Pillot et al., 1996**). Nous nous sommes basés sur un peptide modèle qui est la 7<sup>ème</sup> hélice de l'Apolipoprotéine E (ApoE 270-287). On sait que l'ApoE interagit avec le peptide A $\beta$  29-42 par sa partie C terminale (**Pillot et al., 1999**) et plus particulièrement par sa 7<sup>ème</sup> hélice (**Lins et al., 1999**). ApoE 270-287 inhibe, quoique de manière imparfaite, la fusion engendrée par A $\beta$  29-42. Les travaux de Lins et ses collaborateurs ont permis d'identifier le meilleur complexe entre la 7<sup>ème</sup> hélice de ApoE et A $\beta$  29-42 et de montrer que l'interaction entre les deux peptides est essentiellement de type hydrophobe et que de surcroît elle est spécifique (**Lins et al., 1999**).

La première étape du travail consiste à reproduire cet assemblage et à en caractériser les interactions. Cinq résidus de ApoE 270-287 ont été identifiés comme étant responsables de l'interaction.

Dans la deuxième étape nous avons proposé des mutations pertinentes en terme d'hydrophobicité puisque l'interaction est de type hydrophobe (Lins et al., 1999) et en terme de propension à former une hélice.

Ensuite, 1024 peptides mutants de ApoE 270-287, correspondant à 4 mutations sur 5 positions, ont été créés et les complexes mutants/A $\beta$  29-42 ont été optimisés pour chacun d'entre eux en combinant la DA pour l'optimisation des chaînes latérales et le Monte Carlo (MC) pour l'optimisation de la position relative des peptides. Cette étape a mené à la sélection de 6 peptides.

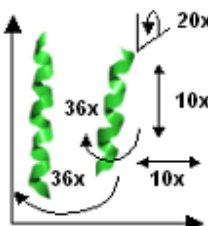
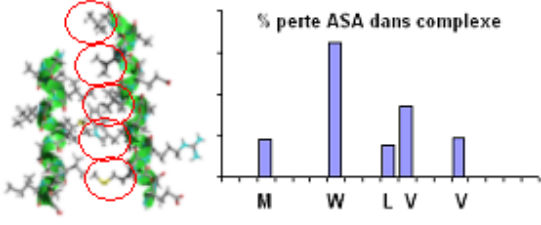
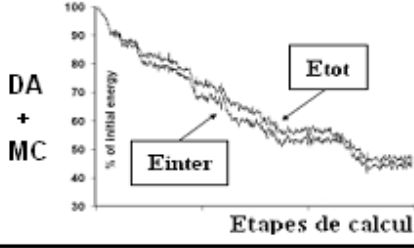
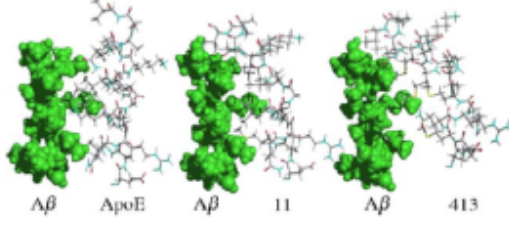
Procédure multi-étapes de PepDesign																	
1	Calcul du complexe par <u>Hypermatrice</u>	 <p><b>5 degrés de liberté.</b>            36 rotations.            36 self-rotations.            20 angles.            10 translations horiz.            10 translations vertic.</p>															
2	Identification des résidus à muter par <u>PEX</u> ( $\delta$ ASA)	 <p>% perte ASA dans complexe</p> <table border="1"> <tr><td>M</td><td>W</td><td>L</td><td>V</td><td>V</td></tr> </table>	M	W	L	V	V										
M	W	L	V	V													
3	Choix des résidus de substitution	<p>E D M Q R O W A G L V E K V Q A A V</p> <p><b>M M M M</b></p> <p><b>L L L L</b></p> <p><b>V V V V</b></p> <p><b>W W W W</b></p>															
4	Optimisation des complexes A $\beta$ <sub>29-42</sub> /mutants	 <p>DA + MC</p> <p>% of initial energy</p> <p>Etot</p> <p>Einter</p> <p>Etapes de calcul</p>															
5	Sélection fine des mutants	 <p>A<math>\beta</math> ApoE A<math>\beta</math> 11 A<math>\beta</math> 413</p>															
6	Tests in vitro	<table border="1"> <thead> <tr> <th></th> <th>Fusion</th> <th>Inhibition</th> </tr> </thead> <tbody> <tr> <td>A<math>\beta</math> seul</td> <td>100%</td> <td>0%</td> </tr> <tr> <td>A<math>\beta</math> + WT</td> <td>89%</td> <td>11%</td> </tr> <tr> <td>A<math>\beta</math> + Mut 11</td> <td>17%</td> <td>83%</td> </tr> <tr> <td>A<math>\beta</math> + mut 413</td> <td>21%</td> <td>79%</td> </tr> </tbody> </table>		Fusion	Inhibition	A $\beta$ seul	100%	0%	A $\beta$ + WT	89%	11%	A $\beta$ + Mut 11	17%	83%	A $\beta$ + mut 413	21%	79%
	Fusion	Inhibition															
A $\beta$ seul	100%	0%															
A $\beta$ + WT	89%	11%															
A $\beta$ + Mut 11	17%	83%															
A $\beta$ + mut 413	21%	79%															

Figure 13. Description schématique de la procédure multi-étapes PepDesign

Enfin deux peptides plus affins que le sauvage ont été identifiés de manière rationnelle en se basant sur l'énergie d'interaction et sur la répartition des interactions le long des peptides. Des essais de fusion de phase lipidique et de leakage ont mis en évidence que ces deux peptides modélisés et prédits comme plus affins pour leur cible réduisent le caractère fusogène de A $\beta$  29-42 de l'ordre de 80%, alors que le peptide modèle Apo E270-287 ne réduit cette fusion que de 11%. Ces résultats expérimentaux indiquent clairement que les peptides modélisés interagissent de manière plus intime avec la cible que ne le fait le partenaire naturel. Ce travail a par ailleurs débouché sur le dépôt d'un brevet en collaboration avec l'équipe du professeur Heinen du Service d'Histologie Humaine à l'Ulg (Patent PCT/EP2005/055050).

Rational design of complementary peptides to the  $\beta$ Amyloid 29–42 fusion peptide: An application of PepDesign. - Decaffmeyer M., Lins L., Charlotiaux B., VanEyck MH, Thomas A., Brasseur R. - *BBA Biomembranes* 2006 mar, 1758(3) 320-7

**Nous ne disposons pas des droits sur cette publication. Nous reprenons ici l'abstract tel que disponible sur Pubmed.**

[http://www.ncbi.nlm.nih.gov/pubmed/16313882?ordinalpos=3&itool=EntrezSystem2.PEntrez.Pubmed.Pubmed\\_ResultsPanel.Pubmed\\_RVDocSum](http://www.ncbi.nlm.nih.gov/pubmed/16313882?ordinalpos=3&itool=EntrezSystem2.PEntrez.Pubmed.Pubmed_ResultsPanel.Pubmed_RVDocSum)

**Abstract:** Peptides in solution currently exist under several conformations; an equilibrium which varies with solvent polarity. Despite or because of this structure versatility, peptides can be selective biological tools: they can adapt to a target, vary conformation with solvents and so on. These capacities are crucial for cargo carriers. One promising way of using peptides in biotechnologies is to decipher their medium-sequence-structure-function relationships and one approach is molecular modelling. Only few "in silico" methods of peptide design are described in the literature. Most are used in support of experimental screening of peptide libraries. However, the way they are made does not teach us much for future researches. In this paper, we describe an "in silico" method (PepDesign) which starts by analysing the native interaction of a peptide with a target molecule in order to define which points are important. From there, a modelling protocol for the design of 'better' peptides is set. The PepDesign procedure calculates new peptides fulfilling the hypothesis, tests the conformational space of these peptides in interaction with the target by angular dynamics and goes up to the selection of the best peptide based on the analysis of complex structure properties. Experimental biological assays are finally used to test the selected peptides, hence to validate the approach. Applications of PepDesign are wide because the procedure will remain similar irrespective of the target which can be a protein, a drug or a nucleic acid. In this paper, we describe the design of peptides which binds to the fusogenic helical form of the C-terminal domain of the Abeta peptide (Abeta29-42).

**3.1.2. Peptides complémentaires au peptide  $\beta$ -Amyloïde 29-42 (Brevet conjoint FSAGx-Ulg).**

***Heinen E. , Decaffmeyer M., Brasseur R. (2006). Complementary peptides for  $\beta$ -Amyloïde 29-42 peptide. WIPO April. Patent PCT/EP2005/055050***

Un brevet relatif à ces travaux a été déposé conjointement entre la Faculté Universitaire des Sciences Agronomiques de Gembloux (FSAGx) et l'Université de Liège (Ulg).

Il porte non pas sur les deux peptides testés expérimentalement mais bien sur les 6 peptides sélectionnés à l'issue de l'étape de criblage. Le brevet protège lesdits peptides ainsi que tous leurs dérivés ou peptidomimétiques, quelle que soit leur mode d'obtention ou leur utilisation. Les revendications du brevet sont reprises ci après.

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
13 April 2006 (13.04.2006)

PCT

(10) International Publication Number  
**WO 2006/037800 A1**

(51) International Patent Classification:

*C07K 14/775* (2006.01)    *C07K 16/18* (2006.01)  
*A61K 38/04* (2006.01)    *G01N 33/50* (2006.01)

(21) International Application Number:

PCT/EP2005/055050

(22) International Filing Date: 6 October 2005 (06.10.2005)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

04077735.1                      8 October 2004 (08.10.2004)    EP

(71) Applicants (for all designated States except US):

**UNIVERSITÉ DE LIÈGE** [BE/BE]; Interface Entreprises-Université, Quai Van Beneden, 25, B-4020 Liège (BE). **FACULTÉ UNIVERSITAIRE DES SCIENCES AGRONOMIQUES DE GEMBOUX** [BE/BE]; Passage des Déportés, 2, B-5030 Gembloux (BE).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **HEINEN, Ernst** [BE/BE]; rue du Tige, 75, B-4453 Villers St Siméon (BE). **BRASSEUR, Robert** [BE/BE]; rue Voie des Gérons, 250, B-5351 Haillot (BE). **DECAFFMEYER, Marc** [BE/BE]; rue Henrioulle, 4, B-1350 Marilles (BE).

(74) Agent: **STERNAGEL, FLEISCHER, GODEMEYER & PARTNER**; An den Gärten 7, 51491 Overath (DE).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM,

AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available):

ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:

— as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))

Published:

— with international search report  
— before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments  
— with sequence listing part of description published separately in electronic form and available upon request from the International Bureau

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: COMPLEMENTARY PEPTIDES FOR  $\beta$ -AMYLOID 29-42 PEPTIDE

(57) Abstract: The invention relates to peptides derived from the wild type ApolipoproteinE3 peptidic fragment 270-287 (ApoE WT) of sequence EDMQRQWAGLVEKVVAAV (SEQ ID NO: 2), having improved interaction with the -Amyloid peptidic fragment 29-42 (A $\beta$  29-42) GAIIGLMVGGVVIA (SEQ ID NO: 1), in a [A $\beta$  29-42/ ApoE mutant n] complex.

WO 2006/037800 A1

### Claims

1. A peptide mutated from the wild type ApolipoproteinE 3 peptidic fragment 270-287 (ApoE WT) of sequence EDMQRQWAGLVEKVQAAV (SEQ ID NO: 2) having improved interaction with the  $\beta$ -Amyloid peptidic fragment 29-42 (A $\beta$  29-42) GAIIGLMVGGVVIA (SEQ ID NO: 1), in a [A $\beta$  29-42/ ApoE mutant n] complex, said peptide being selected from the group consisting in the peptides having the following sequences:

EDMQRQLAGVVEKVQAAV (SEQ ID NO: 3)

EDMQRQLAGLVEKWQAAV (SEQ ID NO: 4)

EDMQRQLAGMWEKVQAAV (SEQ ID NO: 5)

EDVQRQLAGLVEKVQAAV (SEQ ID NO: 6)

EDMQRQMAGLMEKMQAAV (SEQ ID NO: 7)

EDMQRQVAGMWEKVQAAV (SEQ ID NO: 8).

2. An isolated peptide derived from the wild type ApolipoproteinE 3 peptidic fragment 270-287 (ApoE WT) of sequence EDMQRQWAGLVEKVQAAV (SEQ ID NO: 2), wherein the peptide is from 18 to 50 amino acid residues in length and having improved interaction with the  $\beta$ -Amyloid peptidic fragment 29-42 (A $\beta$  29-42) GAIIGLMVGGVVIA (SEQ ID NO: 1), said peptide comprising one or a combination of amino acid sequences selected from the group consisting of:

-EDMQRQLAGVVEKVQAAV (SEQ ID NO: 3)

-EDMQRQLAGLVEKWQAAV (SEQ ID NO: 4)

-EDMQRQLAGMWEKVQAAV (SEQ ID NO: 5)

-EDVQRQLAGLVEKVQAAV (SEQ ID NO: 6)

-EDMQRQMAGLMEKMQAAV (SEQ ID NO: 7) and

-EDMQRQVAGMWEKVQAAV (SEQ ID NO: 8)

-or any chemical derivative, variant, peptidomimetic, multimer thereof.

3. An isolated peptide derived from the wild type ApolipoproteinE 3 peptidic fragment 270-287 (ApoE WT) of sequence EDMQRQWAGLVEKVQAAV (SEQ ID NO: 2), wherein the peptide is from 18 to 50 amino acid residues in length and, when complexed to the  $\beta$ -Amyloid peptidic fragment 29-42 (A $\beta$  29-42) GAIIGLMVGGVVIA (SEQ ID NO: 1), decreases the fusogenic activity of A $\beta$  29-42; said peptide comprising one or a combination of amino acid sequences selected from the group consisting of:

- EDMQRQLAGVVEKVQAAV (SEQ ID NO: 3)
- EDMQRQLAGLVEKWQAAV (SEQ ID NO: 4)
- EDMQRQLAGMWEKVQAAV (SEQ ID NO: 5)
- EDVQRQLAGLVEKVQAAV (SEQ ID NO: 6)
- EDMQRQMAGLMEKMQAAV (SEQ ID NO: 7) and
- EDMQRQVAGMWEKVQAAV (SEQ ID NO: 8)
- or any chemical derivative, variant, peptidomimetic, multimer thereof.

4. The peptide of any one of claims 1 to 3, wherein said peptide comprises an amino acid sequence selected from the group consisting of:

- EDMQRQLAGVVEKVQAAV (SEQ ID NO: 3) and
- EDMQRQMAGLMEKMQAAV (SEQ ID NO: 7)
- or any chemical derivative, variant, peptidomimetic, multimer thereof

5. A recombinant protein, polypeptide or oligopeptide comprising one or a combination of amino acid sequences selected from the group consisting of:

- EDMQRQLAGVVEKVQAAV (SEQ ID NO: 3)
- EDMQRQLAGLVEKWQAAV (SEQ ID NO: 4)
- EDMQRQLAGMWEKVQAAV (SEQ ID NO: 5)
- EDVQRQLAGLVEKVQAAV (SEQ ID NO: 6)
- EDMQRQMAGLMEKMQAAV (SEQ ID NO: 7) and
- EDMQRQVAGMWEKVQAAV (SEQ ID NO: 8)
- or any chemical derivative, variant, peptidomimetic, multimer thereof.



6. A polynucleotide encoding any one of the peptide of claim 1 to 4 or encoding the recombinant protein, polypeptide or oligopeptide of claim 5.
7. A detection method of the  $\beta$ -Amyloid protein or its peptidic fragment  $\beta$ -Amyloid 29-42 (A $\beta$  29-42) GAIIGLMVGGVVIA (SEQ ID NO: 1), wherein said method comprises:
- a step of contacting a biological sample containing the  $\beta$ -Amyloid protein or its peptidic fragment  $\beta$ -Amyloid 29-42 (A $\beta$  29-42) with a peptide or recombinant protein comprising one or a combination of amino acid sequences selected from the group consisting of SEQ ID NOs 3, 4, 5, 6, 7, and 8, or any chemical derivative, variant, peptidomimetic, multimer thereof; and
  - a step of detecting the binding of said peptide or of said recombinant protein to the  $\beta$ -Amyloid protein or to its  $\beta$ -Amyloid peptidic fragment 29-42 (A $\beta$  29-42).
8. A kit for use in the detection of the  $\beta$ -Amyloid protein or its peptidic fragment  $\beta$ -Amyloid 29-42 (A $\beta$  29-42) GAIIGLMVGGVVIA (SEQ ID NO: 1), the kit comprising a peptide or recombinant protein comprising one or a combination of amino acid sequences selected from the group consisting of SEQ ID NOs 3, 4, 5, 6, 7, and 8, or any chemical derivative, variant, peptidomimetic, multimer thereof; the kit further comprising an agent for detecting the binding of said peptide or of said recombinant protein to the  $\beta$ -Amyloid protein or to its  $\beta$ -Amyloid peptidic fragment 29-42 (A $\beta$  29-42).
9. A binding assay for the detection of the  $\beta$ -Amyloid protein or its peptidic fragment  $\beta$ -Amyloid 29-42 (A $\beta$  29-42) GAIIGLMVGGVVIA (SEQ ID NO: 1), the assay involving the use of a peptide or recombinant protein comprising one or a combination of amino acid sequences selected from the group consisting of SEQ ID NOs 3, 4, 5, 6, 7, and 8, or any chemical derivative, variant, peptidomimetic, multimer thereof.

10. The binding assay of claim 9, wherein said assay is an immuno-assay, a radioimmuno-assay, an enzyme-linked immunoabsorbent assay, or a sandwich assay.
11. A medicament comprising one or more peptides or recombinant proteins according to any one of claims 1 to 5, or the polynucleotide according to claim 6.
12. The medicament according to claim 11, comprising one or more peptides or recombinant proteins which comprise one or a combination of amino acid sequences selected from the group consisting of SEQ ID NOs 3, 4, 5, 6, 7, and 8, or any chemical derivative, variant, peptidomimetic, multimer thereof.
13. Use of a peptide or a recombinant protein according to any one of claims 1 to 5, or use of the polynucleotide according to claim 6 for the manufacture of a medicament for preventive and/or therapeutic treatment of Alzheimer's disease.
14. An antibody or antibody fragment binding to an antigen specific for a peptide sequence selected from the group consisting of SEQ ID NOs 3, 4, 5, 6, 7, and 8, or any chemical derivative, variant, peptidomimetic, multimer thereof; and wherein at least four amino acids in said peptide sequence are part of a reactive portion with said antibody.
15. The antibody according to claim 14, wherein said antibody is polyclonal, monoclonal, bispecific, chimeric or antiidiotypic.
16. Use of an antibody or antibody fragment specific for a peptide sequence selected from the group consisting of SEQ ID NOs 3, 4, 5, 6, 7, and 8, or any chemical derivative, variant, peptidomimetic, multimer thereof, and wherein at least four amino acids in said peptide sequence are part of a reactive portion with said antibody; wherein said antibody is used in a binding assay for the detection of the  $\beta$ -Amyloid protein or its peptidic fragment  $\beta$ -Amyloid 29-42 (A $\beta$  29-42) GAIIGLMVGGVVIA (SEQ ID NO: 1).

17. Use of an antibody or antibody fragment specific for a peptide sequence selected from the group consisting of SEQ ID NOs 3, 4, 5, 6, 7, and 8, or any chemical derivative, variant, peptidomimetic, multimer thereof, and wherein at least four amino acids in said peptide sequence are part of a reactive portion with said antibody; wherein said antibody is used in a detection method of the  $\beta$ -Amyloid protein or its peptidic fragment  $\beta$ -Amyloid 29-42 (A $\beta$  29-42) GAIIGLMVGGVVIA (SEQ ID NO: 1); wherein said method comprises:

-a step of contacting a biological sample containing the  $\beta$ -Amyloid protein or its peptidic fragment  $\beta$ -Amyloid 29-42 (A $\beta$  29-42) with a peptide or recombinant protein comprising one or a combination of amino acid sequences selected from the group consisting of SEQ ID NOs 3, 4, 5, 6, 7, and 8, or any chemical derivative, variant, peptidomimetic, multimer thereof; and

-a step of detecting the binding of said peptide or of said recombinant protein to the  $\beta$ -Amyloid protein or to its  $\beta$ -Amyloid peptidic fragment 29-42 (A $\beta$  29-42).

18. A method of treating Alzheimer's disease comprising administering a therapeutic amount of a composition comprising a peptide sequence or a protein containing one or a combination of peptide sequences selected from the group consisting of SEQ ID NOs 3, 4, 5, 6, 7, and 8, or any chemical derivative, variant, peptidomimetic, multimer thereof.

## 3.2. Aspect prédiction de structure.

### *3.2.1. Prédiction de la structure peptidique : Où en sommes nous ?*

*Thomas A., Deshayes S., Decaffmeyer M., Van Eyck MH., Charlotiaux B., Brasseur R. (2006). Prediction of peptide structure: how far are we? Proteins 65(4):889-97.*

Si la fonctionnalité des peptides dépend évidemment de leur séquence et donc des fonctions chimiques qu'ils portent, elle est également grandement modulée par leur structure et donc par la manière dont ces fonctions sont orientées dans l'espace. A ce jour, pas mal de méthodes expérimentales de détermination de structures peptidiques existent. Au nombre de celles-ci on compte la spectroscopie IR à transformée de Fourier (FT-IR), la résonance magnétique nucléaire (RMN), le dichroïsme circulaire (CD) ou encore la diffraction des rayons X (RX). Ces méthodes sont performantes mais dans beaucoup de cas, l'analyse des résultats qu'elles fournissent pour le même peptide met en lumière des discordances. La raison probable de ces discordances est que la structure des peptides n'est que rarement unique et qu'elle dépend autant de leur capacité à se stabiliser que de la nature du milieu, aqueux ou lipidique. En effet, comparés aux protéines qui sont de plus grande taille, les peptides ont moins de possibilités de stabilisation interne, ce qui ne fait qu'accroître l'importance relative de leurs interactions avec le solvant ou avec d'autres partenaires.

Nous l'avons vu, les peptides sont un outil biotechnologique particulièrement prometteur, notamment pour des applications thérapeutiques. Malgré cela, le nombre de méthode de prédiction de la structure peptidique par modélisation reste très restreint. Jusqu'il y a peu seules deux méthodes étaient disponibles.

Premièrement la méthode Rosetta développée par le groupe de David Baker (**Simons et al, 1997**) combine modélisation par homologie et prédiction *de novo* pour les fragments sans homologie. Robetta fournit 10 modèles, dont le meilleur est le 10ème. Cette méthode a rencontré de beaux succès pour la prédiction de la structure tridimensionnelle de protéines (**Bonneau et al., 1992 ; Bradley et al., 1993**). Elle a également permis à l'équipe qui l'a développée de participer avec succès à plusieurs

éditions du CASP (Critical Assessment of techniques for protein Structure Prediction). Elle a aussi permis la construction *de novo* de protéines capables d'adopter un repliement préalablement choisi, le cas le plus connu étant celui de Top7, une protéine de 93 résidus comportant deux hélices  $\alpha$  et cinq brins  $\beta$  qui a été modélisée par l'équipe de David Baker (**Kuhlman et al., 2003**). Toutefois, si ces résultats sont bons pour les protéines, ceux pour les peptides sont médiocres.

La méthode utilise un champ de forces basé sur un potentiel de forces moyen (MFP) dérivé de la structure des protéines. C'est probablement la raison pour laquelle cette méthode donne d'excellents résultats pour la prédiction de structures protéiques et des résultats médiocres pour la prédiction de structures peptidiques, notamment parce que ce champ de forces favorise grandement le repliement des structures pour former des collapsés hydrophobes. L'équipe de développeurs de la méthode est semble-t-il consciente du fait qu'elle est plus dévolue à l'étude de la conformation des protéines que des peptides, puisque le serveur Robetta ne permet pas la prédiction sur des séquences d'une longueur inférieure à 20 résidus. Le serveur Robetta, qui se base sur cette méthode Rosetta, est disponible sur le net (<http://rosetta.bakerlab.org/>) (**Chivian et al., 2003**).

La deuxième méthode de détermination *in silico* de la structure des peptides que nous avons identifiée a été développée par le groupe indien de Raghava. Cette méthode est appelée PEPSTR et est basée dans un premier temps sur la prédiction de la structure secondaire des résidus par la méthode BETATURNS (**Kaur and Raghava, 2004**) et ensuite par l'assignement aux résidus d'angles  $\Phi\Psi$  standards, enfin sur une minimisation énergétique utilisant le champ de forces Amber 6. La méthode permet de traiter des peptides de 7 à 25 résidus et donne une structure unique. A ce jour elle n'a pas été publiée et n'a été citée qu'une fois dans la littérature. Le serveur est toutefois disponible sur internet (<http://www.imtech.res.in/raghava/pepstr/>).

L'article que nous présentons ici expose les résultats obtenus grâce à la méthode PepLook, développée conjointement par le CBMN et la société Biosiris, spin off de la Faculté. La méthode est plus précisément décrite dans le chapitre suivant, relatif aux CPPs. Elle est spécifiquement dédiée à la prédiction *in silico* de peptides et n'est pas actuellement adaptée à la prédiction de structures protéiques.

La méthode est fondée sur l'exploration de l'espace conformationnel du peptide basé d'une part sur le constat qu'il est possible de reconstruire n'importe quelle protéine de la Protein Data Bank (PDB) en utilisant un nombre restreint de couple d'angles  $\Phi\Psi$  (64 couples d'angles en tout) (Etchebest et al., 2005) et d'autre part sur une procédure itérative de type stochastique telle que celle utilisée dans de nombreux domaines scientifiques (Glick et al., 2002).

Cette méthode est modulée par une approche de type Boltzmann qui permet, à chaque itération, de faire varier la probabilité de tirage des différents couples d'angles  $\Phi/\Psi$  pour chaque position et ce, en fonction de leur contribution à l'obtention de structures de bonne ou de mauvaise énergie (figure14). La description de la méthode sera développée au point 3.2.2.

