Université de Reims Champagne-Ardenne

UFR Sciences Exactes et Naturelles

THESE DE DOCTORAT

Ecole doctorale Sciences et Santé

Présentée en vue de l'obtention du grade de

Docteur en Sciences de l'Université de Reims Champagne-Ardenne

Spécialité : **Biomolécules et Dynamique Cellulaire**

par

Gautier MOROY

Etude structurale du domaine d'interaction du récepteur de l'élastine. Approches biochimiques, biophysiques et bioinformatiques.

Soutenue publiquement le Mercredi 14 Décembre 2005 à 15 H 00

devant le jury composé de :

Pr Alain. J.P. ALIX	U.R.C.A.	Reims	Directeur de Thèse
Pr Manuel DAUCHEZ	U.R.C.A.	Reims	Co-directeur
Dr William HORNEBECK	CNRS-URCA	Reims	Examinateur
Dr David PERAHIA	IBBMC CNRS-UPS	Orsay	Rapporteur, examinateur
Dr Roland H. STOTE	Institut LeBel CNRS-ULP	Strasbourg	Rapporteur, examinateur
Dr Stéphanie HÉRY-HUYNH	U.R.C.A.	Reims	Invité, co-encadrante

REMERCIEMENTS

Je voudrais tout d'abord remercier, le Professeur Alain J.P. ALIX, directeur du Laboratoire de Spectroscopies et Structures BioMoléculaires (LSSBM, IFR53 Biomolécules, URCA) pour m'avoir accueilli au sein de son laboratoire et m'avoir encadré durant mes années de DEA et de thèse. Je voudrais exprimer ici ma gratitude pour m'avoir donné toute liberté dans la manière de mener à bien ce travail de thèse. Merci également pour la grande confiance que ses qualités humaines a su instaurer entre nous, pour sa rigueur en toutes circonstances et en tout temps, ainsi que pour ses conseils avisés que je ne saurais oublier.

Que le Professeur Manuel A.M. DAUCHEZ soit remercié pour avoir guidé mes premiers pas dans le domaine de la recherche scientifique.

Je voudrais exprimer aux Docteurs David PERAHIA et Roland STOTE ma reconnaissance pour avoir accepté de juger ce travail, ainsi qu'au Docteur William HORNEBECK pour avoir accepté l'invitation à être membre du jury de cette thèse.

J'aimerais exprimer ma profonde gratitude au Docteur Stéphanie HÉRY-HUYNH, responsable des matériels bioinformatiques de l'équipe, pour ses conseils, ses remarques pertinentes, sa gentillesse et la grande patience dont elle a toujours su faire preuve à mon égard. Je voudrais la remercier pour son aide scientifique sincère et désintéressée dès la première heure et également pour avoir accepté de faire partie de mon jury de thèse. Je lui exprime ici, mes plus vifs remerciements.

Merci à l'ensemble des membres du Laboratoire de Chimie de l'Université de la Basilicate (Potenza, Italie) pour avoir rendu mon séjour de 3 mois aussi agréable du point vue scientifique que relationnel, et particulièrement le Professeur Antonio M. TAMBURRO pour nos discussions scientifiques et son excellent accueil au sein de son laboratoire, le Docteur Angela OSTUNI pour son aide omniprésente, le Docteur Antonietta PEPE pour la prise des spectres RMN, le Docteur Brigida BOCHICCHIO et Marina RUSSO pour leur sincère gentillesse.

Je voudrais remercier les membres de l'Institut de Génétique Médicale (Université de Humboldt, Berlin), de l'Institut de Biochimie (Université de Humboldt, Berlin), du Centre ″Rosalind Franklin″ pour la Recherche Génomique (Wellcome Trust Genome Campus, Cambridge) et de l'Institut ″Max Planck″ de Génétique Moléculaire pour nos travaux en collaboration sur les peptides dérivant de la fibrilline-1.

Que soient également remerciés les membres de l'unité CNRS FRE2715 pour les collaborations scientifiques que nous avons eues ensemble, particulièrement Mademoiselle

Elise PROST, le Docteur Laurent VERDIER, ainsi que le Docteur Jean-Marc NUZILLARD pour l'étude RMN d'un peptide d'intérêt.

J'aimerais remercier les membres du département de Mathématiques-Informatique de la Faculté des Sciences de l'URCA pour les vacations et les postes d'Attaché Temporaire d'Enseignements et de Recherches qui m'ont été accordées.

Je voudrais remercier sincèrement les membres du LSSBM : le Docteur Amina El HDIY, le Docteur Patrick FUCHS pour son aide spontanée, le Docteur Michael MAIZIERES pour sa sympathie et sa constante bonne humeur. Merci également à Monsieur Donat-Jacques DORÉ, au Docteur Nicolas FLOQUET et à Monsieur Hocine YEZID. Un grand merci aussi à Mademoiselle Françoise CHARTON pour m'avoir initié aux techniques expérimentales de spectroscopies Raman et de DCE.

Enfin , je remercie tout le ″groupe du midi″ pour nos longues discussions à propos de tout et de rien et pour les agréables moments que nous avons pu passer ensemble : Stéphanie, Gérard, Cédric, Alex, Etienne, Alix, Nicolas, Maryline, Magda et Clara.

Merci, enfin, au Docteur Eldra DELANNAY pour son éternelle bonne humeur et sa gentillesse sans égale.

RESUME

Ce travail de thèse porte sur l'étude théorique et expérimentale de l'interaction entre des peptides issus de protéines de la matrice extracellulaire (MEC) et leur domaine de fixation sur la sous-unité EBP (Elastin Binding Protein) du récepteur de l'élastine.

Nous avons tout d'abord cherché à caractériser le récepteur EBP, et étudié dans ce but le peptide Sgal (**VVGSPSAQDEASPLS**) issu du domaine de fixation de l'EBP. Les résultats expérimentaux, obtenus par spectroscopies de dichroïsme circulaire électronique et de RMN, nous indiquent que le peptide Sgal est en équilibre entre une structure en hélice polyproline II et une conformation indéterminée. Par simulation de dynamique moléculaire, nous avons obtenu des résultats concordant avec nos résultats expérimentaux et nous proposons qu'un coude β de type I sur la séquence **QDEA**, à cause de sa stabilité au cours des trajectoires de dynamique moléculaire, est important et nécessaire pour l'activité réceptrice de