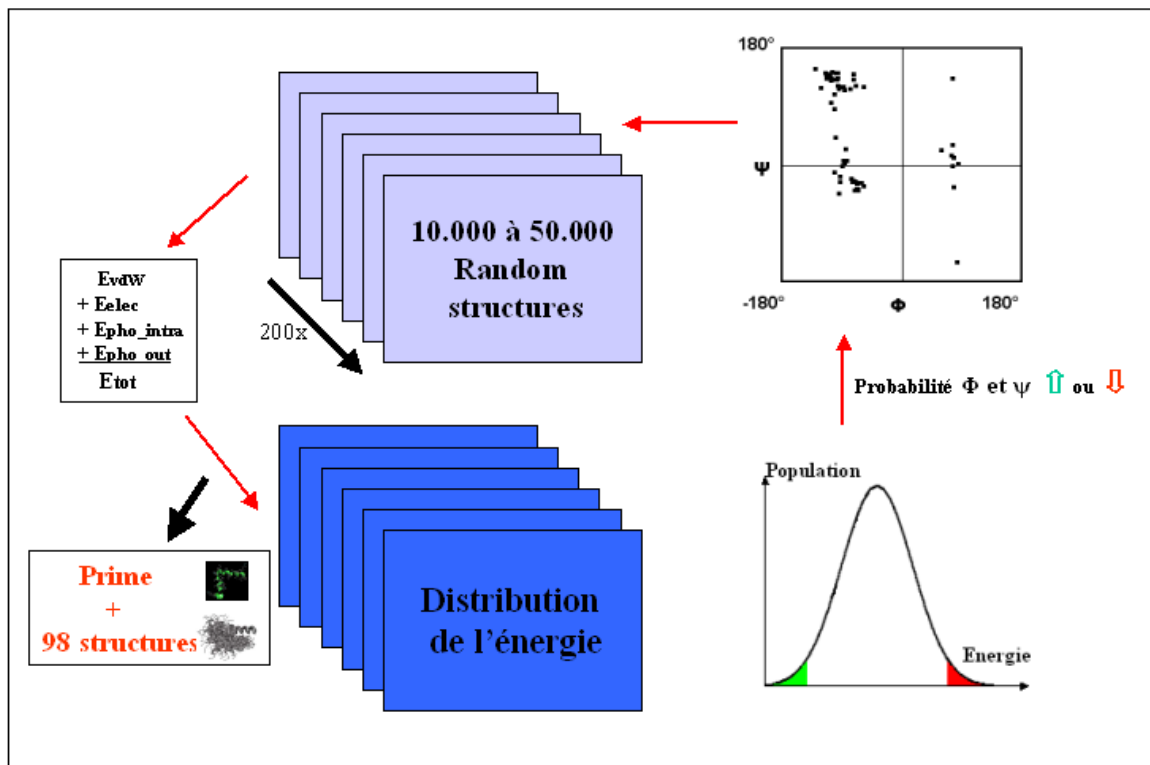


Figure 14. Schéma de principe de la méthode PepLook.

Le résultat final du calcul est d'une part, une structure, la plus stable d'un point de vue énergétique, appelée le Prime, mais aussi 98 autres structures de basse énergie. L'originalité de la méthode est qu'outre la population structurale prédite du peptide, elle donne accès à un score de stabilité et un score de polymorphisme pour le peptide. Ces dernières données nous paraissent essentielles. En effet un peptide de fonction donnée, par

exemple un peptide transporteur, doit nécessairement être capable d'adapter sa structure pour remplir plusieurs tâches différentes. Il devra être capable de lier son cargo mais aussi avoir une certaine solubilité dans un milieu hydrophile (le sang et le cytoplasme) et dans un milieu hydrophobe (la membrane). Nous pensons que la stabilité intrinsèque et le polymorphisme sont vraisemblablement des facteurs déterminants pour la fonctionnalité biologique des peptides, il est donc important d'avoir accès à cette information.

Les travaux décrits dans cet article ont porté sur la comparaison des résultats de prédiction de structure obtenus par les méthodes Rosetta, Pepstr et PepLook pour des peptides de types différents : la magainin 2 qui est un CPP hélicoïdal, le Transportan qui est un CPP de structure mixte hélice/désordonnée (random coil), HA2 qui est le peptide de l'hémagglutinine influenza et qui a une structure désordonnée et hCT9-32 qui est un fragment de la calcitonine humaine. Dans l'article nous montrons que les structures prédites par la méthode PepLook pour les trois premiers peptides correspondent très bien aux structures déterminées expérimentalement en RMN, bien mieux que ne le font les méthodes Pepstr et Rosetta. Pour hCT9-32, l'analyse des résultats PepLook suggère que le peptide est polymorphe, ce qui est en corrélation avec le fait que les résultats expérimentaux sont controversés. Le même type d'analyse suggère que la magainine 2 est monomorphe alors que le transportan est désordonné uniquement dans sa partie N terminale et HA2 dans sa partie C terminale. D'autre part, les scores de stabilité basés sur le calcul des MFP pour les structures calculées de ces différents peptides, sauf HA2, sont du même ordre que ceux des structures obtenues en RMN.

Prediction of Peptide Structure: how far are we? – Thomas A., Deshayes S., Decaffmeyer M., VanEyck MH., Charlotiaux B. and Brasseur R. – *Proteins*. 2006 Dec 1;65(4):889-97.

**Nous ne disposons pas des droits sur cette publication. Nous reprenons ici l'abstract tel que disponible sur Pubmed.**

[http://www.ncbi.nlm.nih.gov/pubmed/17019719?ordinalpos=2&itool=EntrezSystem2.PEntrez.Pubmed.Pubmed\\_ResultsPanel.Pubmed\\_RVDocSum](http://www.ncbi.nlm.nih.gov/pubmed/17019719?ordinalpos=2&itool=EntrezSystem2.PEntrez.Pubmed.Pubmed_ResultsPanel.Pubmed_RVDocSum)

**Abstract:** Rational design of peptides is a challenge, which would benefit from a better knowledge of the rules of sequence-structure-function relationships. Peptide structures can be approached by spectroscopy and NMR techniques but data from these approaches too frequently diverge. Structures can also be calculated in silico from primary sequence information using three algorithms: Pepstr, Robetta, and PepLook. The most recent algorithm, PepLook introduces indexes for evaluating structural polymorphism and stability. For peptides with converging experimental data, calculated structures from PepLook and, to a lesser extent from Pepstr, are close to NMR models. The PepLook index for polymorphism is low and the index for stability points out possible binding sites. For peptides with divergent experimental data, calculated and NMR structures can be similar or, can be different. These differences are apparently due to polymorphism and to different conditions of structure assays and calculations. The PepLook index for polymorphism maps the fragments encoding disorder. This should provide new means for the rational design of peptides.



### ***3.2.2. Le polymorphisme structural, un paramètre nécessaire pour la pénétration cellulaire des peptides : Hypothèse.***

#### **A) Introduction.**

Les CPPs constituent une autre application thérapeutique des peptides. Les membranes plasmiques constituent une barrière naturelle qui sépare le compartiment interne de la cellule de l'extérieur. Ces membranes interdisent le passage de la plupart des composés organiques et sont donc un obstacle à l'entrée des molécules thérapeutiques telles que les peptides, les protéines, les drogues ou encore les acides nucléiques dans les cellules.

De nombreuses méthodes utilisant des systèmes de transport tels que les hydrogels, les liposomes ou encore les nanoparticules ont été développées pour permettre la traversée membranaire de molécules d'intérêt tout en les protégeant de la dégradation enzymatique (**Morishita and Peppas, 2006**). Parmi ces systèmes, les CPPs dont le premier a été identifié en 1994 (**Derossi et al., 1994**), constituent un choix d'avenir (**Mae and Langel, 2006**). Deux stratégies utilisant les CPP sont utilisées.

La première stratégie fait appel à la formation d'un lien covalent entre le transporteur et sa cargaison. Cette technologie utilise des peptides tels que TAT48-60 qui est un fragment du facteur d'activation de transcription de HIV-1 (**Fawell et al., 1994**), la Pénétratine, un peptide de 16 résidus issu de l'homéodomaine d'Antennapedia ou encore le Transportan qui est un peptide chimérique de 27 résidus construit à partir de la partie N-terminale du neuropeptide Galanine et du Mastoparan de venin de guêpe, liés par une lysine (**Pooga et al., 1998**).

La seconde stratégie est de type non covalent. Dans celle-ci le complexe formé entre les peptides vecteurs et la cargaison est stabilisé par des interactions intermoléculaires. Par exemple le peptide Pep-1 (**Morris et al., 2001**), le peptide MPG (**Morris et al., 1997**) et le peptide KALA (**Wyman et al., 1997**) sont identifiés comme étant efficaces pour cette stratégie.

Aucune certitude n'existe à ce jour quant au(x) mécanisme(s) de pénétration des CPPs. Toutefois plusieurs mécanismes ont été proposés, parmi lesquels on retrouve le modèle des micelles inverses (**Derossi et al., 1996**), l'électroporation locale (**Binder et**

al., 2003), le passage par une structure transitoire de pores (Deshayes et al., 2006), la fusion de membrane (Henriques et al., 2004), un modèle impliquant l'endocytose (Jones, 2007; Kaplan et al., 2005; Richard et al., 2005) ou encore une combinaison de ces différents mécanismes (Letoha et al., 2003).

En revanche, il est évident que les CPPs présentent des similarités. En effet, ils sont généralement amphipatiques permettant l'interaction tant avec la membrane qu'avec la cargaison, d'une longueur généralement limitée à 30 acides aminés et disposent d'une charge nette positive, importante pour l'interaction avec les têtes polaires des phospholipides, la solubilité du complexe formé avec la cargaison et l'interaction avec des cargaisons chargées telles que les acides nucléiques. Plus récemment il a été mis en évidence que la conformation du CPP en fonction de l'environnement est aussi un critère important pour son internalisation (Chaloin et al., 1997).

L'hypothèse que nous avons émise en prélude à ce travail est, qu'outre les propriétés communes des CPPs décrites plus hauts, un peptide vecteur doit être capable d'adopter plusieurs conformations pour pouvoir pénétrer une membrane. En d'autres termes qu'il doit être polymorphe. En effet, on peut aisément imaginer qu'un même peptide ne peut présenter une solubilité acceptable dans des milieux aussi différents que l'eau ou la membrane que dans la mesure où il est capable de moduler son hydrophobicité apparente en faisant varier sa structure. Nous avons vérifié cette hypothèse sur deux peptides vecteurs (le Transportan et la Pénétratine) ainsi que sur des variants fonctionnels et non fonctionnels.

## **B) Méthode.**

### ***Méthode Boltzmann- stochastique.***

L'objectif de la méthode est d'obtenir une prédiction sur la structure d'un peptide à partir de sa seule séquence. On imagine aisément qu'il est illusoire de vouloir générer et calculer l'énergie de toutes les conformations théoriques que serait susceptible d'adopter un peptide. En effet, même si on limite le champs d'investigation aux seuls angles  $\Phi\Psi$ , et qu'on ne l'explore que tous les  $10^\circ$ , on aurait  $36 \times 36$ , soit 1296 possibilités pour chaque résidus, ce qui reviendrait à  $1296^{20}$ , soit  $1,8 \times 10^{62}$  possibilités de

conformation pour un peptide de 20 résidus. Ce calcul est impossible dans un délai raisonnable avec les moyens informatiques dont on dispose actuellement.

Dans notre méthode nous avons donc cherché à limiter le nombre de possibilités structurales, c'est à dire le nombre de couples d'angles  $\Phi\Psi$ , afin de rendre le calcul réalisable tout en espérant un résultat cohérent et satisfaisant. Nous avons donc décidé de partir d'un set restreint de couple d'angles  $\Phi\Psi$  dérivés du travail de l'équipe d'Etchebest (Etchebest et al., 2005). Ce groupe a proposé un alphabet structural composé de 16 pentapeptides peptides à partir duquel il est possible de reconstruire la structure de n'importe quelle protéine de la pdb. Les 64 couples d'angles  $\Phi\Psi$  extraits de ces pentapeptides, que nous avons utilisés pour notre travail, sont repris à la figure 15.

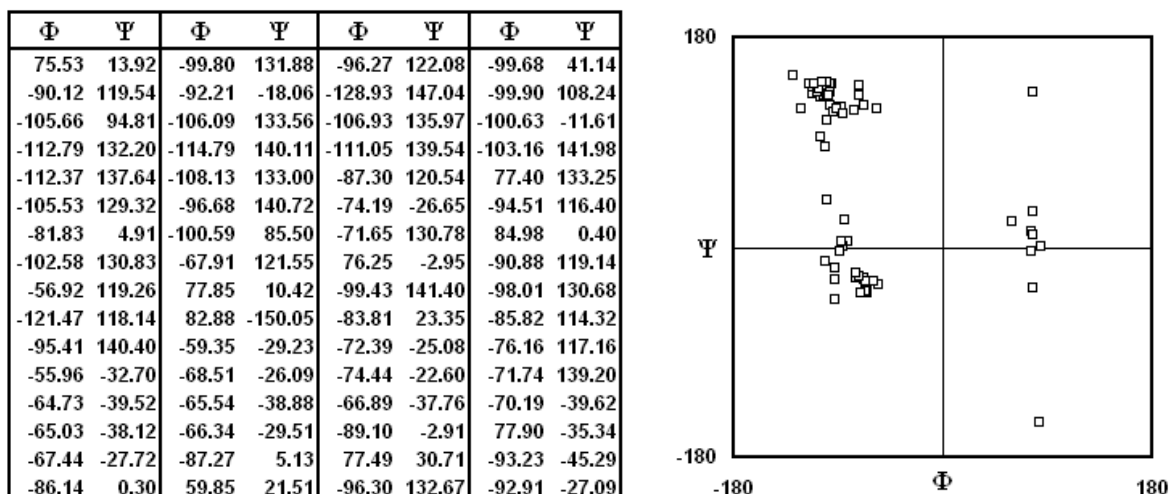


Figure 15. Angles  $\Phi\Psi$  utilisés pour la prédiction de structure.

Par rapport à l'exemple pris plus haut, le nombre de conformations possibles en employant ce set de couples d'angles  $\Phi\Psi$  est réduit à  $64^{20}$  soit  $1,3 \times 10^{36}$ , ce qui est plus raisonnable mais reste toutefois inaccessible avec les puissances de calcul disponibles à ce jour. Il a donc fallu avoir recours à une méthode permettant de réduire encore le nombre de conformations calculées. Nous nous sommes tournés vers une approche de type Boltzmann-stochastique.

Dans cette méthode itérative, la première étape consiste à générer 50000 structures de manière aléatoire à l'aide du set de 64 couples d'angles. Pour chacune de ces structures l'énergie totale est calculée comme la somme de l'énergie de van der Waals

( $E_{vdW}$ ), de l'énergie électrostatique( $E_{elec}$ ), de l'énergie hydrophobe intramoléculeire ( $E_{pho\_intra}$ ) et de l'énergie hydrophobe intermoléculeire ( $E_{pho\_out}$ ) qui simule l'interaction avec le solvant (**équation 7**).

$$(Eq.7) \quad E_{tot} = E_{vdW} + E_{elec} + E_{pho\_intra} + E_{pho\_out}$$

La contribution de van der Waals est calculée à l'aide de la fonction de Lennard-Jones qui décrit l'énergie d'interaction entre deux atomes non-liés (**équation 4,§1.3.2**).

L'équation de Coulomb est utilisée pour le calcul de l'énergie d'interaction électrostatique entre atomes chargés non liés (**équation 5,§1.3.2**).

L'énergie hydrophobe intramoléculeire est calculée suivant une fonction qui simule la décroissance exponentielle en fonction de la distance entre les atomes en interaction (**équation 6,§1.3.2**).

Le champs de forces permet également de calculer l'énergie d'interaction de la structure avec le solvant (**équation 8**), permettant de réaliser des prédictions tant en milieu hydrophobe qu'hydrophile

$$(Eq.8) \quad E_{pho\_out} = \sum_{i=1}^N S_i E_{tr_{Si}}$$

où  $S$  est la surface atomique accessible au solvant calculée par la méthode de Shrake and Rupley avec une précision de 162 points.  $E_{tr_{Si}}$  est l'énergie de transfert de l'atome  $i$  d'une phase hydrophobe vers une phase hydrophile par unité de surface.

Dans une deuxième étape, pour chaque résidu, on classe les 50000 couples d'angles  $\Phi\Psi$  en fonction de l'énergie des structures qu'ils ont engendrées. Lorsqu'on porte le nombre de couples d'angles  $\Phi\Psi$  correspondant à une zone d'énergie on obtient une courbe de type gaussien (**figure 16**).

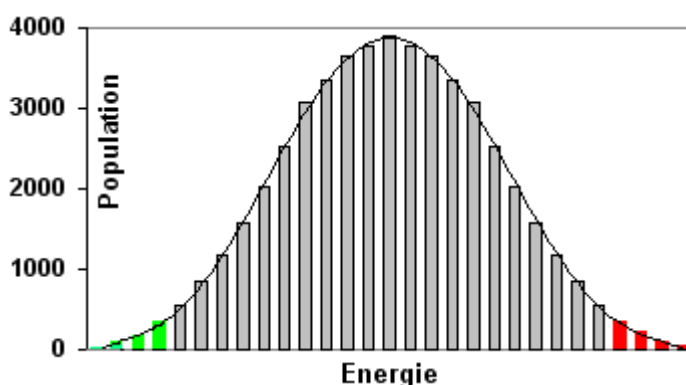


Figure 16. Distribution de l'énergie des structures calculées.

Dans cette gaussienne un grand nombre de couples conduit à des structures d'énergie moyenne (en gris) alors qu'un petit nombre conduit à des structures d'énergie basse, donc intéressantes (en vert), et qu'un autre petit nombre conduisent à des structures d'énergie haute, donc peu intéressantes (en rouge). Pour les itérations suivantes, on va accroître la probabilité de tirage des couples  $\Phi\Psi$  qui apparaissent dans la zone des énergies intéressantes sans apparaître dans la zone des énergies défavorables. Les couples  $\Phi\Psi$  qui au contraire apparaissent dans la zone des énergies défavorables sans apparaître dans la zone des énergies intéressantes auront une probabilité de tirage réduite.

De cette manière, le choix de tirage aléatoire des angles  $\Phi\Psi$  pour chaque résidu se réduit d'itération en itération, conduisant finalement à la génération de structures d'énergies de plus en plus favorables d'un point de vue énergétique. Classiquement, un nombre d'itération équivalent à 10 fois le nombre de résidus du peptide suffit à donner des résultats reproductibles.

Au final, la procédure donne 99 structures de basse énergie dont la meilleure que nous appelons le "PRIME".

### ***Score de stabilité.***

Le score de stabilité est donné en comparant le MFP (Potentiel de Forces moyen) de chaque résidu de la structure du peptide avec la valeur de MFP de référence extraite d'une banque non redondante de 500 structures protéiques repliées (**Lowell et al., 2003**). Sur l'ensemble de ces 500 structures, nous avons calculé les fréquences des différents couples atomiques en fonction de la distance pour différents types atomiques (O, N, Csp2, Csp3,...). L'équation de Boltzman est ensuite utilisée pour lier ces fréquences à une énergie qui décrit la probabilité d'existence d'un état donné par rapport à l'ensemble des états possibles (**équation 9**).

$$\text{(Eq.9) } P_m = [ \exp (-E_m / kT) ] / Z$$

où  $P_m$  est la probabilité d'existence de l'état  $m$ ,  $k$  est la constante de Boltzmann,  $T$  est la température absolue,  $E$  l'énergie associée à l'état  $m$  et  $Z$  la somme des états ou somme de Boltzmann ( $Z = \sum_m \exp(-E_m/kT)$ )

Si le nombre de données est suffisant on peut lier la probabilité ( $P$ ) à une fréquence ( $F$ ). On peut donc calculer des distributions de fréquence en fonction de la distance pour chaque couple d'atomes et en déduire des distributions d'énergie en fonction de la distance (**Figure 17, haut**).

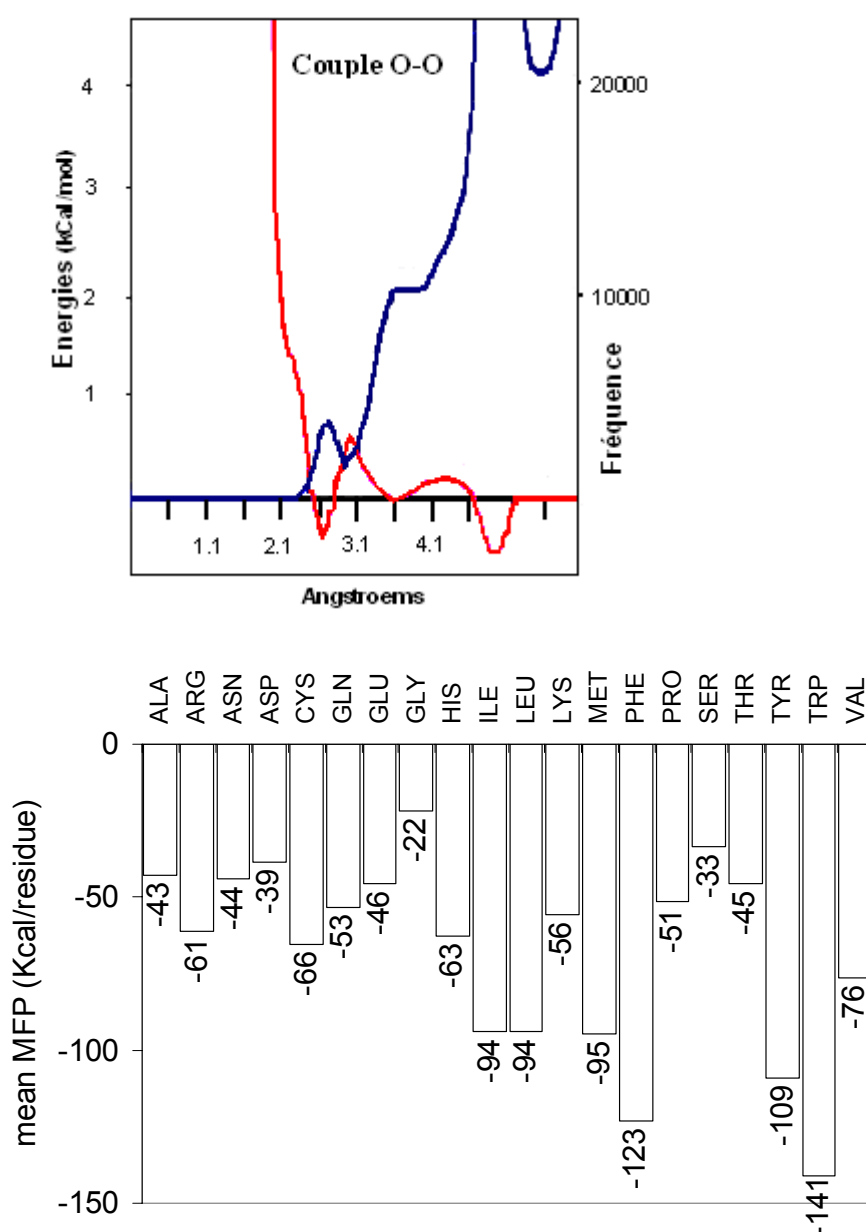


Figure 17. Transformation des fréquences en fonction de la distance en énergie pour le couple O-O (haut). Valeurs de référence des MFP des 20 résidus (Bas).

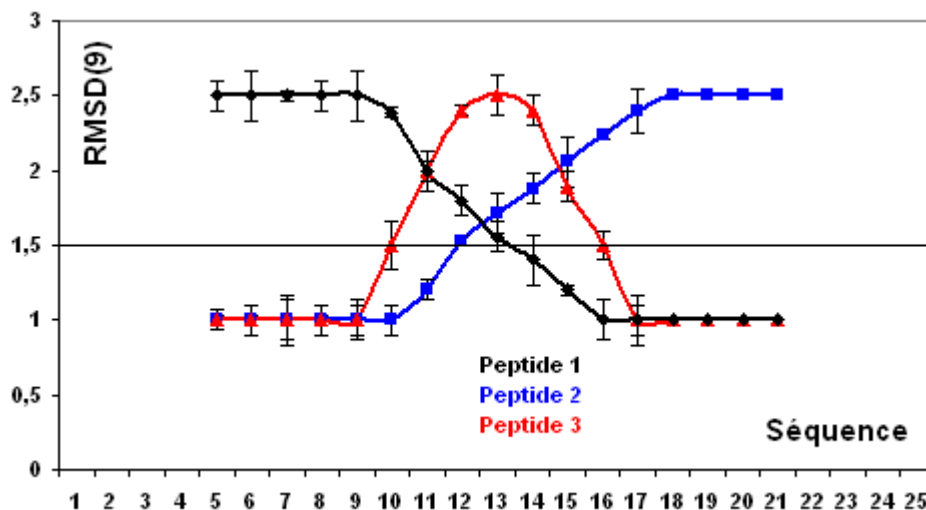
Finalement nous avons calculé une valeur MFP de référence pour chaque type de résidu dans un ensemble de protéines de structure stable en considérant que le MFP d'un acide aminé est la somme des MFPs de ses atomes constitutifs (**figure 17,bas**).

Le score de stabilité est obtenu en comparant la valeur de MFP des résidus d'une structure à la valeur de référence des mêmes résidus. Ce score permet donc d'identifier dans une structure, les résidus en carence de stabilité et donc les résidus susceptibles de l'augmenter en cherchant à lier des partenaires externes.

### *Score de polymorphisme.*

Une originalité de la méthode PepLook est qu'elle ne fournit pas une structure unique mais bien une série de structures dont la meilleure d'un point de vue énergétique est le Prime. Ceci permet de comparer les structures entre elles et surtout de les classer en familles structurales si le peptide est polymorphe.

Pour ce faire le RMSd[9], c'est à dire la déviation de RMS de chaque résidus par rapport à la structure de référence (le Prime) est calculée sur une fenêtre de 9 résidus. La valeur est assignée à l'acide aminé central de la fenêtre. Ceci permet d'avoir une visualisation rapide des zones monomorphes et des zones présentant une certaine labilité structurale (**figure 18**).



**Figure 18. Score de polymorphisme le long de la séquence.**

Dans l'exemple de la figure 18, la variété structurale pour le peptide 1 apparaît en N terminal (RMSd[9] élevé), ce qui signifie que l'on peut aligner localement les 98

structures sur celle du Prime dans la partie C terminale du peptide, ce qui n'est pas le cas pour la partie N terminale. Cette diversité structurale est localisée en C terminal pour le peptide 2 et au centre pour le peptide 3.

***Séquence des peptides étudiés.***

Pénétratine: **RQIKIWFQ<sup>50</sup>NRRMKWKK-NH<sub>2</sub>**  
 Pénétratine (48) **RQIKIFFQ<sup>50</sup>NRRMKWKK-NH<sub>2</sub>**  
 Pénétratine (48-56) **RQIKIFFQ<sup>50</sup>NRRMKFKK-NH<sub>2</sub>**  
 Transportan: **GWTLNSAGYL<sup>10</sup>LGKINLKALA<sup>20</sup>ALAKKIL-NH<sub>2</sub>**  
 TP07: **LNSAGYLLGK<sup>10</sup>INLKALAALA<sup>20</sup>KKIL-NH<sub>2</sub>**  
 TP08: **LLGKINLKAL<sup>10</sup>AALAKKIL-NH<sub>2</sub>**  
 TP10: **AGYLLGKINL<sup>10</sup>KALAALAKKI<sup>20</sup>L-NH<sub>2</sub>**  
 TP13: **LNSAGYLLGK<sup>10</sup>ALAALAKKIL<sup>20</sup>-NH<sub>2</sub>**  
 TP15: **LNSAGYLLGK<sup>10</sup>LKALAALAK-NH<sub>2</sub>**

**C) Résultats et discussion.**

Afin de vérifier notre hypothèse nous avons calculé les structures de deux peptides vecteurs et de leurs variants à l'aide de PepLook. La Pénétratine et des mutants dans lesquels un ou plusieurs tryptophanes ont été remplacés par des phénylalanines : Pénétratine 48 (W48F) et Pénétratine 48-56 (W48F et W56F). Ces simples mutations abolissent partiellement ou totalement la capacité du peptide à rentrer dans les cellules (Derossi et al., 1996 ; Lindberg et al., 2003) (figure 19). Deuxièmement les calculs ont porté sur le Transportan et 5 variants par délétion dont deux sont internalisés (TP07 et TP10) et deux ne le sont pas ou beaucoup moins que le peptide sauvage (TP08, TP13 et TP15) (Soomets et al., 2000) (figure 19).

Peptide	Internalisation
<b>Penetratin</b>	+++
Penetratin (48)	+
Penetratin (48-56)	-
<b>Transportan</b>	+++
TP07	+++
TP08	- / +
TP10	+++
TP13	- / +
TP15	- / +

**Figure 19. Capacité d'internalisation des peptides étudiés.**

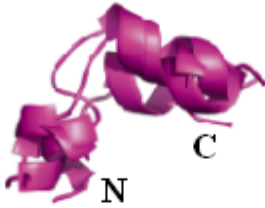


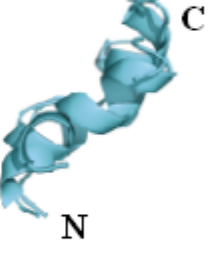
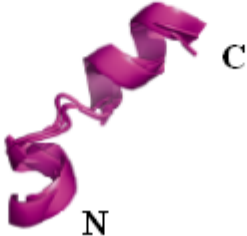
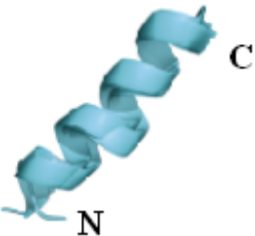


Pour chacun de ces peptides PepLook a généré 98 structures d'énergie basse plus un Prime. Les calculs ont été réalisés en milieu hydrophobe et en milieu hydrophile et ont été répétés trois fois.

### C1) Pénétratine et mutants.

#### *C.1.1. Analyse de la structure des Primes*

Les structures calculées des Primes de la Pénétratine et de ses mutants en milieu hydrophobe et hydrophile pour les 3 essais ont été alignées (**figure 20**).

	Milieu Hydrophobe	Milieu hydrophile
Pénétratine		
Pénétratine 48		
Pénétratine 48-56		

**Figure 20.** Structure des Primes pour la Pénétratine et ses mutants en milieu hydrophobe et hydrophile (alignement des structures résultant de 3 calculs).

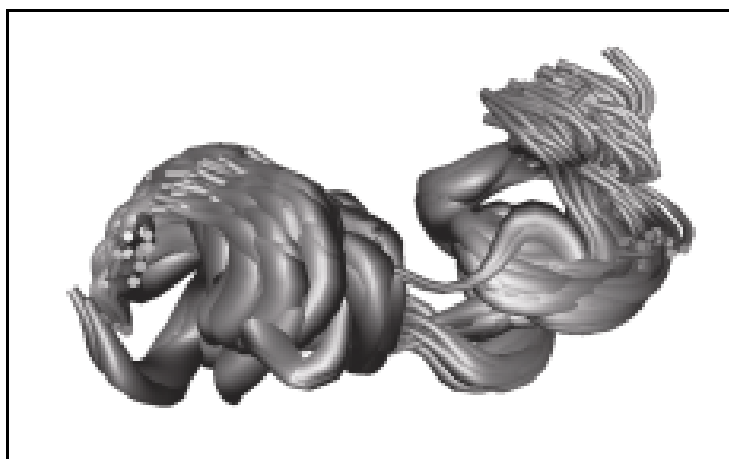
L'analyse de la structure secondaires des Primes générés par PepLook pour chaque peptide indique que, quel que soit le milieu, la Pénétratine présente un taux d'hélicité de 30% essentiellement localisé dans la partie Cterminale. Dans le milieu

hydrophobe, le taux d'hélicité est de 31% pour la Pénétratine 48 et de 44% pour la Pénétratine 48-56. Ce taux monte respectivement à 55% et 75% en milieu hydrophile. Ces résultats montrent donc que la contribution hélicoïdale est plus prononcée pour les mutants que pour le peptide sauvage ce qui est en accord avec les résultats de la littérature (Letoha et al., 2003).

### *C.1.2. Analyse du polymorphisme.*

#### *Pénétratine sauvage.*

Pour cette partie du travail nous avons utilisé les 98 structures générées par PepLook (**Figure 21**) que nous avons comparées à la structure du Prime. Nous avons calculé deux types de RMSd, un RMSd global (gRMSd) afin d'identifier et de faire un regroupement des populations structurales ainsi qu'un RMSd local (RMSd[9]) moyen calculé pour les différentes populations pour mettre à jour les divergences structurales locales.

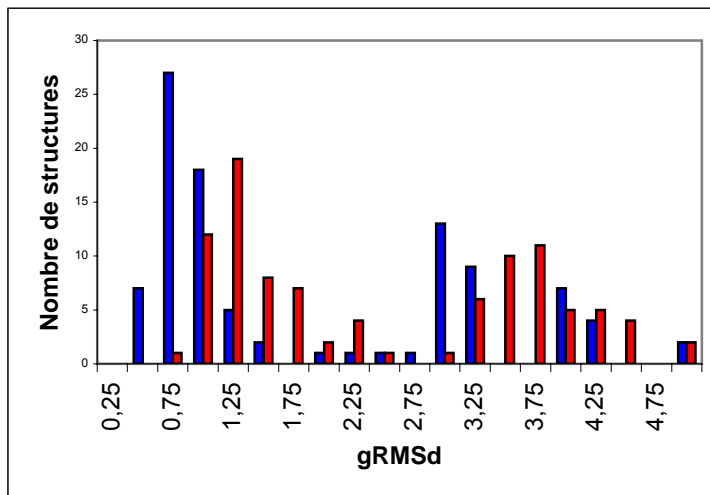


**Figure 21. Alignement structural des 98 structures de la Pénétratine générée par PepLook en milieu hydrophile.**

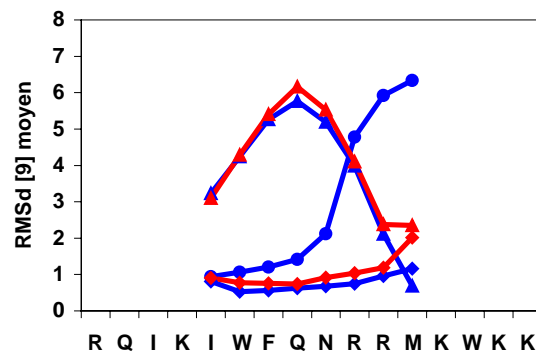
L'analyse du gRMSd met en lumière que pour la Pénétratine, plusieurs populations coexistent, tant en milieu hydrophobe avec 2 populations comprenant respectivement 56% (incluant le Prime) et 44% des structures dont la déviation du RMS global (gRMSd) est centrée sur 3,6Å, qu'en milieu hydrophile avec 3 populations

comprenant respectivement 64% (incluant le Prime), 23% et 13% des structures dont la la déviation du RMS global par rapport au Prime (gRMSd) est centrée sur respectivement 2,9Å et 4,0 Å (**figure22**).

Le RMSd local (RMSd[9]) moyen en milieu hydrophile nous indique premièrement que la première population incluant le Prime montre un RMSd[9] moyen inférieur à 1,5. Ensuite il apparaît que la deuxième population diverge structurellement du Prime dans la partie C terminale et que la troisième diverge du Prime dans la partie centrale. En milieu hydrophobe, la 2ème population diverge structurellement du Prime dans le centre du peptide, les parties Nterminale et Cterminale étant superposables à celle du Prime (**figure23**).

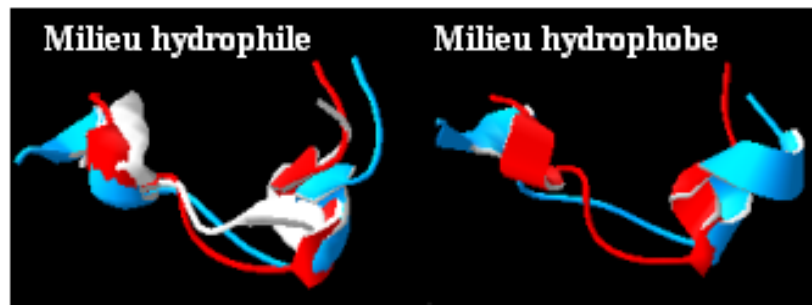


**Figure 22.** Distribution des structures calculées pour la Pénétratine en fonction de leur gRMSd par rapport au Prime en milieu hydrophile (bleu) et en milieu hydrophobe (rouge).



**Figure 23.** RMSd[9] des structures représentatives des différentes populations en milieu hydrophile (bleu) et en milieu hydrophobe (rouge) pour la Pénétratine.

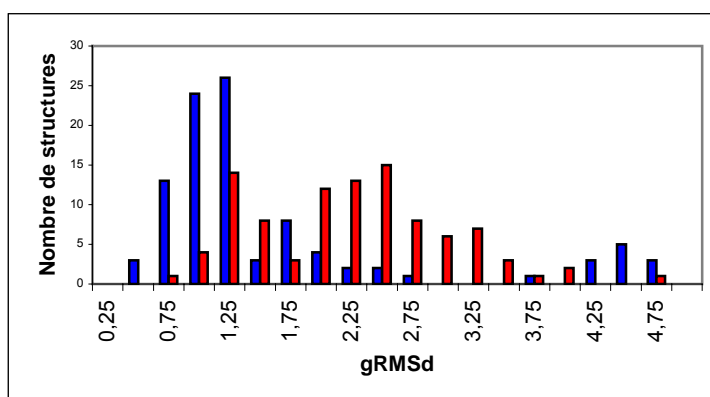
Ceci indique clairement que plusieurs populations structurales existent aussi bien en milieu hydrophobe qu'hydrophile. On peut donc parler de polymorphisme pour la Pénétratine (**figure 24**).



**Figure 24. Structure des représentants des familles structurales de la Pénétratine en milieu hydrophile et hydrophobe (Prime en rouge, 2<sup>ème</sup> population en bleu, troisième en blanc).**

#### *Pénétratine 48.*

La même analyse du polymorphisme a été réalisée sur ce peptide mutant. Elle indique que 3 populations existent en milieu hydrophobe avec respectivement 29% (incluant le prime), 51% et 20% des structures dont les gRMSd sont centrés sur 2,2Å et 3,2 Å, alors qu'en milieu hydrophile on distingue d'une part la population du Prime avec 88% et une seconde population de 12% dont le gRMSd est centré sur 4,3Å (**figure 25**).



**Figure 25. Distribution des structures calculées pour la Pénétratine 48 en fonction de leur gRMSd par rapport au Prime en milieu hydrophile (bleu) et en milieu hydrophobe (rouge).**

L'analyse du RMSd[9] montre que la structure en milieu hydrophobe peut diverger par rapport à celle du Prime tout au long de la séquence, mais de manière très ténue et qu'en milieu hydrophile la divergence est marquée en C terminal (**figure 26**). La Pénétratine 48 montre donc une tendance au polymorphisme moins marquée que celle de la Pénétratine sauvage. Une vue des structures représentant les familles structurales dans les deux milieux est donnée à la figure 27.

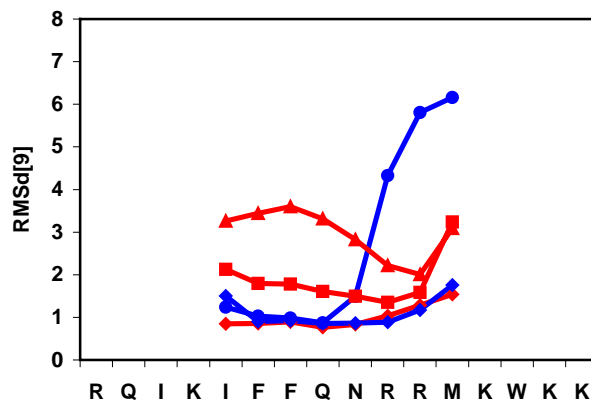


Figure 26. RMSd[9] des structures représentatives des différentes populations en milieu hydrophile (bleu, populations 1 et 2 de bas en haut) et en milieu hydrophobe (rouge, populations 1, 2 et 3 de bas en haut) pour la Pénétratine 48.

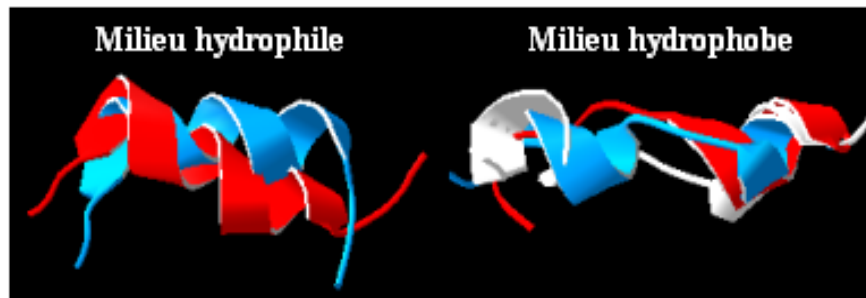


Figure 27. Structure des représentants des familles structurales de la Pénétratine 48 en milieu hydrophile et hydrophobe (Prime en rouge, 2<sup>ème</sup> population en bleu, troisième en blanc).

#### *Pénétratine 48-56.*

Le même type d'analyse mené sur la Pénétratine 46-58 montre qu'en milieu hydrophobe trois populations regroupant respectivement 64% (incluant le Prime), 23% et 13% dont les gRMSd sont centrés sur 2,4Å et 4,1 Å coexistent mais qu'en milieu

hydrophile nous n'avons qu'une population unique correspondant à la structure du Prime (**figure 28**). L'unicité de la population en milieu hydrophile est confirmée par l'analyse du RMSd[9] alors qu'en milieu hydrophobe les populations divergent d'un point de vue structural dans la partie C-terminale et au centre du peptide (**figure 29**). Une vue des structures représentant les familles structurales dans les deux milieux est donnée à la figure 30.

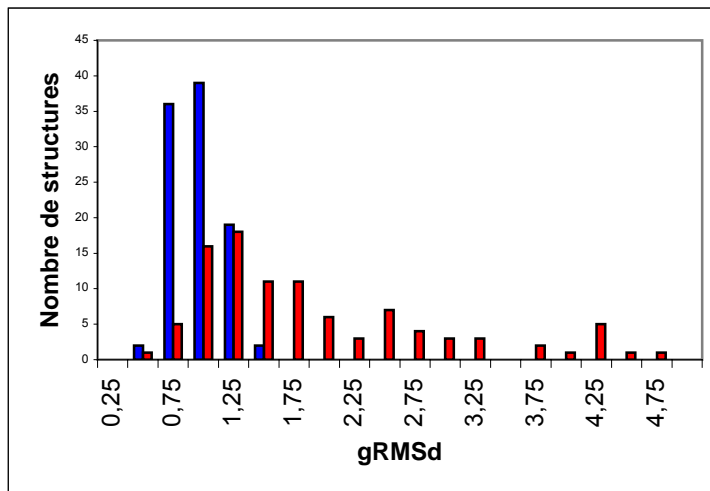


Figure 28. Distribution des structures calculées pour la Pénétratine 48-56 en fonction de leur gRMSd par rapport au Prime en milieu hydrophile (bleu) et en milieu hydrophobe (rouge).

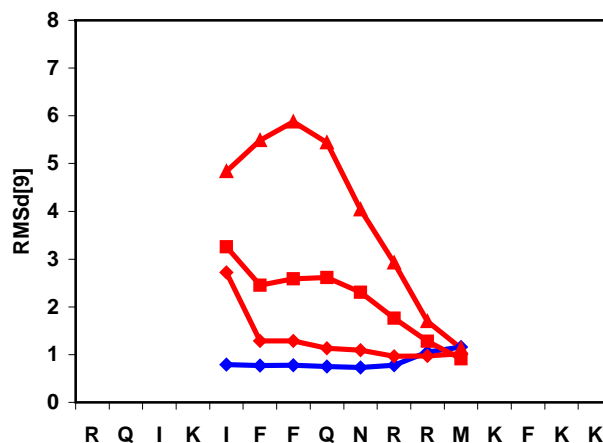
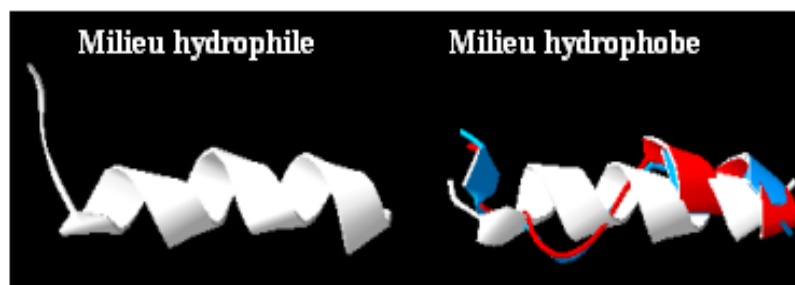


Figure 29. RMSd[9] des structures représentatives des différentes populations en milieu hydrophile (bleu) et en milieu hydrophobe (rouge, populations 1, 2 et 3 de bas en haut) pour la Pénétratine 48-56.



**Figure 30. Structure des représentants des familles structurales de la Pénétratine 48 en milieu hydrophile et hydrophobe (Prime en rouge, 2<sup>ème</sup> population en bleu, troisième en blanc).**

En conclusion, on voit clairement qu'il est possible de corréler les résultats obtenus par PepLook avec les données fonctionnelles. En effet, la Pénétratine sauvage, qui a une excellente capacité à pénétrer les membranes (**figure 19**), est prédite comme polymorphe dans les deux milieux. Le gRMSd moyen calculé sur les 98 structures est de respectivement 1,8Å et 2,4Å en milieu hydrophile et en milieu hydrophobe.


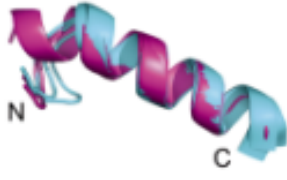
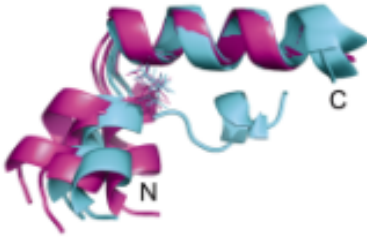
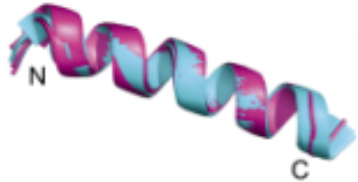

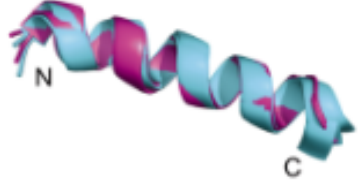
La Pénétratine 48 a une structure plus ordonnée en hélice principalement en milieu hydrophile. D'autre part, le gRMSd moyen calculé sur les 98 structures tombe respectivement à 1,5Å et 2,1Å en milieu hydrophile et en milieu hydrophobe. Ce peptide est donc partiellement polymorphe, mais moins que le peptide sauvage. Il conserve toutefois une capacité réduite à pénétrer les membranes (**figure 19**).

Enfin, la pénétratine 48-56, perd complètement son polymorphisme en milieu hydrophile avec un gRMD moyen de 0,8Å, tandis qu'elle est sensiblement moins polymorphe que le peptide sauvage en milieu hydrophobe (gRMs moyen =1,8). Ce peptide ne présente pas de capacité à traverser les membranes (**figure 19**).

## **C2) Transportan et variants.**

### ***C.2.1. Analyse de la structure des Primes***

Les calculs de structure ont été répétés trois fois. Les structures calculées des Primes du Transportan et de ses variants en milieu hydrophobe et hydrophile ont été alignées (**figure 31**).

Peptides fonctionnels	Peptides non fonctionnels
<p>Transportan</p> 	<p>TP08</p> 
<p>TP07</p> 	<p>TP13</p> 
<p>TP10</p> 	<p>TP15</p> 

**figure 31. Structure des primes pour le Transportan et ses variants en milieu hydrophobe (rose) et hydrophile (bleu) (alignement des structures résultant de 3 calculs).**

L'analyse de la structure des Primes du Transportan indique que le peptide est hélicoïdal à 50% principalement dans sa partie C terminale (L16 à L27), alors que la structure est plus variable dans la partie N terminale avec une amorce d'hélice entre S6 et G8. Les variants fonctionnels TP07 et TP10 adoptent le même type de structure, hélicoïdal en C terminal et variable en Nter, que le Transportan. Par contre les variants négatifs, sont à 70% structurés en hélice  $\alpha$  pour TP08 et à 80% pour TP13 et TP15 (**figure 31**). Il n'y a pas de données expérimentales disponibles pour la structure des variants mais les données de RMN pour le Transportan sont tout à fait en accord avec les structures calculées (**Lindberg et al., 2001**), à savoir : Une partie C-terminale stable et hélicoïdale et une partie N-terminale plus variable, liées par un domaine courbe autour de l'asparagine 15.



### C.2.2. Analyse du polymorphisme.

La même stratégie d'analyse que pour la Pénétratine et ses mutants a été appliquée.

#### *Transportan.*

L'analyse du gRMSd met en lumière que pour le Transportan, plusieurs populations coexistent. En milieu hydrophobe, 4 populations apparaissent : la première constituée du Prime seul, les suivantes comprenant respectivement 56%, 25% et 18% des structures dont le gRMSd est centré sur 4Å, 5,5Å et 7,4Å. En milieu hydrophile, 3 populations comprenant respectivement 56% (incluant le Prime), 26% et 18% des structures dont le gRMSd est centré sur 5,2Å et 7,8Å (**figures 32 et 34**).

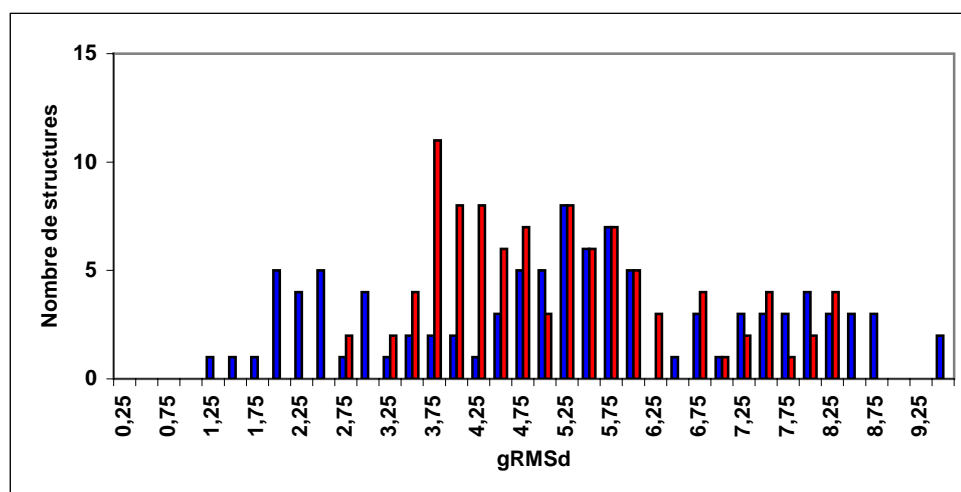


Figure 32. Distribution des structures calculées pour le Transportan en fonction de leur gRMSd par rapport au Prime en milieu hydrophile (bleu) et en milieu hydrophobe (rouge).

Le RMSd local (RMSd[9]) moyen en milieu hydrophile nous indique que la première population (incluant le Prime) diverge légèrement de la structure du Prime dans la partie N-terminale. Les deuxième et troisième populations divergent plus nettement dans cette zone. En milieu hydrophobe, la première population présente un RMSd[9] nul puisqu'elle est constituée uniquement du Prime, tandis que les trois autres populations divergent très nettement dans la partie N-terminale (**figure 33**). Ces résultats indiquent que la structure est très bien conservée dans la partie C-terminale du peptide et plus

variable dans la partie N-terminale. On peut donc dire que le Transportan est un peptide polymorphe et que ce polymorphisme est principalement localisé dans sa partie N-terminale.

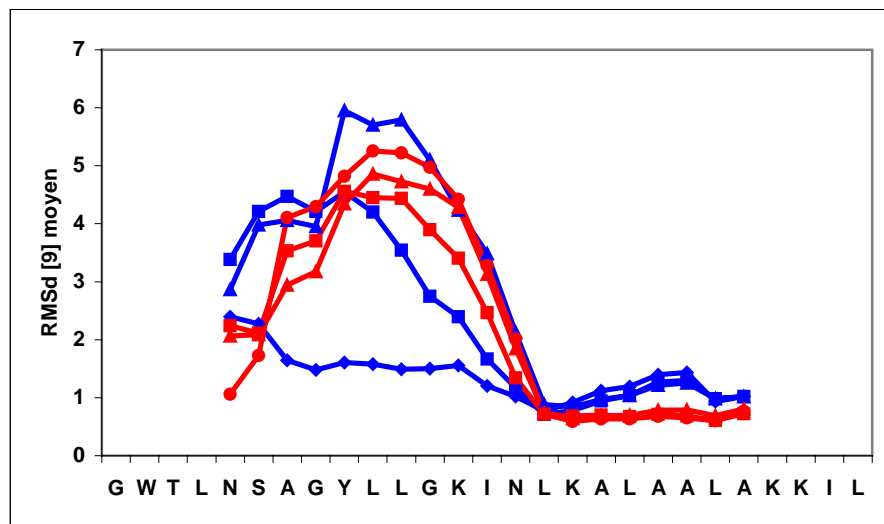


Figure 33. RMSd[9] des structures représentatives des différentes populations en milieu hydrophile (bleu, populations 1, 2 et 3 de bas en haut) et en milieu hydrophobe (rouge, populations 1, 2 et 3 de bas en haut) pour le Transportan.

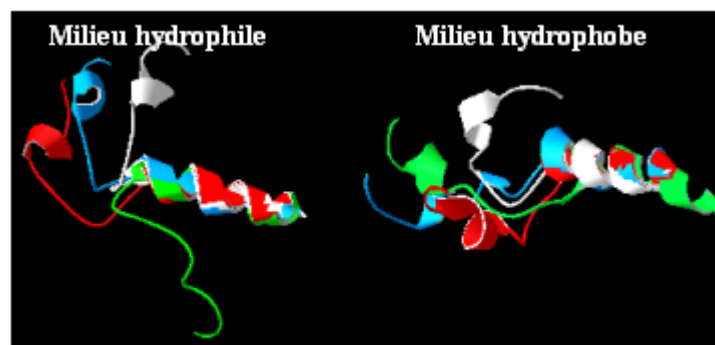
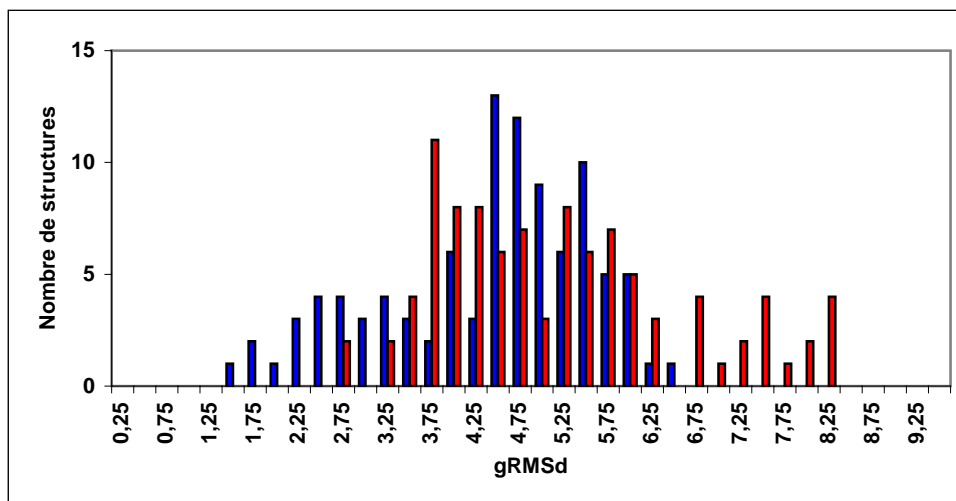


Figure 34. Structure des représentants des familles structurales du Transportan en milieu hydrophile et hydrophobe (Prime en rouge, 2<sup>ème</sup> population en bleu, troisième en blanc, quatrième en vert).

### TP07.

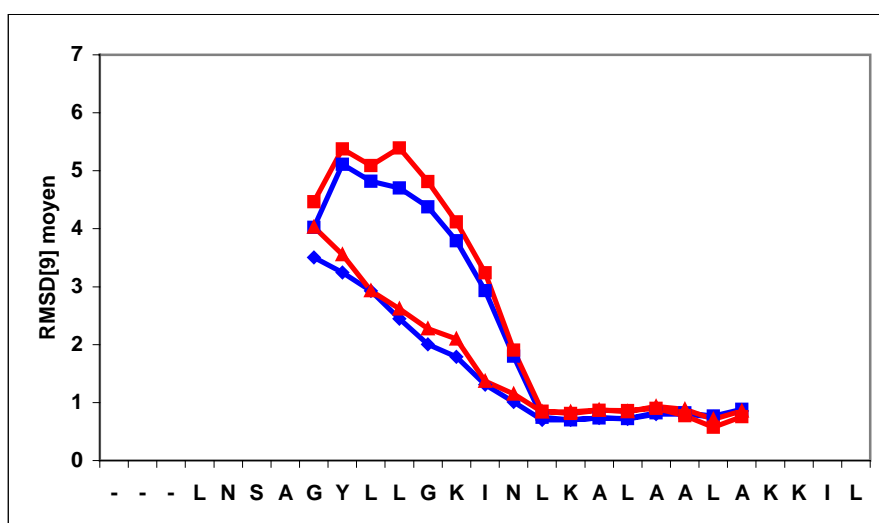
Pour TP07, l'analyse du gRMSd indique que plusieurs populations structurales coexistent. Tant en milieu hydrophobe qu'hydrophile, 2 populations apparaissent : la première inclut le Prime avec respectivement 84% et 28% des structures, la seconde

comprenant respectivement 16% des structures centrées sur un gRMSd de 6,6Å et 72% des structures centrées sur un gRMSd de 4,7Å par rapport au Prime. (**figures 35 et 37**).



**Figure 35.** Distribution des structures calculées pour le TP07 en fonction de leur gRMSd par rapport au Prime en milieu hydrophile (bleu) et en milieu hydrophobe (rouge).

L'analyse du RMSd local (RMSd[9]) moyen dans les deux milieux indique que les deux populations divergent de celle du Prime dans la partie N-terminale (**figure 36**) et est très bien conservée dans la partie C-terminale du peptide. On peut donc dire que le TP07 est un peptide polymorphe et que ce polymorphisme est principalement localisé dans sa partie N-terminale, comme pour le Transportan sauvage.



**Figure 36.** RMSd[9] des structures représentatives des différentes populations en milieu hydrophile (bleu, populations 1 et 2 de bas en haut) et en milieu hydrophobe (rouge, populations 1 et 2 de bas en haut) pour le TP07.

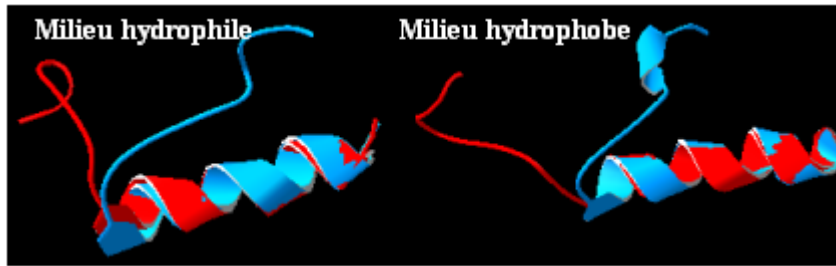


Figure 37. Structure des représentants des familles structurales de TP07 en milieu hydrophile et hydrophobe (Prime en rouge, 2<sup>ème</sup> population en bleu, troisième en blanc, quatrième en vert).

### TP10.

Pour TP10, l'analyse du gRMSd indique, comme pour le Transportan et le TP07, que plusieurs populations structurales coexistent. En milieu hydrophile, trois populations rassemblent respectivement 44% (incluant le Prime), 37% et 19% des structures centrées sur des gRMSd de 2,3Å et 5,0Å, tandis qu'en milieu hydrophobe 2 populations, outre le Prime qui est seul, apparaissent avec respectivement 85% et 14% des structures centrées sur des gRMSd de 2,9Å et 5,8Å. (figures 38 et 40).

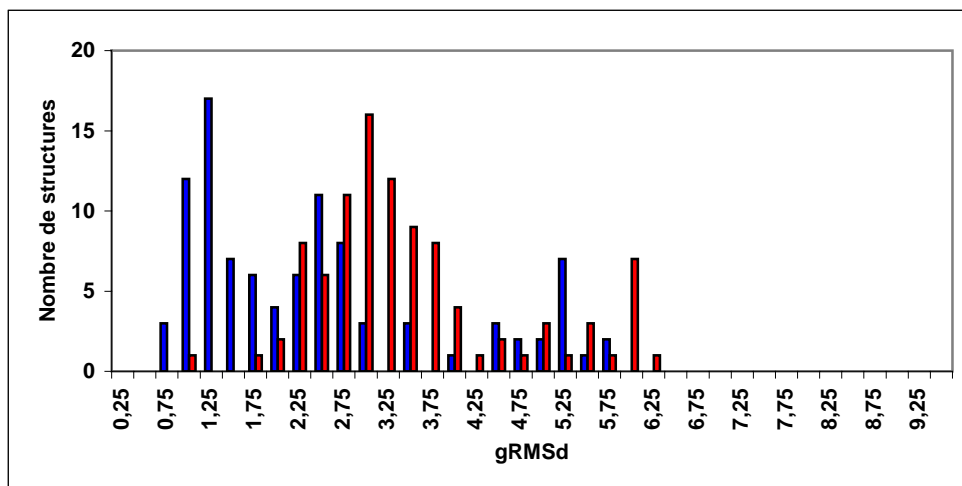


Figure 38. Distribution des structures calculées pour le TP10 en fonction de leur gRMSd par rapport au Prime en milieu hydrophile (bleu) et en milieu hydrophobe (rouge).

Ici aussi, dans les deux milieux, la structure est très bien conservée dans la partie C-terminale du peptide et plus variable dans la partie N-terminale (figure 39). Le TP10

est donc un peptide polymorphe et que ce polymorphisme est principalement localisé dans sa partie N-terminale, comme pour le Transportan sauvage et pour TP07.

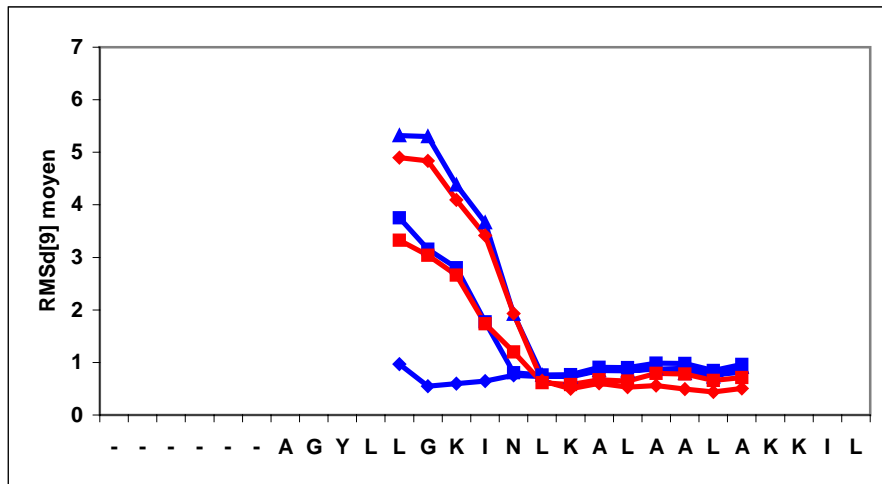


Figure 39. RMSd[9] des structures représentatives des différentes populations en milieu hydrophile (bleu, populations 1, 2 et 3 de bas en haut) et en milieu hydrophobe (rouge, populations 1 et 2 de bas en haut) pour le TP10.

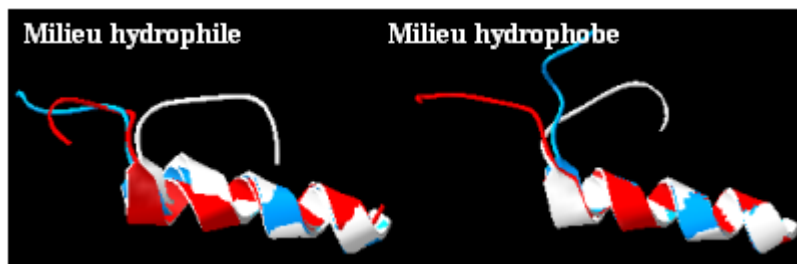
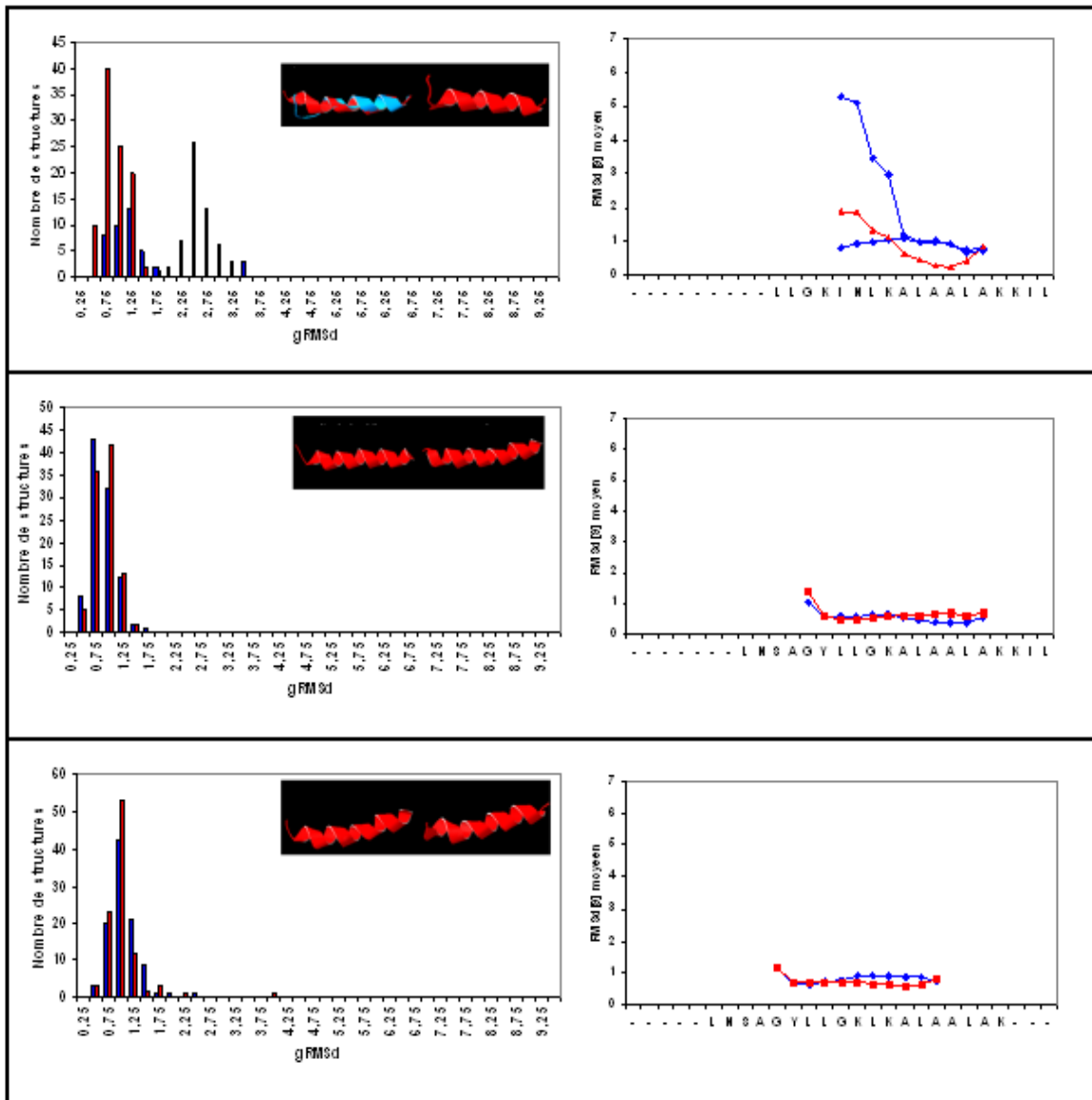


Figure 40. Structure des représentants des familles structurales de TP10 en milieu hydrophile et hydrophobe (Prime en rouge, 2<sup>ème</sup> population en bleu, troisième en blanc, quatrième en vert).

#### *TP08, TP13 et TP15.*

Le même type d'analyse pour TP13 et TP15 indique que, tant en milieu hydrophobe qu'hydrophile, 100% des structures sont regroupées en une population unique avec une déviation de RMS global faible par rapport au Prime, inférieure à 1Å. Le TP08 présente quant à lui deux populations, uniquement en milieu hydrophile : celle du Prime et une deuxième qui présente toutefois un gRMSd moyen de l'ordre de 2,6 Å, ce qui reste faible par rapport au gRMSd calculés pour le Transportan, TP07 et TP10 (figure 41).

On peut donc considérer que, contrairement au Transportan et à ses variants fonctionnels, TP 08, TP13 et TP15 sont des peptides monomorphes.



**Figure 41.** De haut en bas: TP08, TP13 et TP15. A gauche la distribution des structures calculées en fonction de leur gRMSd par rapport au Prime en milieu hydrophile (bleu) et en milieu hydrophobe (rouge). A droite le RMSd[9] des structures représentatives des différentes populations en milieu hydrophile (bleu) et en milieu hydrophobe (rouge). En insert, la structure des représentants des familles structurales en milieu hydrophile (à gauche) et hydrophobe (à droite) (Prime en rouge, 2<sup>ème</sup> population en bleu).

En conclusion il apparaît que les peptides TP7 et TP10 ainsi que le Transportan sauvage qui ont été décrits dans la littérature comme étant spontanément internalisés dans les cellules sont également polymorphes tant en milieu hydrophile qu'en milieu hydrophobe. Ce polymorphisme leur permet d'adapter leur structure en fonction du milieu

dans lequel il se trouvent. A contrario les peptides TP08, TP13 et TP15 qui ne sont pas perméables sont prédits comme étant monomorphes. Le même constat peut être fait pour la Pénétratine qui est perméable et polymorphe et ses mutants Pénétratin 48 et 48-56 qui sont non perméables et monomorphes (**figure 42**).

		gRMSd moyen		
		hydrophile	hydrophobe	Internalisation
<b>Transportan</b>	GWTLNSAGYLLGKINLKAL	5,2	5,1	+++
<b>TP07</b>	GWTLNSAGYLLGKINLKAL	4,3	3,7	+++
<b>TP08</b>	GWTLNSAGYLLGKINLKAL	1,8	0,3	-/+
<b>TP10</b>	GWTLNSAGYLLGKINLKAL	2,2	3,4	+++
<b>TP13</b>	GWTLNSAGYLLGKINLKAL	0,2	0,2	-/+
<b>TP15</b>	GWTLNSAGYLLGKINLKAL	0,4	0,3	-/+
<b>Pénétratin</b>	RQIKIWFQNRMMKWKK	1,8	2,4	+++
<b>Pénétratin 48</b>	RQIKIFFQNRMMKWKK	1,5	2,1	-/+
<b>Pénétratin 48-56</b>	RQIKIFFQNRMMKFKK	0,8	1,8	-

**Figure 42. Tableau récapitulatif des peptides étudiés, de leurs scores de polymorphisme et de leur capacité d'internalisation.**

Nous suggérons donc que la capacité d'internalisation d'un peptide est en relation avec sa capacité à modifier sa structure grâce à ses propriétés de polymorphisme, de manière à pouvoir être soluble autant dans l'eau que dans une phase lipidique.

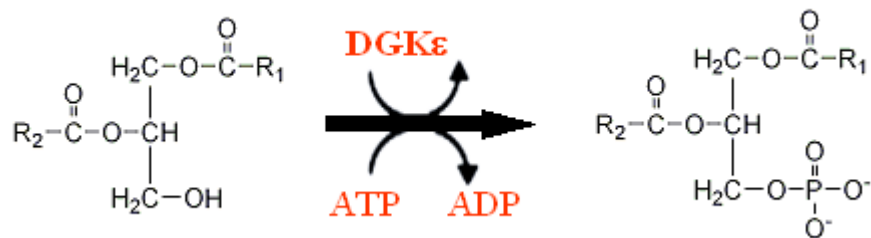
Ce travail a fait l'objet de la rédaction d'un article accepté à BBA (28/1/08) :  
***Structural Polymorphism of two CPP: An important parameter of activity.***  
***(Deshaye S, Decaffmeyer M, Brasseur R and Thomas A).***

### 3.2.3. Détermination de la topologie du segment hydrophobe de la diacylglycerol kinase $\epsilon$ de mammifère dans une membrane cellulaire et sa relation avec les prédictions en modélisation.

#### A) Introduction.

Ce travail a été réalisé dans le cadre d'une collaboration avec l'équipe du Professeur Richard Eband de l'Université McMaster à Hamilton au Canada. Cette équipe s'intéresse particulièrement à la famille des diacylglycerol kinases (DGK) humaines. Dans ce cadre il nous était demandé de vérifier que parmi les isoformes de cette protéine, seule la DGK $\epsilon$  possède un segment suffisamment hydrophobe et long pour servir d'ancre membranaire. Le deuxième objectif était de localiser et d'optimiser la longueur de ce segment potentiellement transmembranaire. Enfin, nous avons proposé d'appliquer notre méthode PepLook à la détermination de la topologie de ce segment hydrophobe de DGK $\epsilon$ .

Cet enzyme est le plus court parmi les 8 isoformes humain ( $\alpha$ ,  $\beta$ ,  $\delta$ ,  $\epsilon$ ,  $\gamma$ ,  $\iota$ ,  $\zeta$ ,  $\theta$ ) et a une spécificité pour les DAGs (Diacylglycérols) portant un groupe arachidonate. Il est impliqué dans la transformation de ces DAGs en acides phosphatidiques (**figure 43**) et est principalement localisé dans les tissus cérébraux et cardiaques, dans les poumons et la rétine.



**Figure 43. Implication de DGK $\epsilon$  dans la conversion des DAGs en acides phosphatidiques.**

Cet isoforme étant le seul à posséder un segment membranaire, il est important d'en connaître le rôle et donc la topologie. Il a par ailleurs été montré expérimentalement par nos collègues Canadiens que ce segment n'est pas impliqué dans la fonction enzymatique de la protéine mais plutôt dans son oligomérisation (**Dicu et al., 2007**).



## B) Méthodes.

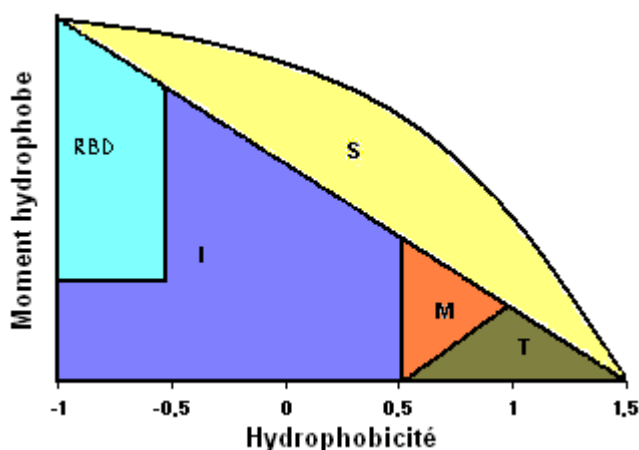
### *Méthode d'Eisenberg.*

La méthode d'Eisenberg permet à partir d'une séquence, de localiser les zones potentiellement membranaires ou solubles d'une protéine (**figure 44**). Pour chaque résidu on place sur un graphe le moment hydrophobe  $\langle \mu \rangle$  en fonction de l'hydrophobicité  $\langle H \rangle$  calculés sur une fenêtre de 11 résidus, en utilisant l'échelle d'hydrophobicité consensus d'Eisenberg (**équations 10 et 11**). L'utilisation de fenêtres de 5 ou 7 résidus permet d'identifier les zones RBD.

$$\langle H_i \rangle = \frac{1}{n} \left( \sum_{n=1}^N h_n \right) \quad \text{équation 10}$$

$$\langle \mu_i \rangle = \left\{ \left[ \sum_{n=1}^N h_n \sin(n\delta) \right]^2 + \left[ \sum_{n=1}^N h_n \cos(n\delta) \right]^2 \right\}^{1/2} \quad \text{équation 11}$$

où  $h_n$  est la valeur de l'hydrophobicité du résidu  $n$ . Les valeurs moyennes  $\langle H_i \rangle$  et  $\langle \mu_i \rangle$  sont calculées sur une fenêtre de 11 résidus et assignées à l'acide aminé central  $i$ . L'angle de gyration  $\delta$  correspond à l'angle moyen entre deux chaînes latérales dans une structure en hélice  $\alpha$  et vaut  $100^\circ$ .



**Figure 44. Représentation du graphe  $\langle \mu_i \rangle$  en fonction de  $\langle H_i \rangle$  dans la méthode d'Eisenberg.**

En fonction du positionnement des résidus sur le graphique d'Eisenberg, on peut prédire la localisation dans la structure de la protéine : Dans une zone RBD, soluble (S), intérieure de la protéine (I), segment membranaire (M) ou transmembranaire (TM).

### ***Méthode IMPALA.***

IMPALA (Integral Membrane Proteins And Lipids Association) est une méthode de simulation implicite d'interaction entre un composé et une membrane. Elle a été développée au CBMN (Ducarme et al., 1998). La simulation du solvant ou des lipides entourant le composé d'intérêt utilise une fonction mathématique empirique permettant de calculer la manière dont ce composé va s'insérer dans l'environnement, notamment en fonction de son profil d'hydrophobicité. On imagine aisément que les simulations explicites, type dynamique moléculaire, pourraient donner des résultats plus proches de la réalité biologique si les conditions de calcul sont adéquates mais elles sont difficile à définir et ces simulations sont extrêmement coûteuses en temps de calcul. A l'inverse, la simulation implicite permet l'obtention de résultats réalistes en des délais raisonnables.

IMPALA utilise classiquement la procédure de Monte Carlo pour tester l'insertion d'un composé dans une membrane modélisée ainsi que son orientation par rapport au plan de la bicouche. Toutefois, dans notre cas nous avons forcé le segment peptidique étudié à explorer toutes les profondeurs de Z de  $-40$  à  $+40\text{\AA}$  et 2000 rotations pour chacune des positions de Z.

Dans cette méthode, les propriétés de l'interface sont considérées comme constantes dans son plan (XY) et ne varient que selon Z, l'épaisseur de la membrane. L'interface lipide/eau est décrite par une fonction  $C(z)$  qui varie le long de l'axe Z qui est la normale à l'interface, et dont l'origine se situe au centre de la bicouche.  $C(z)$  est une fonction empirique qui varie de 0 dans la phase hydrophobe à 1 dans la phase hydrophile (équation 12, figure 45).

$$C_{(z)} = 1 - \frac{1}{1 + e^{\alpha(|z| - z_0)}} \quad \text{équation 12}$$

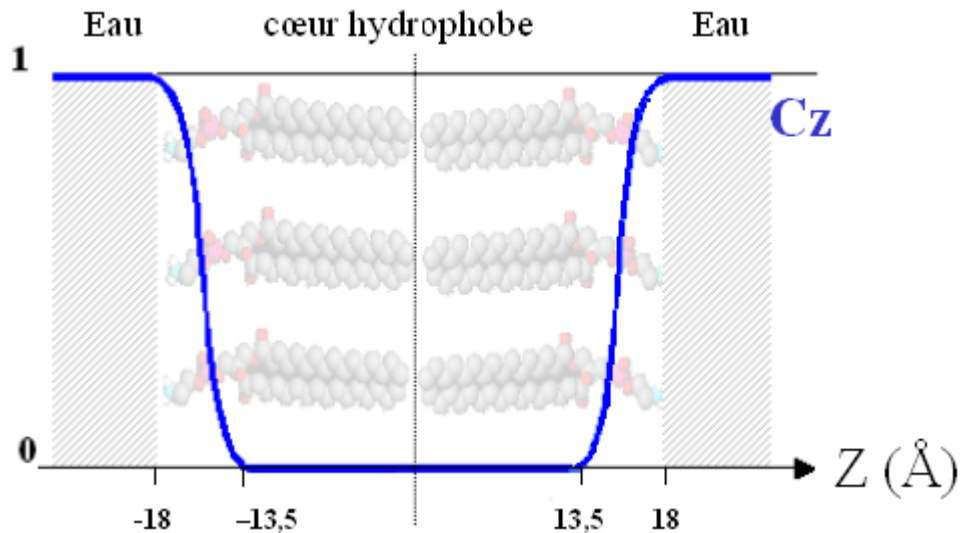
où  $\alpha = 1.99\text{\AA}^{-1}$  et  $z_0 = 15.75\text{\AA}$  (milieu des têtes polaires) sont calculés de manière à ce que  $C_{(|z|=18\text{\AA})} = 1$  et  $C_{(|z|=13.5\text{\AA})} = 0$ .

Il s'agit donc d'une fonction symétrique (**figure 45**), avec respectivement :

$C_{(z)} = 0$  de  $-13.5 \text{ \AA}$  à  $13.5 \text{ \AA}$ , donc dans le coeur hydrophobe.

$C_{(z)} = 1$  entre  $18 \text{ \AA}$  et  $\infty$  et entre  $-18 \text{ \AA}$  et  $-\infty$ , donc dans le solvant.

$0 \leq C_{(z)} \leq 1$  donc au niveau des têtes polaires.



**Figure 45. Représentation de la membrane (valeur de  $C(z)$  en bleu).**

Deux types de contraintes sont utilisées pour analyser l'insertion d'une molécule. La contrainte de perturbation lipidique et la contrainte hydrophobe, représentant respectivement les contributions mécaniques et chimiques au problème d'insertion.

- Contrainte de perturbation lipidique ( $E_{lip}$ ): Les interactions entre lipides dans une membrane sont favorables puisqu'ils sont organisés de manière à appairer d'une part leurs têtes polaires et d'autre part leurs chaînes acylées. L'insertion d'un composé étranger perturbe cette organisation. Cette perturbation est d'autant plus importante que le composé est volumineux et qu'il offre une grande surface de contact (**équation 13**).

$$E_{lip} = a_{lip} \sum_{i=1}^N S_i C(Z_i) \quad \text{équation 13}$$

où  $a_{lip}$  est une constante empirique positive et  $S_i$  la surface accessible de l'atome  $i$ .

- **Contrainte hydrophobe:** Les atomes hydrophobes préfèrent se situer dans un environnement apolaire (dans le cœur hydrophobe de la membrane). A contrario, cet environnement est défavorable pour des atomes hydrophiles. Cette contrainte hydrophobe est fonction de l'hydrophobicité de l'atome considéré (donnée par son énergie de transfert) et de son accessibilité (donnée par la surface accessible) (équation 14).

$$E_{pho} = - \sum_{i=1}^N S_{(i)} E_{tr(i)} C_{(z_i)} \quad \text{équation 14}$$

où  $E_{tr(i)}$  est l'énergie de transfert de l'atome  $i$ .

### C) Résultats et discussion.

#### C1) Recherche des segments potentiellement transmembranaires.

Dans la première étape de ce travail nous avons recherché les segments potentiellement transmembranaires parmi les 8 isoformes de la DGK humaine à l'aide de la méthode d'Eisenberg. Nous avons pu confirmer que seule la DGK $\epsilon$  possède un segment hydrophobe suffisamment long que pour être transmembranaire. Il est situé dans la partie N-terminale de la protéine, de G19 à F37 (**figure 46**). Il nous a paru toutefois que ce segment était probablement un petit peu court pour traverser la membrane s'il était sous forme d'hélice. De plus, la suite de la séquence, est clairement hydrophobe (**WCSL**) et la position de ces résidus sur le graphe d'Eisenberg est essentiellement due à l'influence de résidus hydrophiles qui suivent (**QRSRRQ**). Cette hypothèse est confirmée par le fait que le segment membranaire prédit par la méthode d'Eisenberg s'allonge si on réduit la fenêtre de calcul.

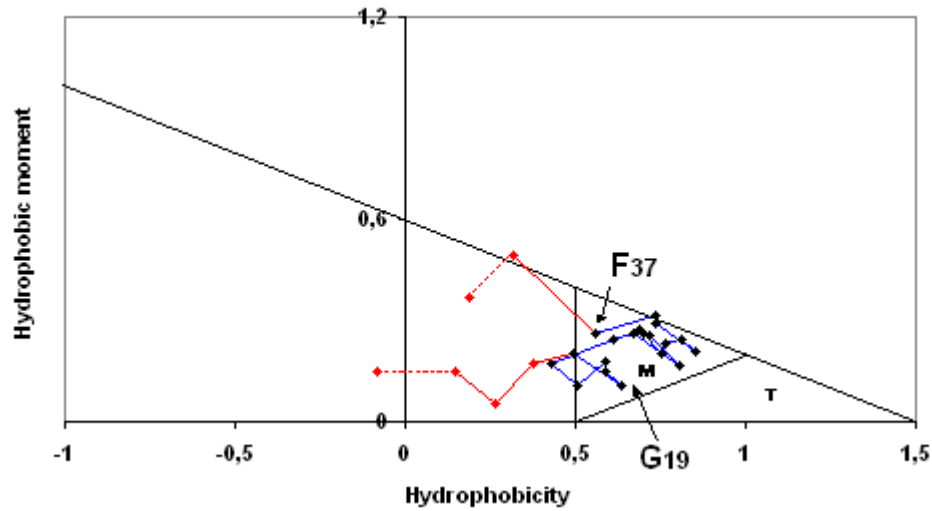


Figure 46. Graphe d'Eisenberg pour DGKε. En bleu, les résidus prédits comme membranaires et en rouge, les zones non membranaires amont et aval.

Dans une deuxième étape nous avons cherché à optimiser la longueur du segment. A cet effet, la méthode IMPALA à été utilisée pour étudier l'insertion du segment G19-F37 et de segments allongés en C et N-terminal (figure 47).

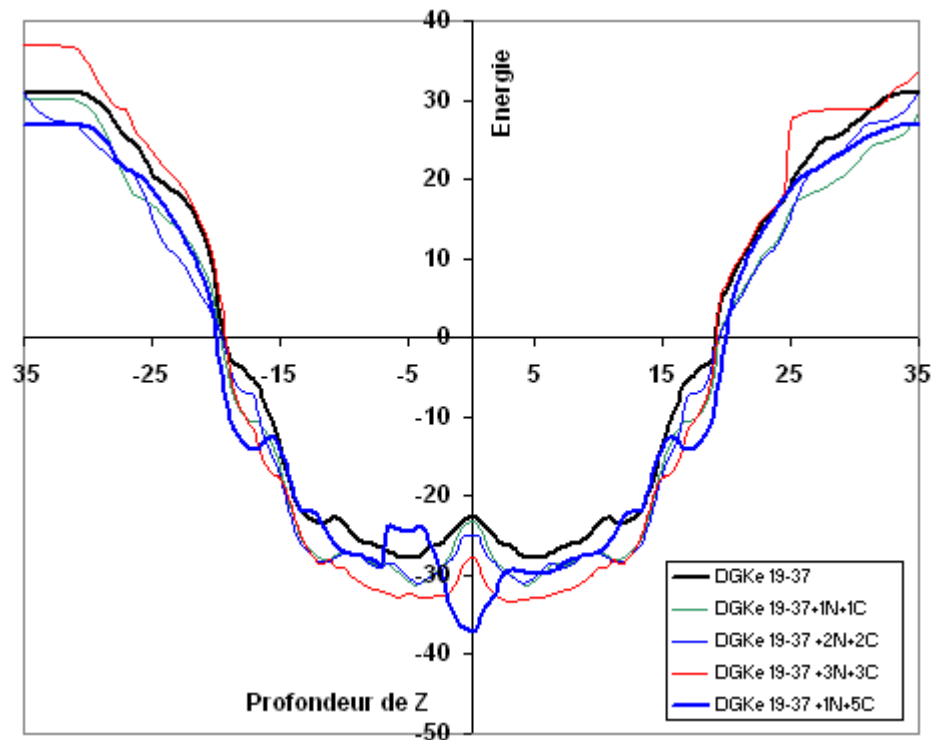
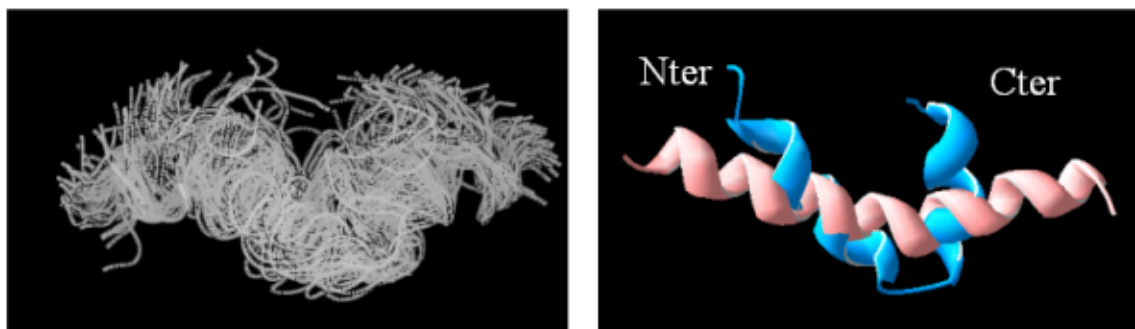


Figure 47. Energie d'interaction des segments allongés de DGKε en fonction de Z.

Le travail a été réalisé sur toutes les combinaisons de segments allongés de +1 résidus à +6 résidus tant du côté N-terminal que du côté C-terminal. Ces segments ont été

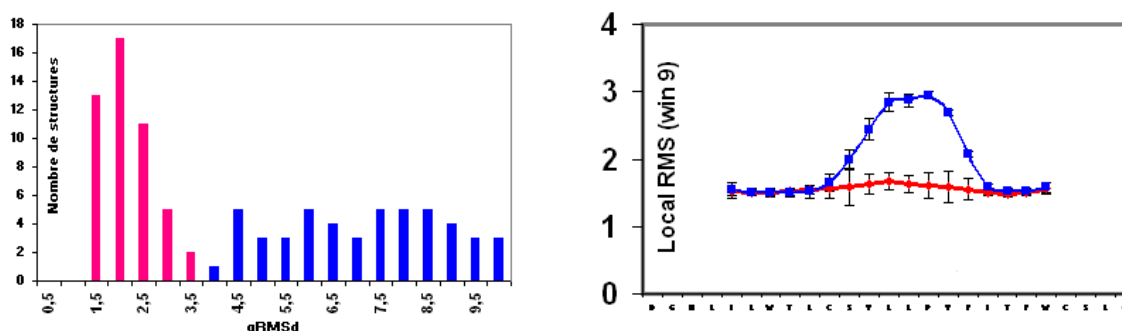




**Figure 49.** Les 99 structures de basse énergie calculées par PepLook (gauche). Structures représentatives des deux populations structurales avec le Prime en rose et le représentant de la 2<sup>ème</sup> population en bleu (droite).

L'analyse de la structure secondaire de tous les résidus des 99 modèles de D18-Q42 montre une très large contribution hélicoïdale (en moyenne 74%). Le reste de la structure est répartie en une contribution random coil à 21% et étendue à 5% conformations (**figure 51**).

On observe en réalité deux populations structurales bien distinctes qui comportent chacune 50% des structures (**figure 50, gauche**). La première, avec un RMSd moyen de 1,9Å, est celle du Prime et est composée de structures hélicoïdales tout au long du segment tandis que la deuxième population présente un RMSd moyen de 6,8 Å et est représentée par le modèle 89 sous la forme d'une hélice brisée en son centre pour donner une forme en U (ou forme «U bent») (**figure 49**). L'analyse du RMSd[9] indique que les modèles de la deuxième population peuvent se superposer à la structure du Prime aux extrémités et non au centre du segment (**figure 50, droite**). D'autre part, la déviation standard nous informe sur le fait que la variabilité structurale à l'intérieur de chacune des deux populations est faible.



**Figure 50.** Distribution des structures calculées en fonction de leur gRMSd par rapport au Prime (première population en rouge, deuxième en bleu) (gauche). RMSd [9] le long de la séquence. La structure de référence est le Prime. La courbe rouge représente la valeur moyenne et la déviation standard pour la population du Prime. La

courbe bleue représente la valeur moyenne par rapport au Prime et la déviation standard pour la seconde population (droite).

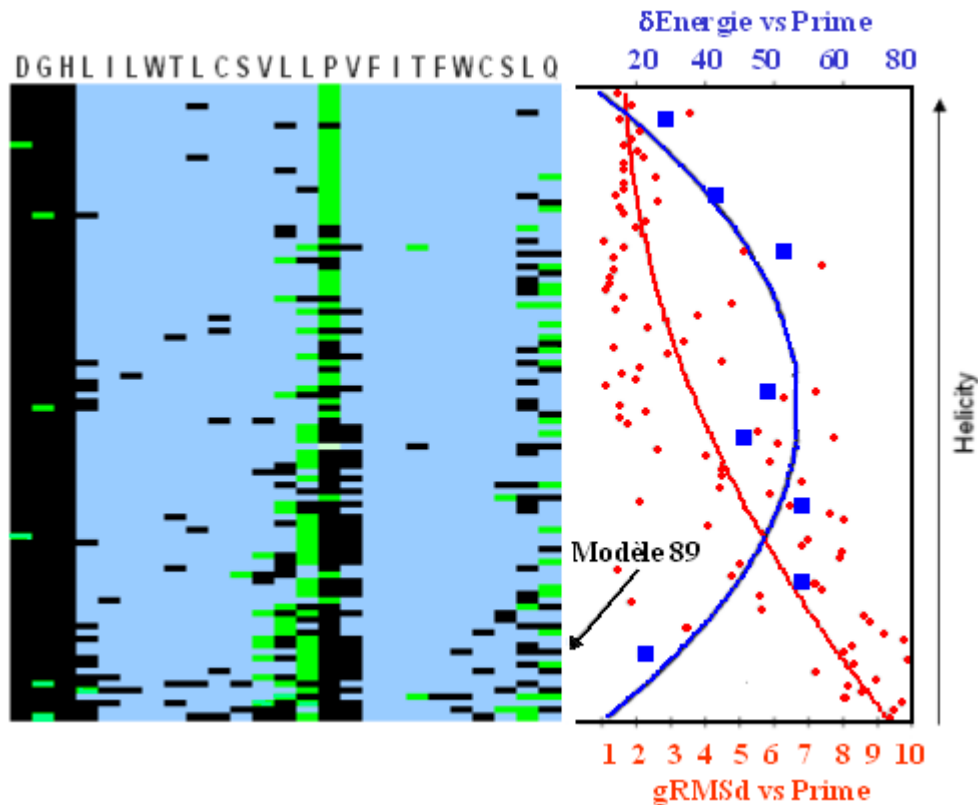


Figure 51. Les 99 modèles PepLook classés selon leur taux d'hélicité décroissant de haut en bas. Le bleu est utilisé pour les résidus en hélice, le noir pour le random coil et le vert pour les autres structures. Le modèle le plus hélicoïdal est le Prime, le modèle 89 représente la seconde population (gauche). Les points et la courbe rouges représentent le gRMSd (en Å) par rapport au Prime et les points et la courbe en bleu, le delta d'énergie (en kcal) par rapport au Prime des 8 peptides les plus stables.

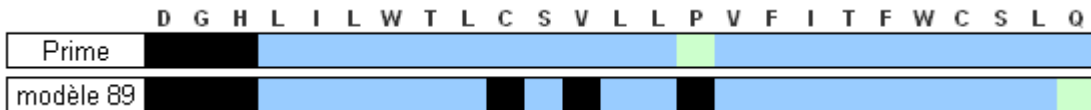
Notons d'une part que l'augmentation du gRMSd est due à une perte d'hélicité dans certaines zones du segment, et d'autre part que les représentants des deux populations correspondent à deux minima énergétiques alors que les structures intermédiaires sont moins favorables de ce point de vue (**figure 51**). Ceci indique que deux structures énergétiquement stables coexistent.

### C.1.2. Analyse de la structure secondaire de DGKε 18-42.

Si l'on regarde plus particulièrement la structure secondaire des représentants des deux populations (le Prime et le modèle 89), on observe que le Prime est presque complètement hélicoïdal sauf sur trois résidus en N-ter (D18-H20) et sur la proline en



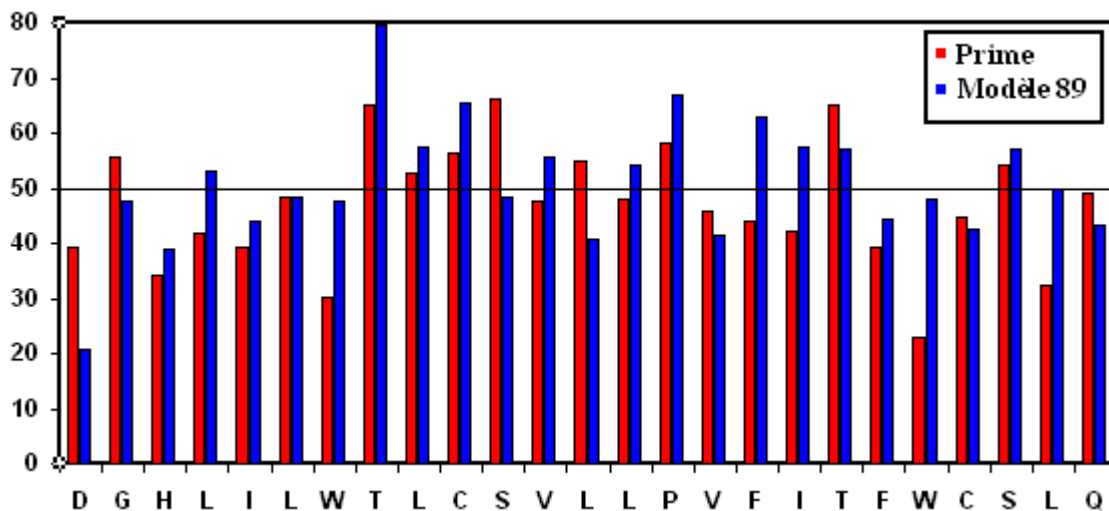
position 32 (**figure 52**). Cette proline induit une très légère déviation de l'hélice (**figure 49, droite**). Le modèle 89 quant à lui présente une structure secondaire qui s'apparente plus à du random coil en son centre, ce qui permet au peptide d'adopter une conformation en U (**figure 49, droite**). Ces résultats sont en conformité avec les données expérimentales obtenues par CD (**Glukhov et al., 2007**)



**Figure 52. Structure secondaire du Prime et du modèle 89 (zoom de la figure 51).**

### *C.1.3. Analyse de la stabilité de DGK $\epsilon$ 18-42.*

Comme nous l'avons précédemment écrit à l'occasion de la description de la méthode PepLook, un score de stabilité par comparaison du MFP au MFP de référence extrait d'un set de protéines repliées est calculé. On considère que les résidus présentant un score inférieur à 50-60% sont en carence de stabilité et sont donc candidats potentiels à une interaction intermoléculaire.



**Figure 53. Score de stabilité (en %MFP) du Prime(rouge) et du modèle 89 (bleu).**

L'analyse du score de stabilité montre que la plupart des résidus sont candidats à une interaction avec un partenaire externe puisque le score de stabilité moyen est de 51% pour le modèle 89 et de 47% pour le Prime. Les parties N-terminale (D18-W24) et C-terminale (V33-Q42) du Prime ont un score de stabilité plus faible (42%) que le centre du

peptide (56%). Dans la forme «U bent» (ou forme en U – modèle 89 – 2<sup>ème</sup> population) certains des résidus qui sont instables lorsque le peptide adopte la structure du Prime, sont mieux stabilisés par des interactions intramoléculaires qui ne sont pas possible dans la forme hélicoïdale : L21, W24, T25, V29, L31, F34, W38 et L41. Deux résidus hydrophobes (L30 et V33) sont toutefois moins stabilisés dans cette conformation et restent candidats pour une interaction extramoléculaire.

#### ***C.1.4. Recherche de partenaires externes pour DGKε 18-42.***

Nous avons précédemment montré que le segment DGKε 18-42 est inséré dans la membrane. Dès lors il y a des partenaires potentiels naturellement présents dans l'environnement du segment : la DPPC, le cholestérol ou le segment 18-42 d'une autre DGKε. Nous avons analysé l'interaction des deux formes de DGKε 18-42 avec les différents partenaires potentiels. A l'aide de la procédure Hypermatrice, décrite dans l'article relatif au design de peptides complémentaires, nous avons calculé l'énergie d'interaction de tous ces partenaires potentiels entre eux.

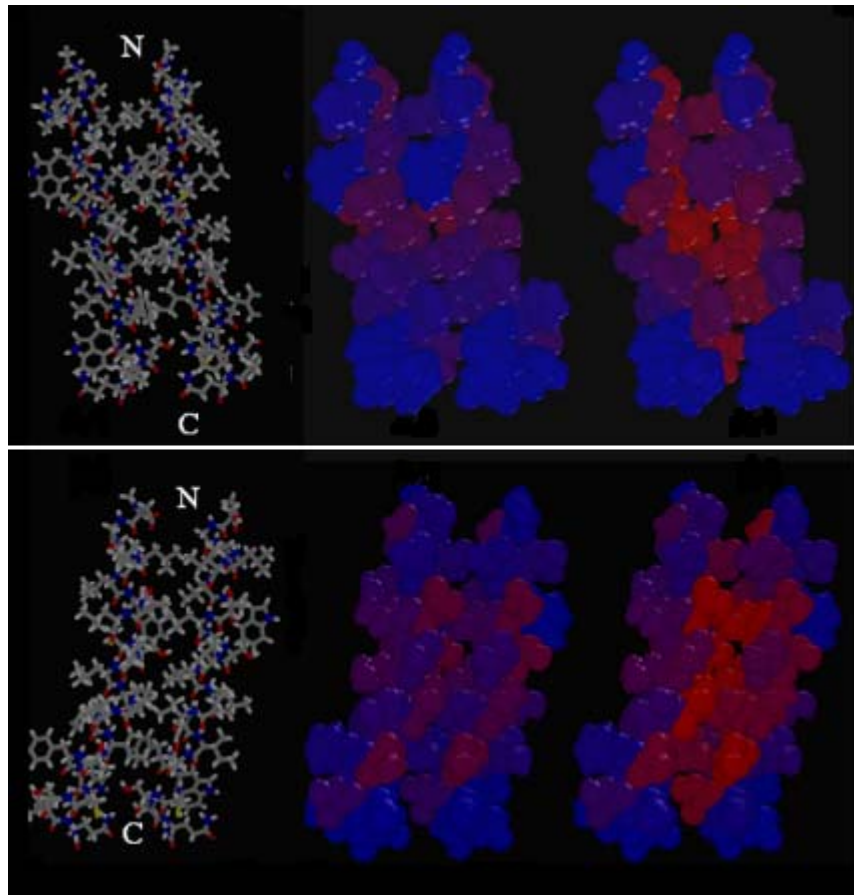
Les résultats indiquent que la forme «U bent» gagne plus de stabilité lorsqu'elle interagit avec le cholestérol qu'avec la DPPC (gain respectif de 75 kCal et 67 kCal). De plus le cholestérol gagne également plus à interagir avec le segment DGKε 18-42 sous la forme «U bent» qu'à s'auto-associer (gain respectif de 75 kCal et 52 kCal).

A contrario, le segment DGKε 18-42, lorsqu'il adopte une structure hélicoïdale, s'associe plus volontiers avec la DPPC qu'avec le Cholestérol (gains respectifs de 107 kCal et 71 kCal). On peut donc en conclure que la forme «U bent» s'associe préférentiellement avec le cholestérol, ce qui n'est pas le cas pour le Prime. D'autre part, si l'on se place du point de vue des lipides, tant la DPPC que le Cholestérol gagnent plus de stabilité en s'associant au segment DGKε 18-42, quelle que soit sa structure, qu'en formant des homo-dimères.

Ensuite, nos résultats montrent que dans sa structure Prime, le segment DGKε 18-42 gagne 209 kCal (soit un gain de stabilité de 25%) en se dimérisant pour seulement 122kCal pour la forme «U bent» ( gain de stabilité de 13%). Ceci indique que la tendance à la dimérisation est réelle pour ce peptide lorsqu'il adopte une conformation sous forme d'hélice. La trimérisation de la forme Prime apporte 17% de gain supplémentaires en

stabilité. Ces résultats sont en accord avec les données expérimentales qui montrent qu'un peptide similaire (L21-W38) a clairement tendance à dimériser, et dans une moindre mesure de trimériser lorsqu'il se conforme en hélice (Glukhov et al., 2007).

Une analyse plus fine à l'aide de l'outil PEX (Thomas et al., 2001) nous a permis d'identifier les résidus du peptides qui ont gagné en terme de stabilité dans la dimérisation de la forme hélicoïdale : Il s'agit de L21, W24, T25, S28, I35, et W38 sur un monomère et G19, I22, L23, L26, V29, L30, P32 V33, T36, F37 et S40 sur le second monomère (figure 54).



**Figure 54. Gain de stabilité pour le dimère de sous forme Prime. A gauche, une vue du dimère tel que calculé par Hypermatrice. Au milieu, représentation chromatique du score de stabilité des résidus lorsqu'ils étaient dans le monomère (de bleu (<35%) à rouge (>85%)). A droite, représentation chromatique du score de stabilité des résidus lorsqu'ils sont dans le dimère. Le panneau du bas représente une rotation de 180° par rapport à celui du haut.**

En conclusion, les deux conformations (hélice et «U bent») coexisteront et pourraient s'insérer dans la membrane. La forme hélice a plutôt tendance à dimériser et n'a pas plus de préférence pour le Cholestérol pour la DPPC, alors que la forme «U bent»

aura moins tendance à dimériser mais liera préférentiellement le cholestérol plutôt que la DPPC.

Habituellement les protéines membranaires sont classifiées en fonction du nombre de fois qu'elles traversent la membrane : polytopiques pour plusieurs traversées, bitopiques pour une traversée unique et monotopique lorsqu'un segment est suffisamment hydrophobe pour lier la membrane sans toutefois la traverser de part en part (**Blobel, 1980**). Les protéines monotopiques sont rares (**Fowler et al., 2007**).

Nous avons ensuite cherché à savoir si la DGK $\epsilon$  est une protéine monotopique ou bitopique et donc s'il est possible que son segment membranaire, plutôt que de traverser la membrane de part en part, adopter une conformation en «U bent» telle que nous l'avions prédite. Une recherche bibliographique nous indique qu'au moins une protéine de ce type existe : la Cavéoline (**Spisni et al, 2005**).

Afin de tenter de répondre à ces questions, l'équipe de R. Epand a réalisé des mesures de microscopie de fluorescence sur des fibroblastes transfectés par un vecteur d'expression de DGK $\epsilon$  marqué en N-terminal par un épitope de type FLAG (et co-transfecté par pmRFP-C1 qui produit une protéine soluble qui fluoresce dans le visible et servira de témoin d'expression). Ensuite, les cellules ont été incubées avec un anticorps anti-FLAG puis avec un second anticorps fluorescent. Ces mesures montrent que, sauf si l'on perméabilise les cellules au Triton X-100, l'anticorps fluorescent n'est pas visible, ce qui indique que le FLAG, et donc la partie N-terminale (séquence amont au segment membranaire (1 à 17)) de la protéine, est inaccessible aux anticorps et donc intracellulaires (**figure 55**). Plusieurs hypothèses étaient à priori envisageables quant à la topologie de la protéine.

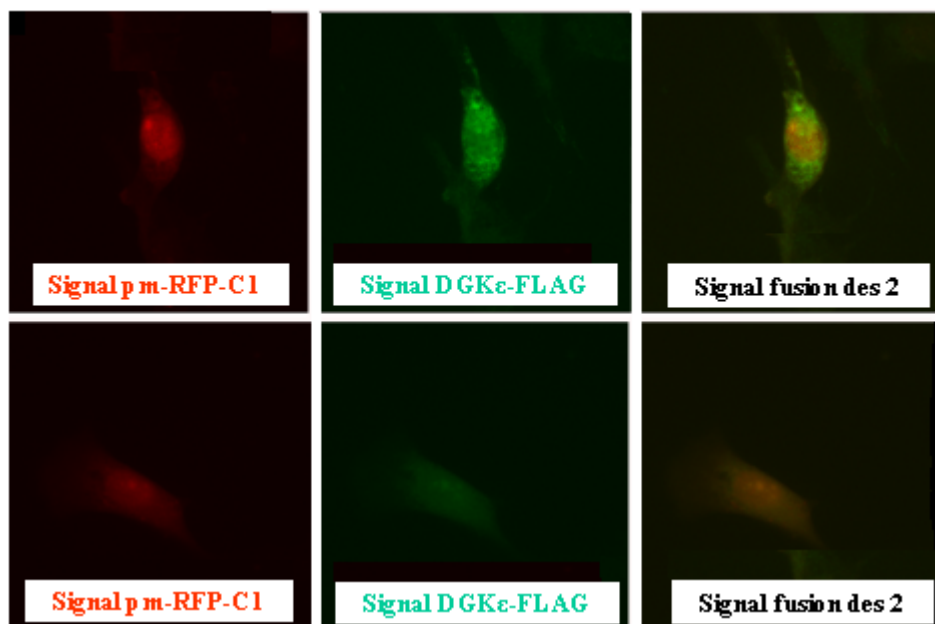
La première hypothèse est que la protéine serait bitopique avec sa partie soluble dans la région intra-cytoplasmique (**figure 56**). Dans ce cas le FLAG situé en amont du segment transmembranaire aurait nécessairement dû donner un signal en fluorescence, ce qui n'est pas le cas.

La deuxième hypothèse est que la protéine est un ectoenzyme bitopique, auquel cas sa partie soluble serait extra-cytoplasmique et le FLAG pointerait vers le cytoplasme (**figure 56**), expliquant l'absence de signal en fluorescence. Cette hypothèse doit être écartée puisque toute la machinerie protéique impliquée dans la régulation de la

conversion des diacylglycérols en acides phosphatidiques est cytoplasmique. La DGK $\epsilon$  est partie prenante dans ce mécanisme de régulation, et son site actif est sur sa partie soluble qui ne peut dès lors qu'être située dans le cytoplasme.

La troisième hypothèse consiste à dire que la partie soluble de DGK $\epsilon$  est dans le cytoplasme et que le segment hydrophobe D18-Q42 adopte une conformation en forme de U telle que suggérée par nos résultats de modélisation (**figure 56**). En effet cette conformation du segment membranaire explique l'absence de signal en fluorescence.

Une dernière hypothèse consisterait à imaginer que la partie soluble de la protéine est extra-cytoplasmique et que le segment membranaire adopte la conformation que nous avons prédite (**figure 56**). Cette solution doit être écartée puisque que la protéine n'est pas un ecto-enzyme, pour les raisons évoquées plus haut, et que d'autre part, cette topologie permettrait d'obtenir un signal en microscopie de fluorescence puisque le FLAG pointerait vers le solvant.



**Figure 55. Résultats de microscopie de fluorescence: en haut, après perméabilisation, en bas sans perméabilisation.**

En conclusion, la seule topologie possible est celle dans laquelle la DGK $\epsilon$  est un enzyme monotopique avec, tant la partie C- que la partie N-terminale, dans le cytoplasme. Cette topologie n'est possible que si le segment membranaire adopte une conformation en U telle que nous l'avons prédite.

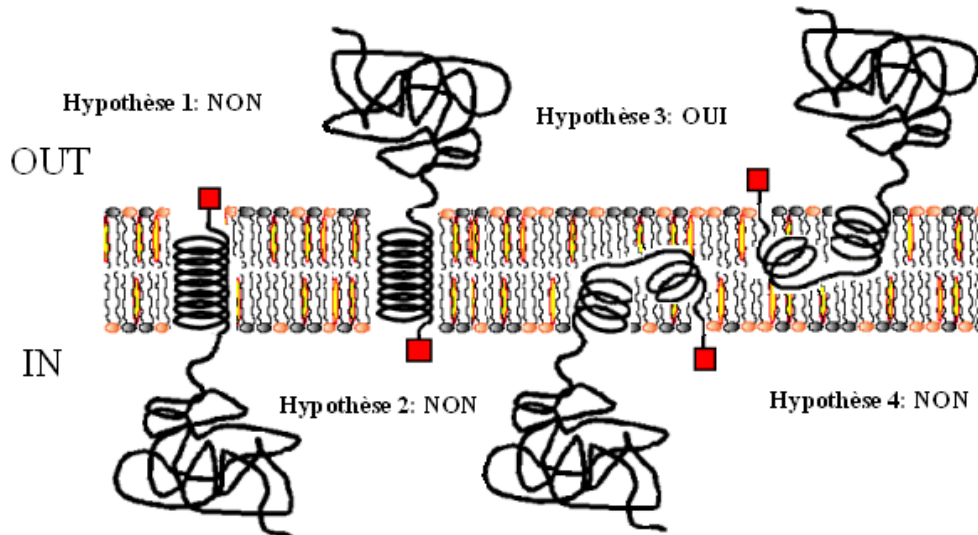


Figure 56. Les 4 hypothèses concernant la topologie de DGK $\epsilon$ . ■ = FLAG.

### *Rôle de la proline 32.*

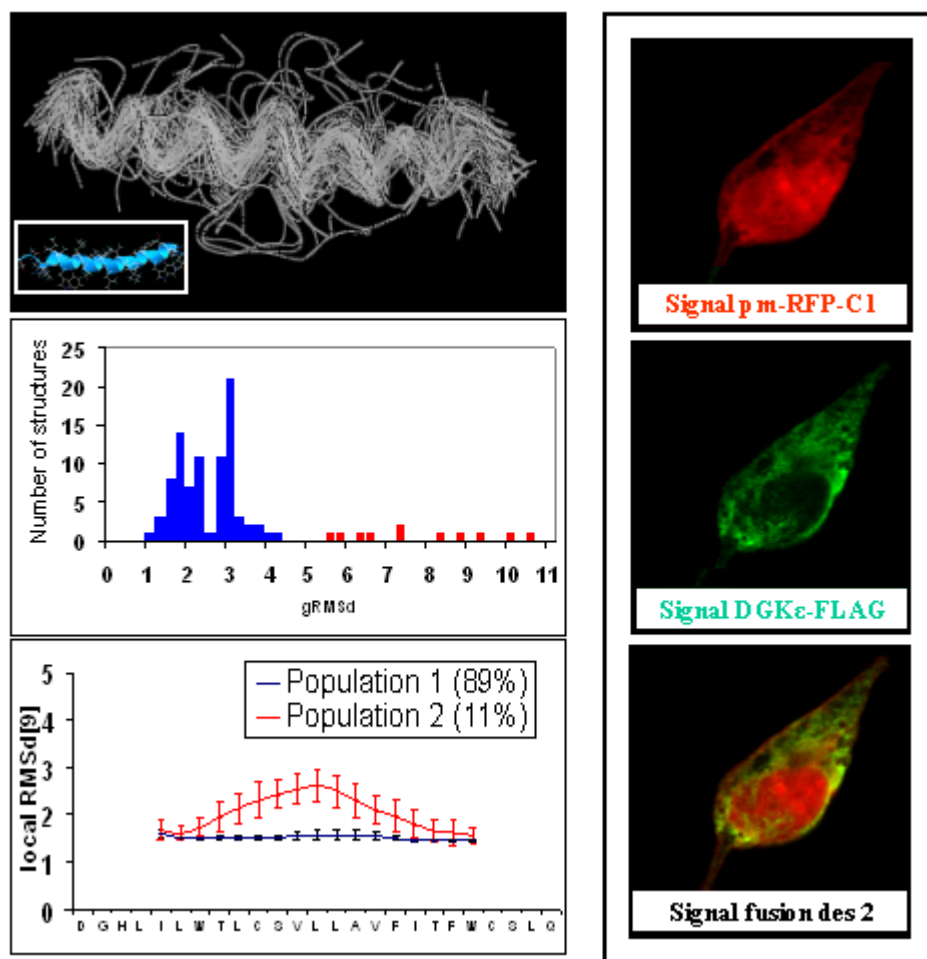
A priori, on retrouve peu de prolines dans les hélices. Toutefois certaines hélices de protéines membranaires polytopiques comportent des prolines sans que ça n'affecte sévèrement la conformation en hélice. D'autre part, un autre exemple de segment hydrophobe adoptant une forme «U bent» existe. En effet, la Cavéoline présente un segment de ce type au centre de sa séquence, ce segment hydrophobe contient également une proline.

Il semble évident que dans notre cas, la P32 joue un rôle déterminant dans la déstabilisation de l'hélice et est nécessaire pour permettre au segment membranaire d'adopter la conformation en U. On sait en effet qu'en raison de sa structure cyclique incluant la fonction amine, la proline ne peut pas adopter des angles  $\psi$   $\theta$  compatibles avec la formation d'une hélice. D'autre part, elle ne dispose pas d'un donneur d'hydrogène (NH) et ne peut donc pas contribuer à la stabilisation de l'hélice en participant à la formation d'un pont hydrogène.

Pour vérifier le rôle de la proline dans la déstabilisation de l'hélice conduisant à une conformation en U, nous avons réalisé une analyse PepLook sur le segment membranaire avec la proline mutée en alanine. Les résultats de cette analyse nous indiquent que la mutation abolit la propension du segment à se conformer en «U bent». En effet, même si deux populations existent (**figure 57**), la première contient des structures

hélicoïdales et est largement prédominante (89 % de l'ensemble), tandis que la seconde qui rassemble des structures autres que des hélices ne représente que 11% de l'ensemble. Parmi ces dernières, moins de la moitié présentent une conformation en U telle que celle observée pour le D18-Q42 sauvage.

Ces résultats sont confortés par l'analyse expérimentale. En effet les données en microscopie de fluorescence sur cellules non perméabilisées indiquent que les anticorps ont accès au FLAG (**figure 57**), ce qui n'était pas le cas pour la protéine sauvage (**figure 55**). Ceci signifie que la mutation de la proline en alanine a pour effet de favoriser la forme en hélice, permettant dès lors de faire pointer le FLAG à l'extérieur du cytoplasme.



**Figure 57.** A gauche : de haut en bas : vue des 99 structures calculées par Peplook (Prime en insert), distribution des structures par gRMSd, RMSd[9] pour les deux populations. A droite : résultats de microscopie de fluorescence sur cellules non perméabilisées.

En conclusion, on peut dire que nos prédictions sont corrélées aux résultats expérimentaux (**Glukhov et al., 2007**) de nos collègues. Les mesures en CD (dichroïsme circulaire) menées sur un peptide libre correspondant au segment membranaire montrent

que son taux d'hélicité varie en fonction du milieu. Dans le SDS le peptide se structure intégralement en hélice. En milieu membranaire (PC:PS=9:1), la contribution hélicoïdale est bien moindre que dans le SDS et diminue encore drastiquement en présence de Cholestérol, ce qui indique que la forme en U est favorisée dans la membrane et particulièrement dans les zones riches en Cholestérol. Il a également été montré en DSC (differential scanning calorimetry) que dans ces conditions membranaires, c'est à dire lorsque le peptide adopte la forme en U, il a une interaction préférentielle avec le Cholestérol plutôt qu'avec les phospholipides (**Glukhov et al., 2007**).

D'autre part, le peptide a tendance à s'oligomériser dans les conditions où il adopte une forme hélicoïdale (**Glukhov et al., 2007**). Nos collègues canadiens étudieront dans un proche avenir les conséquences fonctionnelles de la topologie qui a été mise en lumière.

Ce travail fait l'objet d'un article actuellement en soumission : ***Decaffmeyer M., Dicu A., Thomas A., Topham MK., Brasseur R. and. Epand RM. (2007). Determination of the topology of the hydrophobic segment of mammalian diacylglycerol kinase epsilon in a cell membrane and its relationship to predictions from modeling. (en soumission).***



## **4. CONCLUSION GENERALE ET PERSPECTIVES**

Les technologies liées aux peptides méritent assurément que l'on s'y intéresse. En effet, nous avons montré dans cette thèse tout l'intérêt que leur portent les sociétés biotechnologiques et plus particulièrement l'industrie pharmaceutique qui y trouve une nouvelle source de molécules thérapeutiques innovantes susceptibles tantôt de contrebalancer leurs problèmes de pipe-line, tantôt de remplacer certains antibiotiques dont l'activité décroît avec l'apparition de souches résistantes. A ce jour pourtant peu de méthodes *in silico* sont disponibles que ce soit pour la prédiction de structure ou la prédiction de liaison à une cible.

L'équipe du CBMN a donc cherché à développer deux méthodes *in silico* innovantes qui sont autant d'outils d'aide à la décision dans le domaine de la conception de peptides. D'une part, avec PepDesign, nous avons montré qu'il est possible de modéliser, *in silico* et de manière rationnelle, un peptide affiné pour une cible. Les résultats expérimentaux obtenus avec les peptides modélisés montrent que la méthode de prédiction est efficace. Notons toutefois qu'elle nécessite d'avoir accès à la structure du complexe de départ et n'est donc utilisable que si l'on dispose d'un peptide complémentaire « template ».

Dans le même ordre d'idée, puisqu'il a été possible de modéliser un peptide complémentaire affiné pour A $\beta$ 29-42, il sera possible également de concevoir un peptide complémentaire pour d'autres cibles peptidiques telles que par exemple le peptide de fusion de SIV qui dispose lui aussi d'un partenaire naturel pouvant servir de « lead ». D'autre part, et dans la mesure où l'équipe du CBMN a aussi démontré qu'il était possible de prédire un peptide non naturel avec une propriété donnée (Lins et al., 2006), rien ne s'oppose plus à la conception d'un peptide complémentaire non naturel. Le programme a été adapté pour pouvoir utiliser des acides aminés non naturels et la prédiction d'un peptide anti-SIV fait l'objet d'un travail de fin d'études qui est en cours.

D'autre part, la méthode PepDesign est complètement transposable à la conception d'un peptide ayant pour cible autre chose qu'un peptide. Par exemple, elle peut être utilisée pour concevoir un peptide interférant dans l'interaction entre deux protéines. Un travail de ce type, en collaboration avec l'UCL et l'Ulg, est également en cours au CBMN. Il consiste à proposer des peptides capables de perturber l'interaction

entre HoxA1 et Pbx, deux protéines dont l'implication dans la formation d'un complexe avec l'ADN est impliquée dans certaines formes du cancer du sein.

Enfin, PepDesign pourrait également être utilisé pour prédire et/ou améliorer, dans une stratégie non covalente, l'interaction entre le vecteur et son cargo.

Concernant la méthode PepLook, nous avons montré qu'elle donne accès non seulement à la structure d'un peptide, mais aussi à sa labilité et à sa distribution structurale qui nous semblent essentielles pour la compréhension des relations structure/fonction de ces molécules. Le score de stabilité proposé permet en outre d'identifier les résidus potentiellement candidats à une interaction avec un partenaire externe que ce soit un peptide, une protéine, une drogue ou encore un phospholipide membranaire. La comparaison de nos résultats avec les données RMN relatives à plusieurs peptides indique que la méthode que nous avons développée est de meilleure qualité que les méthodes de prédiction existantes (**Thomas et al, 2006**). D'autre part, c'est la seule méthode qui donne accès à des informations sur le polymorphisme de la molécule étudiée. Toutefois la méthode peut encore être améliorée, par exemple en « spécialisant » les couples d'angles  $\Phi\Psi$  pour chaque type de résidus.

A titre d'application, elle a permis de vérifier l'hypothèse de l'importance de la versatilité structurale pour la capacité d'internalisation cellulaire des peptides. C'est le cas des CPPs (Cell Penetrating Peptides) pour lesquels le polymorphisme semble être une propriété essentielle pour l'internalisation. Ceci devrait nous permettre à l'avenir d'être capable d'assister des équipes qui conçoivent ces molécules, d'une part en proposant des séquences qui conduisent à des peptides polymorphes et d'autre part en proposant, tôt dans le processus de conception, d'écarter des séquences qui ne le feraient pas.

D'autre part l'utilisation de la méthode PepLook a permis de proposer une topologie cohérente pour le segment transmembranaire de DGK $\epsilon$ . Nous envisageons de reproduire le même type d'analyse pour la Caveolin, une autre protéine pour laquelle une topologie de type monotopique avec une forme en U a été suggérée (**Spini et al., 2005**). Ce travail est en cours actuellement en collaboration avec l'équipe du Professeur R. Epanand au Canada.

L'utilisation thérapeutique des peptides est souvent conditionnée par une modification préalable, par exemple pour éviter ou amenuiser l'effet de la dégradation

biologique induite par les peptidases. Ces modifications consistent souvent en une glycylation, une cyclisation ou encore en l'utilisation d'acides aminés D, voire même de résidus non naturels ou de composés mixtes organique/peptidiques. Par comparaison entre les structures prédites des peptides sauvages et des peptides modifiés, il devrait être possible, en utilisant PepLook, d'étudier l'influence de ces modifications sur la structure et donc sur la fonctionnalité d'un peptide. Un travail allant dans ce sens est en cours au CBMN afin d'étudier l'influence de la  $\beta$ -méthylation des phénylalanines de la substance P sur la structure de ce peptide. Nous espérons pouvoir corréler nos résultats avec les données fonctionnelles données dans la littérature (Birney et al, 1995 ; Sagan et al., 2001). Dès lors nous espérons à moyen terme être capables de prédire l'effet biologique d'autres peptides portant ce type de modifications, à partir de la seule séquence.

### Articles et brevets publiés durant cette thèse.

- Membrane Topology of the Escherichia coli AmpG Permease Required for Recycling of Cell Wall Anhydromuropeptides and AmpC  $\beta$ -Lactamase Induction. - Chahboune A., Decaffmeyer M., Basseur R. and Joris B. - *Antimicrobial Agents and Chemotherapy* Mar.2005 1145-49
- TP0453, a concealed outer membrane protein of Treponema pallidum, enhances membrane permeability. - Hazlett KR, Cox DL, Decaffmeyer M, Bennett MP, Desrosiers DC, La Vake CJ, La Vake ME, Bourell KW, Robinson EJ, Basseur R, Radolf JD. - *J Bacteriol.* 2005 Sep;187(18):6499-508.
- Rational design of complementary peptides to the  $\beta$ Amyloid 29–42 fusion peptide: An application of PepDesign. - Decaffmeyer M., Lins L., Charlotheaux B., VanEyck MH, Thomas A., Basseur R. - *BBA Biomembranes* 2006 mar, 1758(3) 320-7
- Brevet n° PCT/EP2005/055050. Complementary Peptides for  $\beta$ -Amyloid 29-42 Peptide. (WO/2006/037800)13/04/2006. Heinen E., Decaffmeyer M; and Basseur R. [http://www.wipo.int/pctdb/en/wo.jsp?LANGUAGE=ENG&KEY=06/037800&ELEMENT\\_SET=F](http://www.wipo.int/pctdb/en/wo.jsp?LANGUAGE=ENG&KEY=06/037800&ELEMENT_SET=F)
- Prediction of Peptide Structure: how far are we? – Thomas A., Deshayes S., Decaffmeyer M., VanEyck MH., Charlotheaux B. and Basseur R. – *Proteins.* 2006 Dec 1;65(4):889-97.
- Nouveaux outils pour la prediction de structures peptidiques. Van Eyck MH, Decaffmeyer M, Basseur R. *Spectra Analyse* 2007 sept 257:25-29I
- PepLook: An innovative in silico tool for determination of structure, polymorphism and stability of peptides (2007).Annick Thomas, Sébastien Deshayes, Marc Decaffmeyer, Marie-HélèneVan Eyck, Benoit Charlotheaux B, Robert Basseur *Biopolymers: 88 (4): 532-532*
- Les Médicaments Peptidiques: Mythe ou Réalité? Marc Decaffmeyer , Annick Thomas, Robert Basseur. *BASE (Accepted, In Press)*
- Molecular determinants of peptide/membrane interactions: “in silico” and biophysical approaches (Review) Lins L., Decaffmeyer M., Thomas A., Basseur R. (In revision)
- Predicted conformation of the hydrophobic segment of mammalian diacylglycerol kinase epsilon and its topology in a cell membrane. Marc Decaffmeyer, Armela Dicu, Annick Thomas, Matthew K. Topham, Robert Basseur and Richard M. Epan. (In preparation)
- Structural Polymorphism of two CPP: An important parameter of activity. Deshayes S, Decaffmeyer M, Basseur R and Thomas A. (Accepted at BBA on 28/1/2008)

# **Annexe**

Article de synthèse publié dans la revue BASE

# Les médicaments peptidiques : mythe ou réalité ?

Marc Decaffmeyer, Annick Thomas, Robert Brasseur

Gembloux Agricultural University – FUSAGx. Centre de Biophysique moléculaire numérique. Passage des Déportés, 2. B-5030 Gembloux (Belgique). E-mail : decaffmeyer.m@fsagx.ac.be

Reçu le 15 mai 2007, accepté le 14 septembre 2007.

Les peptides trouvent de nombreux débouchés d'application dans le domaine biotechnologique, parmi lesquels des applications pharmaceutiques. La progression du nombre de nouvelles petites molécules déposées à la FDA comme substances médicamenteuses s'essouffle. Dans ce contexte, de grosses sociétés pharmaceutiques investissent dans la recherche peptidique pour ouvrir de nouvelles perspectives thérapeutiques. Même s'ils sont utilisés comme agents thérapeutiques depuis près d'un siècle sous leur forme naturelle, l'utilisation des peptides reste parcimonieuse bien qu'elle ait connu un développement significatif depuis quelques dizaines d'années, notamment grâce à la mise au point de méthodes de production, chimiques en phase solide ou biologiques telles qu'en *phage display*. Les peptides présentent nombre d'avantages par rapport aux médicaments traditionnels que sont les petites molécules, toutefois leur utilisation pharmaceutique se heurte à des limites, qu'elles soient liées à la faible capacité des peptides à traverser les barrières membranaires, à leur faible durée de demi-vie ou à la difficulté de comprendre leurs mécanismes d'action. Des solutions peuvent être apportées pour surmonter ces problèmes. La modélisation moléculaire a un rôle important à jouer dans ce contexte, que ce soit pour la prédiction *in silico* de la structure, de la stabilité et de la variété structurale des peptides ou pour l'étude et l'amélioration de leur capacité à lier une cible.

**Mots-clés.** Peptide, modélisation, structure, *in silico*, médicaments.

**Peptides as drugs: myth or reality?** Peptides find many applications in the fields of biotechnology, among which pharmaceutical applications. Progression of the number of traditional drugs deposited at the FDA is slowing down. It's the reason why many big pharmaceutical companies invest more and more in peptides to open new therapeutic perspectives. Even if some peptides have been used as therapeutic agents in their natural bioactive form for nearly one century, a significant development of the number of peptides has occurred since a few tenths of years, in particular due to the development of methods of production, chemical as solid phase as well as biological as *phage display*. Peptides have a number of advantages when compared to traditional drugs like small molecules, however their development encounters some limitations, which are relative to a low ability to cross membrane barriers and a low half-life time and a difficulty to understand their mechanism of action. A series of solutions can be brought to overcome these problems. Molecular modelling has an important role to play in this context, notably for the *in silico* determination of the structure, the stability and the polymorphism of peptides or for the study and the improvement of the capacity to bind a target.

**Keywords.** Peptide, design, structure, *in silico*, drugs.

## 1. INTRODUCTION

Les peptides sont utilisés depuis de nombreuses années dans le domaine de la biotechnologie, par exemple pour la production d'anticorps à l'usage des laboratoires de biologie. Ces molécules seront plus que probablement intensivement utilisées dans un futur proche pour d'autres applications et plus particulièrement dans le domaine de l'industrie pharmaceutique. On pense notamment à la conception de peptides régulateurs d'activité protéique. L'utilisation de peptides antimicrobiens (AMPs) est promise à un grand avenir, notamment afin de contrer la perte d'efficacité des

antibiotiques classiques (Gordon et al., 2005 ; Marr et al., 2006). Les peptides vecteurs (CPPs) constituent une autre possibilité d'utilisation thérapeutique. Ils sont capables de faire pénétrer des composés thérapeutiques aussi variés que des drogues, des siRNA ou d'autres peptides dans des cellules cibles (Deshayes et al., 2005 ; Mae et al., 2006). Parmi les méthodes d'avenir, pour aider à la conception de peptides performants, on compte la modélisation moléculaire.

Plusieurs programmes de modélisation *in silico*, principalement des méthodes de *docking*, existent pour la conception de petites molécules, telles que Dock, AutoDock, FlexX et autres méthodes similaires

(Ghosh et al., 2006). A contrario, et en dépit de l'avenir prometteur des peptides, leur modélisation *in silico* est longtemps restée une voie peu active. Nous verrons toutefois que quelques méthodes ont été développées récemment.

Enfin, si la fonctionnalité des peptides dépend de leur séquence et donc des fonctions chimiques qu'ils portent, elle est également grandement modulée par leur structure. Les méthodes de détermination expérimentales sont souvent lentes et coûteuses, les problèmes techniques qui y sont liés sont nombreux et les résultats qu'elles donnent sont souvent divergents. Jusque récemment, il n'existait pas de méthode satisfaisante de détermination *in silico* de la structure des peptides.

## 2. PEPTIDES ET DÉRIVÉS PEPTIDIQUES : MÉDICAMENTS DE DEMAIN ?

Le médicament idéal est un mythe : il s'agirait d'une molécule de faible poids moléculaire, soluble, capable de mimer chimiquement et structurellement de manière quasi parfaite un ligand naturel sans en avoir les effets, d'être éliminée rapidement de l'organisme après son action mais sans induire d'effet secondaire. En outre, sa production devrait être facile et peu coûteuse (Edwards et al., 1999). Finalement, son administration au patient devrait être aussi aisée que possible.

La chimie du médicament est basée depuis toujours sur l'exploitation et la transformation de substances médicinales naturelles, notamment de produits issus des plantes, copiant en cela l'utilisation qu'en faisaient nos ancêtres. Cette source s'épuise. Nous en voulons pour preuve l'épuisement des pipelines des laboratoires pharmaceutiques décrit par les analystes financiers américains comme le *pipeline problem*. Ceci se traduit par le fait que de nombreux brevets d'exploitation arrivent en fin de vie alors qu'un nombre de plus en plus limité de molécules nouvelles sont présentées à la Food and Drug Administration (Surowiecki, 2004). Un très récent rapport du Congrès Américain indique que le rapport entre le nombre des nouvelles molécules actives déposées et les budgets R&D des firmes décroît de manière significative depuis plus de 20 ans (Howlett, 2006). Ceci indique que l'innovation en la matière est de plus en plus ardue et coûteuse. En effet, le développement d'une nouvelle molécule thérapeutique coûte entre 200 et 800 millions d'euros (Anon., 2003 ; Goozner, 2004) et prend environ 10 ans.

Un moyen de plus en plus envisagé par les laboratoires pharmaceutiques pour proposer des alternatives thérapeutiques originales est de concevoir des médicaments de type peptidique ou peptidomimétique.

On peut classer les peptides thérapeutiques en trois catégories (Sato et al., 2006) : premièrement, les

peptides naturels, aussi appelés peptides bioactifs qui sont soit produits par l'organisme, soit dérivés de protéines (Watt, 2006) ; deuxièmement, les peptides issus de librairies génétiques par *phage display* (Sergeeva et al., 2006) ; enfin, les peptides issus de librairies chimiques produites par synthèse en phase solide (Shin et al., 2005). On peut également citer les peptides vecteurs qui ne sont pas à proprement parler des molécules médicamenteuses mais qui servent à transporter des molécules thérapeutiques (Mae et al., 2006). Parmi ces peptides transporteurs, on retrouve par exemple la Pénétratine dérivée de l'homéodomaine d'Antennapédia, conçue pour la délivrance de protéines bioactives (Derossi et al., 1994) ou le Transportan, peptide chimérique issu de la Galanine et du Mastoparan, conçu pour transporter des ARN interférants (Pooga et al., 1998), etc. Finalement, une dernière voie de recherche mettant en œuvre des peptides est celle qui les utilise non pas comme produits finis mais comme outils pour concevoir des peptidomimétiques, petites molécules qui miment les propriétés bioactives du peptide (Patch et al., 2002).

On peut dater la découverte du premier peptide bioactif à plus de 80 ans. C'est en effet en 1923 que Banting et Macleod obtinrent le Prix Nobel de Médecine pour leur découverte de l'insuline dont on connaît le rôle dans le traitement du diabète. À cette époque, et jusque dans les années soixante, la synthèse de peptides par voie chimique classique prenait jusqu'à plusieurs mois, rendant impossible toute exploitation industrielle. Ce n'est qu'à partir de 1963, année du développement par Merrifield de la méthode de synthèse peptidique par voie solide (Merrifield, 1963) pour lequel il reçut le Prix Nobel de Chimie en 1984, et aussi grâce à la mise au point de techniques de séparation/purification performantes telles que l'HPLC (Chromatographie Liquide Haute Performance) (Chen et al., 1995) ou la SPE (extraction de phase solide) (Kamysz et al., 2004), que l'utilisation industrielle de peptides non issus du milieu naturel est devenue envisageable. D'autre part, depuis une vingtaine d'années, la mise au point de la technique du *phage display* (Smith, 1985) puis les progrès dans la maîtrise de cette méthode ont permis la production de librairies de peptides performantes menant à la sélection de peptides bioactifs (Cortese et al., 1996).

Depuis ces percées technologiques, les peptides, au même titre que les protéines, sont considérés comme des molécules thérapeutiques d'avenir. En 2004, plus de 20 % des médicaments appartenant au top 200 des ventes étaient à base de protéines ou de peptides, avec des chiffres d'affaires atteignant les 40 milliards \$ (McGee, 2005), soit environ 10 % du chiffre global de l'industrie pharmaceutique. En 2004, entre 600 et 700 peptides étaient en stade de développement et plus de 150 étaient à différents stades de phase



clinique (McGee, 2005). On les trouve dans différents domaines tels que le traitement de certaines formes de cancers, du SIDA, de l'ostéoporose, de maladies neuro-dégénératives (Lien et al., 2003).

Les peptides offrent plusieurs avantages par rapport aux petites molécules qui constituent les médicaments traditionnels. Le premier avantage est que, représentant souvent la plus petite partie fonctionnelle d'une protéine, ils offrent une efficacité, une sélectivité et une spécificité que les concepteurs de petites molécules ont du mal à atteindre (Hummel et al., 2006). À titre d'exemple, on peut citer les travaux récemment menés sur NAP, un peptide de 8 résidus, le plus petit fragment actif de ADNP (*Activity Dependent Neuroprotective Protein*) qui montre une activité neuroprotectrice (Gozes et al., 2006). Ce peptide est actuellement en phase d'essais cliniques.

Un deuxième avantage, et non des moindres, est que les produits de dégradation des peptides sont des acides aminés, ce qui élimine grandement les risques de toxicité (Loffet, 2002).

Troisièmement, les peptides s'accumulent peu dans les tissus en raison de leur temps de demi-vie court, ce qui, il est vrai, constitue aussi un désavantage quand il est question de durée d'action.

La conception de peptides à usage thérapeutique reste toutefois extrêmement compliquée et soulève un grand nombre de challenges. Premièrement, le coût de production est plus élevé que celui des petites molécules et les quantités nécessaires sont plus importantes (Marx, 2005), principalement lorsque la voie d'administration au patient n'est pas orale. Deuxièmement, les peptides n'obéissent pas aux fameuses règles des cinq (Lipinski et al., 2001), notamment concernant leur poids moléculaire. Selon ces règles, les peptides seraient incapables de passer du tractus digestif vers le système circulatoire. Donc, si l'on s'en tient à ce critère classiquement utilisé dans l'industrie pharmaceutique, les peptides sont considérés comme étant peu absorbables par voie orale. Pour cette raison, l'administration des peptides thérapeutiques est souvent intraveineuse, ce qui provoque inévitablement un manque de confort pour le patient. Toutefois, des stratégies alternatives de délivrance existent, faisant notamment appel à l'administration via les muqueuses (Prego et al., 2005), nasales (Maggio, 2005) ou pulmonaires (Skyler et al., 2001) par exemple. D'autres groupes ont mis au point des enrobages permettant une libération retardée du peptide et donc l'espacement des injections. C'est le cas de la société Roche qui a mis sur le marché le Pegasys<sup>TM</sup> qui est une formulation pegylée d'Interféron qui permet de limiter sensiblement le nombre d'injections. D'autres encore, telle que la société Bioject, ont développé des propulseurs à gaz à haute pression permettant au peptide de traverser la peau sans requérir à une injection (Stout et al., 2007).

Enfin, la technologie eligen<sup>R</sup> développée par la société Emisphere permet d'enrober le peptide pour lui permettre un transfert du tube digestif vers le sang, autorisant une prise orale (Malkov et al., 2005). Finalement, les stratégies faisant appel à des transporteurs (*carrier peptides*) sont également en plein développement.

Le troisième aspect qu'il faut aborder ici est lié à la biodégradabilité des peptides due aux peptidases du système gastro-intestinal, des reins, du foie et du sérum. Cette dégradation, qui a pour conséquence une durée de demi-vie faible, de l'ordre de quelques minutes voire au mieux de quelques heures, peut être contrée de différentes manières. Une des voies les plus utilisées est celle ayant recours à la modification des parties C et N terminales (par N-acétylation et N-amidation, par exemple) des peptides pour empêcher l'action des exopeptidases. Est aussi envisagée l'utilisation d'acides aminés D ou non-naturels, lesquels sont moins sujets à l'action de ces enzymes (Fauchère et al., 1992). Les cyclisations, que ce soit par des ponts disulfures, la formation d'anneaux lactames ou des liaisons Nter-Cter, sont également utilisées pour contrecarrer l'action des protéases (Werle et al., 2006). Enfin la Pegylation, c'est-à-dire la fixation d'unités de poly(éthylène glycol), est une stratégie de plus en plus utilisée (Caliceti et al., 2003 ; Harris et al., 2003 ; Veronese et al., 2005). Certains groupes ont par ailleurs montré qu'il est possible de conserver la fonction d'un peptide alors qu'il est partiellement composé d'acides aminés D (Benkirane et al., 1993) ou même qu'il est possible de modéliser un peptide de fonction donnée avec des acides aminés non-naturels (Lins et al., 2006).

Ces transformations, on le voit, mènent à la conception de molécules qui s'apparentent de moins en moins à des peptides et de plus en plus à des petites molécules qui portent les fonctions chimiques essentielles mises en lumière par l'utilisation des peptides. On parlera de peptidomimétiques ou de pseudo-peptides. Pour ces produits, le peptide n'est pas une finalité mais bien un modèle de conception. Classiquement, la conception d'un peptidomimétique se déroule en plusieurs étapes :

- identification de la protéine cible
- identification de l'effecteur (ligand, peptide bioactif ou protéine partenaire)
- recherche du segment bioactif le plus court
- utilisation d'analogues peptidiques structurellement contraints en méthylant le C $\alpha$  (Sagan et al., 2001) ou le C $\beta$  (Birney et al., 1995), voire en le cyclisant
- établissement d'un pharmacophore et finalement
- criblage de banques de petites molécules respectant les propriétés chimico-spatiales du pharmacophore (Hummel et al., 2006).

### 3. LA MODÉLISATION AU SERVICE DE LA CONCEPTION DE PEPTIDES

La modélisation *in silico* de peptides à visée thérapeutique peut se diviser en deux voies : basée sur le ligand ou basée sur la structure de la cible. La modélisation basée sur la structure de la cible nécessite évidemment d'identifier cette cible puis d'en déterminer la structure, ensuite d'identifier le site d'interaction, l'idéal étant de réaliser une co-cristallisation de la cible avec son ligand naturel, finalement de cribler des banques de données à la recherche de molécules s'adaptant au mieux chimiquement et structurellement au site de liaison (Anderson, 2003). La modélisation basée sur le ligand est parfois une voie obligatoire, c'est notamment le cas lorsque la structure de la cible est inaccessible. Beaucoup de récepteurs importants sont des protéines membranaires qui sont notoirement difficiles à cristalliser. Dans ce cas, il est nécessaire de disposer de la structure d'un ligand ou d'une protéine interagissant avec la cible, on parlera de *lead*. Ensuite, il s'agira de transposer cette structure en un pharmacophore (empreinte chimico-spatiale du peptide actif) puis de réaliser un criblage virtuel de banques de peptides ou de petites molécules donnant un bon score QSAR (relation quantitative activité/structure) (Parvu, 2003) ou une bonne réponse fonctionnelle. Dans les deux cas, ces techniques sont principalement adaptées au criblage de molécules rigides.

### 4. ASPECT DESIGN DE LA SÉQUENCE

Nous avons vu que des méthodes expérimentales de design de peptides existent par criblages de banques fournies par des phages ou par synthèse en phase solide. Ces méthodes sont longues et onéreuses.

Par contre, un nombre réduit de méthodes de modélisation de peptides *in silico* existent actuellement. Parmi ces méthodes, on peut parler de l'algorithme MIMETIC qui a été utilisé notamment pour la conception d'un peptide complémentaire à la Thrombomodulin en vue de moduler la fonction de la Procarboxypeptidase R (Shimomura et al., 2003) ou d'un peptide complémentaire à des segments de la transcriptase inverse de HIV-1 (Campbell et al., 2002). Dans cette méthode, des peptides sont générés de manière itérative et un score tenant compte de propriétés physico-chimiques permet de les classer.

La deuxième famille de méthodes de design *in silico* de peptides fait appel à la dynamique moléculaire comme cela a été fait par exemple pour la conception d'un peptide antagoniste de l'interleukin-6 humaine (hIL-6) (Yang et al., 2005). Cette méthode requiert malheureusement de très longs temps de calcul et ne permet dès lors de tester qu'un nombre très restreint de peptides.

Par ailleurs, les méthodes de *docking*, généralement utilisées pour le design de petites molécules (Dock (Kuntz et al., 1982), AutoDock (Goodsell et al., 1990)), sont mal adaptées au design de peptides.

La méthode de modélisation *in silico* PepDesign, la plus récente en date, permet la conception rapide de peptides capables de se lier à une molécule cible. La conception d'un peptide anti-AB29-42, peptide impliqué dans la formation des plaques séniles caractéristiques de la maladie d'Alzheimer, a été réalisée en utilisant cette méthode (Decaffmeyer et al., 2006). Ces peptides, ainsi que tous leurs dérivés, ont été protégés par le dépôt d'un brevet (Heinen et al., 1996) pour toute application thérapeutique ou à visée diagnostique de la maladie d'Alzheimer. Cette méthode de design est applicable pour la modélisation d'un peptide ciblant un peptide, une protéine ou toute molécule dont on pourrait connaître la structure tridimensionnelle. Elle pourrait être utilisée par exemple pour concevoir un peptide qui perturberait l'interface entre deux protéines et donc empêcherait la formation d'un complexe protéine/protéine impliqué dans le développement d'une maladie. Utilisée sur des acides aminés naturels, elle est en outre également adaptée à la modélisation de peptides composés d'acides aminés transformés (D, non naturels, etc.). Ceci pourrait permettre dans le futur de tester l'influence des modifications requises, par exemple pour contrecarrer les problèmes de digestion enzymatique, sur l'efficacité de reconnaissance et de liaison à la cible.

### 5. ASPECT STRUCTURAL

Nous savons que la structure du peptide est cruciale, notamment parce qu'elle doit impérativement correspondre à une conformation active en présence de la cible. D'autre part, il paraît essentiel d'être capable de vérifier si les transformations (modification du squelette peptidique ou des chaînes latérales) laisseront au pseudo-peptide la possibilité d'adopter la conformation requise pour sa fonction.

À ce jour, plusieurs méthodes expérimentales de détermination de structure des peptides existent, au nombre desquelles on compte le dichroïsme circulaire (DC) (Holzwarth et al., 1965), la spectroscopie infrarouge à transformée de Fourier (FT-IR) (Goormaghtigh et al., 1990), la cristallographie aux rayons X ou encore la résonance magnétique nucléaire (RMN) (Bechinger et al., 1999). Ces techniques sont souvent lourdes à mettre en œuvre et sont également confrontées à différentes limitations par exemple de cristallisation, d'agrégation, etc. D'autre part, la réalité veut que les données relatives à la structure de peptides fournies à partir de ces différentes techniques divergent souvent.

Ce constat indique que l'idée reçue selon laquelle une séquence donne naissance à une structure unique dans un solvant donné, si elle est cohérente pour la plupart des protéines, ne trouve pas forcément à s'appliquer aux peptides. Ceci est notamment dû au fait qu'en comparaison des protéines, les peptides disposent d'un nombre plus restreint de possibilités de stabilisation intramoléculaire et également qu'une même séquence peptidique peut mener à plusieurs solutions structurales énergétiquement comparables, donc probables.

Si différentes méthodes expérimentales de détermination des structures peptidiques existent, à notre connaissance seules trois méthodes de détermination *in silico* sont accessibles à ce jour. La première méthode de détermination *in silico* de structure est appelée ROSETTA et a été développée à la fin des années 1990 par une équipe américaine (Simons et al., 1997). Elle est basée sur l'utilisation mixte de la modélisation par homologie à des structures connues de segments protéiques pour la détermination de la conformation locale et de prédiction *de novo* en utilisant un champ de force basé sur des potentiels moyens (MFP). Cette méthode donne dix structures dont la plus probable porte le n° 10. Le serveur Robetta qui utilise la procédure Rosetta est accessible sur le net (Chivian et al., 2003). ROSETTA a été beaucoup utilisée pour la prédiction de la structure tridimensionnelle de protéines (Bonneau et al., 2002 ; Bradley et al., 2003), et ce avec de beaux succès notamment à l'occasion de diverses éditions du CASP (Critical Assessment of techniques for protein Structure Prediction). Elle a aussi été utilisée avec succès pour construire *de novo* des protéines capables d'adopter un repliement préalablement choisi, le cas le plus connu étant celui de Top7, une protéine de 93 résidus comportant deux hélices  $\alpha$  et cinq brins  $\beta$  qui a été modélisée par l'équipe de David Baker (Kuhlman et al., 2003).

Malheureusement, si cette méthode a fait ses preuves pour les protéines, elle n'est pas adaptée aux peptides, probablement en raison du champ de force utilisé, basé sur les MFP, c'est-à-dire sur une analyse statistique des structures de protéines. Ce champ de force a en effet tendance à replier les peptides comme des protéines et conduit à des structures globulaires (collapse hydrophobe) qui ne concordent pas avec les données obtenues expérimentalement (Thomas et al., 2006). Quoi qu'il en soit, le serveur Robetta ne permet pas la prédiction sur des séquences d'une longueur inférieure à 20 résidus, ce qui indique clairement que la méthode est plus destinée à l'étude de protéines que de peptides.

La deuxième méthode connue de détermination *ab initio* de la structure des peptides avait été développée précédemment par un groupe indien. Cette approche, appelée PEPSTR, est basée dans un premier temps sur la prédiction de la structure secondaire des

résidus par la méthode BETATURNS (Kaur et al., 2004). Ensuite, des angles  $\Phi\Psi$  standard sont assignés aux résidus avant une minimisation énergétique utilisant le champ de force Amber6. La méthode permet de traiter des peptides de 7 à 25 résidus. Elle donne une structure unique. À notre connaissance, cette méthode n'a pas été publiée et n'a été citée qu'une fois dans la littérature, même si le serveur est disponible en ligne ([www.imtech.res.in/raghava/pepstr/](http://www.imtech.res.in/raghava/pepstr/)).

La dernière méthode en date a été développée par la société Biosiris S.A. en collaboration avec l'équipe de R. Brasseur à Gembloux (Thomas et al., 2006). Cette méthode appelée PepLook est spécialement dévolue à la prédiction *in silico* de peptides et n'est pas à ce jour adaptée à la prédiction de structures protéiques. Elle est basée sur l'exploration de l'espace conformationnel du peptide en utilisant un set de 64 couples d'angles  $\Phi\Psi$  qui permet la reconstruction de n'importe quelle protéine de structure connue (Etchebest et al., 2005). Une procédure itérative de type stochastique telle que celle utilisée dans de nombreux domaines scientifiques (Glick et al., 2002) est employée et est en outre modulée par une approche de type Boltzmann qui permet, étape après étape, de moduler la probabilité d'utilisation des différents couples d'angles  $\Phi/\Psi$  pour chaque position en fonction de leur contribution à l'occurrence de structure de bonne ou de mauvaise énergie. Cette méthode donne la structure la plus stable d'un point de vue énergétique, appelée le Prime, mais aussi 98 autres structures de basse énergie.

Cette méthode est originale en ce que, outre le fait de donner la structure la plus stable, elle donne accès à des paramètres sur la stabilité du peptide, sur sa propension à lier des partenaires externes, ainsi que sur son polymorphisme (ou désordre ou diversité structurale). En effet, un score de stabilité est établi en comparant le MFP (Mean Force Potential) de chaque résidu du Prime à des valeurs de référence calculées à partir d'une banque non-redondante de près de 500 structures de protéines repliées (Lovell et al., 2003), ce qui permet également d'identifier les acides aminés les plus enclins à créer des interactions extramoléculaires et donc à chercher des partenaires afin d'augmenter leur stabilité. Enfin, un score de polymorphisme (ou de désordre) est également établi par le calcul le long de la séquence d'un RMSd[9] (Root Mean Square deviation sur une fenêtre de neuf résidus) de chacune des 98 structures de meilleure énergie par rapport au Prime.

Cette dernière donnée est loin d'être anecdotique si l'on veut bien concevoir qu'un peptide donné doit forcément être capable d'adapter sa structure pour remplir plusieurs tâches différentes. Si on prend le cas d'un peptide capable de passer une membrane plasmique pour remplir sa fonction, il doit simultanément être soluble dans un milieu hydrophile (le sang et le cyto-

plasme), dans un milieu hydrophobe (la membrane) et enfin probablement s'organiser dans une troisième structure en présence de sa cible. Il ne peut présenter une solubilité adéquate dans différents milieux que si son hydrophobicité apparente change, et donc si sa structure change. Le polymorphisme est donc plus que probablement un facteur déterminant pour la fonctionnalité biologique des peptides, et à ce titre il semble important d'avoir accès à cette information.

Cette hypothèse a été proposée pour les peptides vecteurs. Les peptides vecteurs fonctionnels connus à ce jour disposent de caractéristiques communes telles qu'une taille maximale de 30 résidus, une certaine amphipaticité et une charge nette positive. Toutefois, leur mode de fonctionnement et plus particulièrement la manière dont ils sont capables de traverser des membranes biologiques, même si elle a été très largement étudiée par des méthodes biophysiques (Fischer et al., 2005), n'est pas encore élucidée. S'il est établi que l'internalisation des peptides est conditionnée par leur amphiphilicité et les charges qu'ils portent, elle est également largement dépendante de leur structure (Deshayes et al., 2005). PepLook a été récemment utilisé pour étudier les propriétés de deux de ces peptides et de leurs mutants. Ces travaux montrent clairement que le polymorphisme structural est une caractéristique essentielle pour la pénétration cellulaire puisqu'il permet au peptide d'adapter sa conformation à son environnement (Deshayes et al., soumis). Ce constat n'est pas anodin et pourrait être à la base de nouvelles méthodes de design de peptides vecteurs.

Enfin, la même méthode a été également utilisée dans un autre domaine pour prédire l'organisation du segment membranaire de la DGK $\epsilon$  humaine (D19-Q23), protéine membranaire monotopique de structure non connue et impliquée dans la transformation du diacylglycerol en acide phosphatidique.

Finalement, dans un avenir proche, la méthode sera utilisée pour vérifier l'effet des modifications chimiques sur la structure des peptides.

## 6. CONCLUSION

De nombreux groupes pharmaceutiques comptent sur le développement de peptides, voire de leur utilisation dans la conception de petites molécules, comme nouvelles voies thérapeutiques ou de recherches thérapeutiques. Ceci a été rendu possible grâce aux avancées technologiques des dernières décennies. Comparés aux petites molécules, les peptides présentent des qualités indéniables en termes d'activité notamment. Toutefois, l'utilisation des peptides en médecine se heurte à différents problèmes liés à leur nature chimique ou biologique, principalement dus à leur sensibilité aux dégradations biologiques et à leur faible capacité à

traverser les barrières membranaires. De nombreux groupes ont mis au point des stratégies permettant soit de nouvelles voies d'administration non orales, soit des mécanismes de protection pour rendre les peptides moins bio-vulnérables. La conception de peptides vecteurs pour le transport de substances thérapeutiques quelles qu'elles soient est également dans l'air du temps. Enfin, depuis peu, de nouveaux outils de modélisation *in silico* voient le jour. Ils viendront en appui des méthodes expérimentales, notamment pour le design de peptides mais aussi pour la gestion et la prédiction de leurs structures, leurs sélectivités, leurs affinités.

## Remerciements

Marc Decaffmeyer a été supporté financièrement par le Ministère de la Région Wallonne (DGTRE : contrat PEPSEIN). Annick Thomas est Directeur de recherche à l'INSERM (France) et Robert Brasseur est Directeur de recherche au FNRS.

## Bibliographie

- Anon., 2003. Coût de recherche et développement du médicament : la grande illusion. *Prescrire*, **23**(244), 782-787.
- Anderson A.C., 2003. The process of structure-based drug design. *Chem. Biol.*, **10**, 787-797.
- Bechinger B. et al., 1999. Peptide structural analysis by solid-state NMR spectroscopy. *Biopolymers*, **51**(3), 174-190. Review.
- Benkirane N. et al., 1993. Antigenicity and immunogenicity of modified synthetic peptides containing D-amino acid residues. Antibodies to a D-enantiomer do recognize the parent L-hexapeptide and reciprocally. *J. Biol. Chem.*, **268**(35), 26279-26285.
- Birney D.M. et al., 1995. Use of beta-methylphenylalanine (beta MeF) residues to probe the nature of the interaction of substance P with its receptor: effects of beta MeF-containing substance P analogs on rabbit iris smooth muscle contraction. *J. Medic. Chem.*, **38**(13), 2478-2482.
- Bonneau R. et al., 2002. *De Novo* prediction of three-dimensional structures for major protein families. *J. Mol. Biol.*, **322**(1), 65.
- Bradley P. et al., 2003. Rosetta predictions in CASP5: successes, failures, and prospects for complete automation. *Proteins*, **53**, 457-468.
- Caliceti P. & Veronese F.M., 2003. Pharmacokinetic and biodistribution properties of poly(ethylene glycol)-protein conjugates. *Adv. Drug Delivery Rev.*, **55**(10), 1261-1277.
- Campbell W. et al., 2002. A novel genetic algorithm for designing mimetic peptides that interfere with the function of a target molecule. *Microbiol. Immunol.*, **46**(3), 211-215.

- Chen H. & Horvath C.S., 1995. High-speed high-performance liquid chromatography of peptides and proteins. *J. Chrom. A*, **705**, 3-20.
- Chivian D. et al., 2003. Automated prediction of CASP-5 structures using the Robetta server. *Proteins*, **53**(Suppl. 6), 524-533, <http://robetta.bakerlab.org/>, (15/12/2005)
- Cortese R. et al., 1996. Selection of biologically active peptides by phage display of random peptide libraries. *Curr. Opin. Biotechnol.*, **7**(6), 616-621.
- Decaffmeyer M. et al., 2006. Rational design of complementary peptides to the betaAmyloid 29-42 fusion peptide: an application of PepDesign. *Biochim. Biophys. Acta*, **1758**(3), 320-327.
- Derossi D., Joliot A.H., Chassaing G. & Prochiantz A., 1994. The third helix of the Antennapedia homeodomain translocates through biological membranes. *J. Biol. Chem.*, **269**(14), 10444-10450.
- Deshayes S., Morris M.C., Divita G. & Heitz F., 2005. Cell-penetrating peptides: tools for intracellular delivery of therapeutics. *Cell. Mol. Life Sci.*, **62**(16), 1839-1849. Review.
- Deshayes S., Decaffmeyer M., Brasseur R. & Thomas A., 2008. Structural polymorphism, a parameter required for cellular uptake of peptides. Submitted.
- Edwards C.M.B., Cohen M.A. & Bloom S.R., 1999. Editorial: Peptides as drugs. *Q. J. Med.*, **92**, 1-4.
- Etchebest C., Benros C., Hazout S. & de Brevern A.G., 2005. A structural alphabet for local protein structures: improved prediction methods. *Proteins*, **59**(4), 810-827.
- Fauchere J.L. & Thurieau C., 1992. Evaluation of the stability of peptides and pseudopeptides as a tool in peptide drug design. *Adv. Drug Res.*, **23**, 127-159.
- Fischer R., Fotin-Mleczek M., Hufnagel H. & Brock R., 2005. Break on through to the other side-biophysics and cell biology shed light on cell-penetrating peptides. *Chem. Biochem.*, **6**(12), 2126-2142. Review.
- Ghosh S., Nie A., An J. & Huang Z., 2006. Structure-based virtual screening of chemical libraries for drug discovery. *Curr. Opin. Chem. Biol.*, **10**(3), 194-202. Review.
- Glick M., Rayan A. & Goldblum A., 2002. A stochastic algorithm for global optimization and for best populations: A test case of side chains in proteins. *Proc. Natl Acad. Sci. USA*, **99**, 703-708.
- Goodsell D.S. & Olson A.J., 1990. Automated docking of substrates to proteins by simulated annealing. *Proteins*, **8**(3), 195-202.
- Goormaghtigh E., Cabiaux V. & Ruyschaert J.M., 1990. Secondary structure and dosage of soluble and membrane proteins by attenuated total reflection Fourier-transform infrared spectroscopy on hydrated films. *Eur. J. Biochem.*, **193**(2), 409-420.
- Goozner M., 2004. *The \$800 million pill. The truth behind the cost of new drugs*. Berkeley, USA: University of California Press.
- Gordon Y.J., Romanowski E.G. & McDermott A.M., 2005. A review of antimicrobial peptides and their therapeutic potential as anti-infective drugs. *Curr. Eye Res.*, **30**(7), 505-515. Review.
- Gozes I. & Spivak-Pohis I., 2006. Neurotrophic effects of the peptide NAP: a novel neuroprotective drug candidate. *Curr. Alzheimer Res.*, **3**(3), 197-199. Review.
- Harris J.M. & Chess R.B., 2003. Effect of pegylation on pharmaceuticals. *Nat. Rev. Drug Discovery*, **2**(3), 214-221.
- Heinen E., Decaffmeyer M. & Brasseur R., 1996. *Complementary peptides for  $\beta$ -Amyloid 29-42 peptide*. Patent PCT/EP2005/055050.
- Holzwarth G. & Doty P., 1965. The ultraviolet circular dichroism of polypeptides. *J. Am. Chem. Soc.*, **87**, 218-228.
- Howlett C., 2006. *A CBO (Congressional Budget Office) Study: Research and development in the pharmaceutical industry*. Washington: Congress of the United States.
- Hummel G., Reineke U. & Reimer U., 2006. Translating peptides into small molecules. *Mol. Biosyst.*, **2**, 499-508.
- Kamysz W. et al., 2004. Fast and efficient purification of synthetic peptides by solid-phase extraction. *Acta Chrom.*, **14**, 180-186.
- Kaur H. & Raghava G.P., 2004. Neural network method for prediction of beta-turn types in proteins using evolutionary information. *Bioinformatics*, **20**(12), 2751-2758.
- Kuhlman B. et al., 2003. Design of a novel globular protein fold with atomic-level accuracy. *Science*, **302**(5649), 1364-1368.
- Kuntz I.D. et al., 1982. A geometric approach to macromolecule-ligand interactions. *J. Mol. Biol.*, **161**(2), 269-288.
- Lien S. & Lowman H.B., 2003. Therapeutic peptides. *Trends Biotechnol.*, **21**(12), 556-562.
- Lins L. et al., 2006. *De novo* design of peptides with specific lipid-binding properties. *Biophys. J.*, **90**(2), 470-479.
- Lipinski C.A., Lombardo F., Dominy B.W. & Feeney P.J., 2001. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development setting. *Adv. Drug Delivery Rev.*, **46**, 3.
- Loffet A., 2002. Peptides as drugs: is there a market? *J. Peptide Sci.*, **8**(1), 1-7.
- Lovell S.C. et al., 2003. Structure validation by Calpha geometry: phi, psi and Cbeta deviation. *Proteins*, **50**(3), 437-450.
- Mae M. & Langel U., 2006. Cell-penetrating peptides as vectors for peptide, protein and oligonucleotide delivery. *Curr. Opin. Pharmacol.*, **6**(5), 509-514. Review.
- Maggio E.T., 2005. Recent developments in intranasal drug delivery technology are creating new vistas for peptide and protein therapeutics. *Drug Delivery Companies Report*, **13**, 29-33.
- Malkov D. et al., 2005. Oral delivery of insulin with the eligen technology: mechanistic studies. *Curr. Drug Delivery*, **2**, 191-197.
- Marr A.K., Gooderham W.J. & Hancock R.E., 2006. Antibacterial peptides for therapeutic use: obstacles and

- realistic outlook. *Curr. Opin. Pharmacol.*, **6**(5), 468-472.
- Marx V., 2005. Watching peptide drugs grow up. *Chem. Eng. News*, **83**(11), 17-24.
- McGee P., 2005. First successes turn tide for peptide therapeutics. *Drug Discovery Dev.*, **4**.
- Merrifield B., 1963. Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J. Am. Chem. Soc.*, **85**, 2149-2154.
- Parvu L., 2003. QSAR - a piece of drug design. *J. Cell. Mol. Med.*, **7**(3), 333-335.
- Patch J.A. & Barron A.E., 2002. Mimicry of bioactive peptides via non-natural, sequence-specific peptidomimetic oligomers. *Curr. Opin. Chem. Biol.*, **6**(6), 872-877. Review.
- Pooga M., Hallbrink M., Zorko M. & Langel U., 1998. Cell penetration by transportan. *FASEB J.*, **12**(1), 67-77.
- Prego C., Garcia M., Torres D. & Alonso M.J., 2005. Transmucosal macromolecular drug delivery. *J. Control Release*, **101**, 151-162.
- Sagan S. et al., 2001. Calpha methylation in molecular recognition. Application to substance P and the two neurokinin-1 receptor binding sites. *Eur. J. Biochem.*, **268**(10), 2997-3005.
- Sato A.K., Viswanathan M., Kent R.B. & Wood C.R., 2006. Therapeutic peptides: technological advances driving peptides into development. *Curr. Opin. Biotechnol.*, **17**(6), 638-642.
- Sergeeva A. et al., 2006. Display technologies: Application for the discovery of drug and gene delivery agents. *Adv. Drug Delivery Rev.*, **15**, 1622-1654.
- Shimomura Y. et al., 2003. Modulation of procarboxypeptidase R (ProCPR) activation by complementary peptides to thrombomodulin. *Microbiol. Immunol.*, **47**(3), 241-245.
- Shin D.S., Kim D.H., Chung W.J. & Lee Y.S., 2005. Combinatorial solid phase peptide synthesis and bioassays. *J. Biochem. Mol. Biol.*, **38**(5), 517-525.
- Simons K.T., Kooperberg C., Huang E. & Baker D., 1997. Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and Bayesian scoring function. *J. Mol. Biol.*, **268**(1), 209-225.
- Skyler J.S. et al., 2001. Efficacy of inhaled human insulin in type 1 diabetes mellitus: a randomised proof-of-concept study. *The Lancet*, **357**, 331-335.
- Smith G.P., 1985. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science*, **228**(4705), 1315-1317.
- Stout R.R. et al., 2007. Needle-free injections using a spring-powered device for subcutaneous, intramuscular and intradermal injections. *Drug. Delivery Techn.*, **7**(2).
- Surowiecki J., 2004. The Pipeline Problem. *The New Yorker*, 16 Feb.
- Thomas A. et al., 2006. Prediction of peptide structure: how far are we? *Proteins*, **65**(4), 889-897.
- Veronese F.M. & Pasut G., 2005. PEGylation, successful approach to drug discovery. *Drug Discovery Today*, **10**(21), 1451-1458.
- Watt P.M., 2006. Screening for peptide drugs from the natural repertoire of biodiverse protein folds. *Nature Biotech.*, **24**, 177-183.
- Werle M. & Bernkop-Schnurch A., 2006. Strategies to improve plasma half-life time of peptide and protein drugs. *Amino Acids*, **30**(4), 351-367.
- Yang Z. et al., 2005. Structure-based design and characterization of a novel IL-6 antagonist peptide. *Mol. Immunol.*, **42**(9), 1015-2121.

(67 réf.)

## **Références bibliographiques**

Ayoub M and Scheidegger D (2006). Peptide drugs, overcoming the challenges, a growing business. *Chemistry today* 24(4):46-48

Bauduin G. Biologie du diabète (2001). *Med.Nucl.* 25(2) :55-60

Benkirane N., Friede M., Guichard G., Briand JP., Vanregenmortel M.H.V. and Muller S. (1993). Antigenicity and immunogenicity of modified synthetic peptides containing D-amino acid residues. Antibodies to a D-enantiomer do recognize the parent L-hexapeptide and reciprocally. *J Biol. Chem.* Dec 15;268(35):26279-85.

Berendsen HJC, van der Spoel D, van Drunen R (1995). GROMACS: a message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* 91:43–56.

Binder H. and Lindblom G. (2003). Charge-dependent translocation of the Trojan peptide penetratin across lipid membranes, *Biophys. J.* **85**, 982-995.

Birney DM, Cole DC, Crosson CE, Kahl BF, Neff BW, Reid TW, Ren K, Walkup RD (1995). Use of  $\beta$ -methylphenylalanine ( $\beta$ MeF) residues to probe the nature of the interaction of substance P with its receptor: effect of  $\beta$ MeF-containing substance P analogs on rabbit iris smooth muscle contraction. *J.Medic.Chem.*38:2478-82

Blobel, G. (1980) Intracellular protein topogenesis. *Proc. Natl. Acad. Sci. U. S. A* 77, 1496-1500.

Bonneau R., Strauss C., Rohl C., Chivian D., Bradley P., Malmstrom L., Robertson T.,

Baker D. (2002). De Novo Prediction of Three-dimensional Structures for Major Protein Families. *J. Mol.Biol* Sep 6 322: (1) 65

Bradley P., Chivian D., Meiler J., Misura K.M., Rohl C., Schief W., Wedemeyer W.J., Schueler-Furman O., Murphy P., Schonbrun J., Strauss C., Baker D. (2003). Rosetta predictions in CASP5: Successes, failures, and prospects for complete automation. *Proteins* 53: Oct 15 (S6) 457-68.

Branden and Tooze (1999). Introduction to Protein structures. *Ed Garland* (2d edition).

Brasseur R., Goormaghtigh E., and Ruyschaert J.M. (1981). Theoretical conformational analysis of phospholipid bilayers, *Biochem. Biophys. Res. Commun.* 103 :301-310

Brasseur R, Lorge P, Espion D, Goormaghtigh E, Burny A, Ruyschaert JM (1988). The mode of insertion the paramyxovirus F1 N\_terminus into lipid matrix, an initial step in host cell/virus fusion *Virus Genet.* 1:325-332



Brasseur (1995). Simulating the folding of small proteins by use of the local minimum energy and the free solvation energy yields native-like structures. *J.Mol.Graph.* 13:312-22.

Brasseur (2000). Tilted peptides : a motif for membrane destabilization (Hypothesis). *Mol.Memb.Biol.* 17:31-40

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

Chaloin L, Vidal, P, Heitz A, VanMau N, Mery J, Divita G. and Heitz F. (1997). Conformations of primary amphipathic carrier peptides in membrane mimicking environments. *Biochemistry* **36**, 11179-11187.

Chen H. and Horvath C.S. (1995). High-speed high-performance liquid chromatography of peptides and proteins *J. Chrom.A.* 705:3-20

Chivian D., Kim D.E., Malmstrom L., Bradley P., Robertson T., Murphy P., Strauss C.E., Bonneau R., Rohl C.A., Baker D. (2003). Automated prediction of CASP-5 structures using the Robetta server. *Proteins.*;53 Suppl 6:524-33.

Clare DA and Swaisgood HE. Bioactive milk peptides: a prospectus (2000). *J.Dairy Sci.* 83:1187-95

Cortese R., Monaci P., Luzzago A., Santini C., Bartoli F., Cortese I., Fortugno P., Galfre G., Nicosia A., Felici F. (1996). Selection of biologically active peptides by phage display of random peptide libraries. *Curr Opin Biotechnol.* Dec;7(6):616-21.

Creighton TE (1993). *Proteins: Structures and Molecular Properties.* Ed Freeman (2d edition).

Derossi D., Joliot A., Chassaing G., Prochiantz A. (1994). The third helix of the Antennapedia Homeodomain Translocates through Biological Membranes. *JBC* 14:10444-50.

Derossi, D., Calvet, S., Trembleau, A., Brunissen, A., Chassaing, G. and Prochiantz, A. (1996). Cell internalization of the third helix of the Antennapedia homeodomain is receptor-independent, *J. Biol. Chem.* **271**, 18188-18193.

Deshayes S., Plenat T., Charnet P., Divita G., Molle G., Heitz F. (2006). Formation of transmembrane ionic channels of primary amphipathic cell-penetrating peptides. Consequences on the mechanism of cell penetration, *Biochim. Biophys. Acta.* **1758**, 1846-1851.

Dicu AO., Topham MK., Attaway L. and Epanand R. Role of the hydrophobic segment of DGK $\epsilon$ . *Biochemistry* **46**:6109-17

Dill KA and Bromberg S (2003) Molecular driving forces. *Ed Garland Science*.

Ducarme P, Rahman M, Brasseur R. (1988) Impala: A simple restraint field to simulate the biological membrane in molecular structure studies, *Proteins: Structure, Function and Genetics*. **30**, 357-71

Edwards CMB, Cohen MA, Bloom SR (1999). Editorial: Peptides as drugs. *Q.J.Med* **92**:1-4

Eisenberg D, Weiss RM, Terwilliger TC (1982). The helical hydrophobic moment: a measure of the amphiphilicity of a helix. *Nature* **299**:371-74

Etchebest C., Benros C., Hazout S., de Brevern A.G. (2005). A structural alphabet for local protein structures: improved prediction methods. *Proteins*. Jun 1;59(4):810-27.

Fauchere J-L. and Thurieau C. (1992) Evaluation of the stability of peptides and pseudo-peptides as a tool in peptide drug design. *Adv Drug Res* **23**: 127-159.

Finking R and Marahiel (2004). Biosynthesis of nonribosomal peptides. *Annu.Rev.Micribiol.* **58**:453-88

Fowler, P. W., Balali-Mood, K., Deol, S., Coveney, P. V., and Sansom, M. S. P. (2007) Monotopic Enzymes and Lipid Bilayers: A Comparative Study, *Biochemistry* **46**, 3108-3115.

Froimowitz (1993).HyperChem: a software package for computational chemistry and molecular modeling. *Biotechniques* **14**(6):1010-3.

Ghosh S., Nie A., An J., Huang Z. (2006). Structure-based virtual screening of chemical libraries for drug discovery. *Curr Opin Chem Biol.* **10**(3):194-202. Review.

Gladyshev VN and Hatfield DL (1999). Selenocysteine-containing proteins in mammals. *J. Biomed.Science* **6**(3) 151-160.

Glick M., Rayan A., Goldblum A. (2002) A stochastic algorithm for global optimization and for best populations: A test case of side chains in proteins. *Proc. Natl. Acad. Sci. U.S.A.* **99** : 703-708.

Glukhov, E., Shulga, Y., Eband, R. F., Dicu, A. O., Topham, M. K., Deber, C. M., and Eband, R. M. (2007). Membrane interactions of the hydrophobic segment of diacylglycerol kinase epsilon, *Biochim. Biophys. Acta*, in press.

Gordon YJ, Romanowski EG, McDermott AM (2005). A review of antimicrobial peptides and their therapeutic potential as anti-infective drugs. *Curr Eye Res.* Jul;30(7):505-15.

Goozner M (2004) The \$800 million pill. The truth behind the cost of new drugs. *University of California Press.*

Gozes I, Spivak-Pohis I.(2006) Neurotrophic effects of the peptide NAP: a novel neuroprotective drug candidate. *Curr Alzheimer Res.* Jul;3(3):197-9. Review.

Hao B, Gong WM, Ferguson TK, James CM, Krzycki JA, Chan MK (2002). A new UAG-encoded residue in the structure of a methanogen methyltransferase. *Science* 296 (5572): 1462-66.

Heinen E., Decaffmeyer M., Brasseur R. (1996) Complementary peptides for  $\beta$ -Amyloid 29-42 peptide (Patent PCT/EP2005/055050)

Horth M, Lambrecht B, Chuah Lay Khim M, Bex F, Clothilde C, Ruyschaert JM, Burny A, Brasseur R (1991) Theoretical and functional analysis of the SIV fusion peptide. *EMBO J.* 10: 2747-55

Hartmann R and Meisel H. Food-derived peptides with biological activity: from research to food applications (2007). *Curr. Opin. Biotechnol.* 18:163-69

Hruby JH. (2002) Designing peptide receptor agonists and antagonists. *Nature Reviews Drugs Discovery* 1 (11): 847-58

Hummel G, Reineke U, Reimer U (2006). Translating peptides into small molecules. *Mol. Biosyst.*, 2 499-508

Jones AT. (2007). Macropinocytosis: searching for an endocytic identity and role in the uptake of cell penetrating peptides. *J Cell Mol Med.* **11**, 670-84.

Kamysz W., Okroj M., Lemicka E., Ossowski T. and Lukasiak (2004). Fast and efficient purification of synthetic peptides by solid-phase extraction. *Acta Chrom.* 14: 180-86

Kaplan, IM. Wadia, JS. and Dowdy, SF. (2005). Cationic Tat peptide transduction domain enters cells by macropinocytosis, *J Control Release* **102**, 247-253.

Karp G (1996). Cell and molecular biology, concept and experiments. Ed Willey & Sons.

Kasher R., Balass M., Scherf T.; Fridkin M., Fuchs S., Katchalski-Katzir E. Design and synthesis of peptides that bind alpha-bungarotoxin with high affinity, *ChemBiol* 8 (2001) 147-155

Katchalski-Katzir E., Kasher R., Balass M., Scherf T., Harel M., Fridkin M., Sussman J.L. and Fuchs S., Design and synthesis of peptides that bind  $\alpha$ -bungarotoxin with high affinity and mimic the three-dimensional structure of the binding-site of acetylcholine receptor, *Biophysical Chemistry*, Volume 100, Issues 1-3, 15 January (2003) 293-305.

Kaur H., Raghav G.P. (2004) Neural Network Method for prediction of beta-turn types in proteins using evolutionary information. *Bioinformatics* Nov 1; 20(12):2751-8

Konturek PC and Konturek SJ (2003). The history of gastrointestinal hormones and the Polish contribution to elucidation of their biology and relation to nervous system. *J Physiol Pharmacol*. 2003 Dec;54 Suppl 3:83-98.

Kreil G (1997). D-Amino acids in animal peptides. *Annu.Rev.Biochem* 66:337-45.

Kuhlman B., Dantas G., Ireton G.C., Varani G., Stoddard B.L., Baker D. (2003). Design of a novel globular protein fold with atomic-level accuracy. *Science*. Nov 21;302(5649):1364-8.

Lathan PW (1999) Therapeutic peptides revisited. *Nature Biotech.* 17:755-57

Letoha, T., Gaal, S., Somlai, C., Czajlik, A., Perczel, A. and Penke, B. (2003). Membrane translocation of penetratin and its derivatives in different cell lines. *J. Mol. Recognit.* **16**, 272-279.

Lien S and Lowman HB (2003). Therapeutic peptides. *Trends Biotechnol.*;21(12):556-62.

Lins L., Thomas-Soumarmon A., Pillot T., Vandekerckhove J., Rosseneu M., and Brasseur R. (1999). Molecular determinant of the interaction between the C-terminal domain of the Alzheimer's  $\beta$ -amyloid peptide and apolipoprotein E  $\alpha$ -helices, *J.Neurochem.* 73 758-769.

Lins L., Charleatoux B., Heinen C., Thomas A., Brasseur R. (2006). De novo design of peptides with specific lipid-binding properties, *Biophys.J.* **90**(2):470-479.

Lindahl (2003) S GE. One hundred years of nobel prizes and 150 years of anesthesiology. *J.Anesth.* 17:42-45

- Lindberg, M., Jarvet, J., Langel, U. and Graslund, A. (2001). Secondary structure and position of the cell-penetrating peptide transportan in SDS micelles as determined by NMR. *Biochemistry* **40**, 3141-3149.
- Lindberg, M., Biverstahl, H., Graslund, A. & Maler, L. (2003). Structure and positioning comparison of two variants of penetratin in two different membrane mimicking systems by NMR. *Eur. J. Biochem.* **270**, 3055-3063.
- Loffet, A (2002). Peptides as drugs: Is there a market? *J.Pept.science* 8 (1): 1-7.
- Lovell S.C., Davis I.W., Arendall W.B. 3rd, de Bakker P.I., Word J.M., Prisant M.G., Richardson J.S., Richardson D.C. (2003). Structure validation by C $\alpha$  geometry: phi,psi and C $\beta$  deviation. *Proteins*. Feb 15;50(3):437-50.
- Merrifield B. (1963). Solid Phase Peptide Synthesis.I.The Synthesis of a Tetrapeptide. *J.Am.Chem.Soc* 85 .2149-54
- Mae M. and Langel U. Cell-penetrating peptides as vectors for peptide, protein and oligonucleotide delivery, *Curr. Opin. Pharmacol.* **6**, 509-514 (2006).
- Maggio ET (2005). Recent developments in intranasal drug delivery technology are creating new vistas for peptide and protein therapeutics. *PharmaVenture* June 13.
- Malkov D, Angelo R, Wang HZ, Flanders E, Tang H, Gomez-Orellana I (2005). Oral delivery of insulin with the eligen technology: mechanistic studies.*Curr Drug Deliv.* Apr;2(2):191-7.
- Marx V (2005). Watching peptide drugs grow up. *Chem. Engin. News* 83 (11): 17-24.
- McGee P.(2005). First Successes Turn Tide for Peptide Therapeutics. *Drug Discovery & Development*. Ed April <http://www.dddmag.com>
- Med Ad News (2005). May
- Mehta NM (2004). Oral delivery and recombinant production of peptide hormones. *Biopharm Int* June
- Morishita M. and Peppas NA. Is oral route possible for peptide and protein drug delivery. *Drug Disc. Today* **11**, 905-910 (2006).
- Morris MC., Vidal P., Chaloin L., Heitz F. and Divita G. (1997). A new peptide vector for efficient delivery of oligonucleotides into mammalian cells. *Nucleic Acids Res.* **25**, 2730-2736.

Morris MC., Depollier J., Mery J., Heitz F. and Divita G. (2001). A peptide carrier for the delivery of biologically active proteins into mammalian cells. *Nat Biotechnol.* **19**, 1173-1176.

Moss GP (1974). The nomenclature of peptides hormones. IUPAC. <http://www.chem.qmul.ac.uk/iubmb/misc/phorm.html>

Pillot T, Goethals M, Vanloo B, Talusot C, Brasseur, Vandekerckhove J, Rosseneu M, Lins L (1996). Fusogenic properties of the C-terminal domain of the Alzheimer beta-amyloid peptide. *JBC* 271:28757-65

Pineau C, Jegou B, Bourgeon F, Nicolas J, Bastide Y, Mescam Y (2006). (WO/2006/061406) antimicrobial peptides, polypeptides comprising same, gene coding for said peptides, vectors, transformed organisms and compositions containing same. <http://www.wipo.int/pctdb/en/wo.jsp?wo=2006061406>.

Pooga M., Hallbrink M., Zorko M. and Langel U. (1998). Cell penetration by transportan, *FASEB J.* **12**, 67-77.

Prescrire (2003). Coût de recherche et développement du médicament: la grande illusion. *Prescrire* 23 (244) : 782-87

Rahman M, Lins L, Thomas A, Brasseur (1997). Are amphipatic asymmetric peptides ubiquitous structures for membrane destabilisation? *J.Mol.Model.* 3:203-15

Ramachandran GN, Ramakrishnan C, Sasisekharan V(1963). Stereochemistry of polypeptide chain configurations. *J. Mol Biol.* 7:95-9

Richard, JP, Melikov, K, Brooks, H, Prevot, P, Lebleu, B, and Chernomordik, LV. (2005). Cellular uptake of unconjugated TAT peptide involves clathrin dependent endocytosis and heparan sulfate receptors, *J. Biol. Chem.* **280**, 15300-15306.

Sagan S., Lequin O., Frank F., Convert O., Ayoub M., Lavielle S., Chassaing G. (2001). Calpha methylation in molecular recognition. Application to substance P and the two neurokinin-1 receptor binding sites. *Eur J Biochem.* 268(10):2997-3005.

Shoji M., Cerebrospinal fluid Abeta 40 and Abeta 42: natural course and clinical usefulness (2002), *Front Biosci.* 7 997-1006.

Sieber SA et Marahiel MA (2003). Learning from nature's drug factories: nonribosomal synthesis of macrocyclic peptides. *J.Bac.* Dec: 7036-43.

Simons K.T., Kooperberg C., Huang E., Baker D. (1997). Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and Bayesian scoring function. *J.Mol.Biol.* Apr 25; 268(1):209-25.

Skyler JS, Cefalu WT, Kourides IA, Landschulz WH, Balagtas CC, Cheng SL and Gelfand RA (2001). Efficacy of inhaled human insulin in type 1 diabetes mellitus: a randomised proof-of-concept study. *The Lancet*, 357 Feb:331-335

Smith G.P. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science*. Jun 14;228(4705):1315-7.

Soomets, U., Lindgren M., Gallet X., Hällbrink M., Elmquist A., Balaspiri L., Zorko M., Pooga M., Brasseur R., Langel U (2000). Deletion analogues of transportan. *Biochim. Biophys. Acta.* **1467**, 165-176.

Spisni E., Tomasi V., Cestaro A. and Tosatto S. (2005). Structural insights into the function of human caveolin 1. *Biochem. Biophys. Res. Comm.* **338**:1383-1390

Stout RR, Gutierrez MJ, Freeland MD, Fein M, Taylor D, Davis Y, Quiuroz R, Correa O, Machado A, Landau S, Garcia M, Arencibia E, Chinery-Hesse N, Wai C and Macias (2007). Needle-free injections using a spring-powered device for subcutaneous, intramuscular and intradermal injections. *Drug Deliv. Techn.* Feb Vol 7 n°2

Surowiecki (2004). Financial Page. New Yorker: 16-02-2004

Thomas A., Bouffieux O., Geurickx D., and Brasseur R. (2001). Pex, analytical tools for PDB files. I. GF-Pex: basic file to describe a protein, *Proteins* 43:28-36.

Van Heijenoort J (2001). Formation of the glycan chains in the synthesis of bacterial peptidoglycan. *Glycobiology* 11 (3) 25-26.

Werle M and Bernkop-Schnürch A (2006). Strategies to improve plasma half life time of peptide and protein drugs. *Amino Acids* 30:351-67

Wyman TB., Nicol F., Zelphati O., Scaria PV., Plank C. and Szoka FC Jr. (1997). Design, synthesis, and characterization of a cationic peptide that binds to nucleic acids and permeabilizes bilayers. *Biochemistry.* **36**, 3008-3017.

Yang Z., Feng J., Li Y., Hu M., Song L., Yu M., Qin W., Shen B. (2005), Structure-based design and characterization of a Novel IL-6 antagonist peptide, *Molecular Immunology* 42:1015-21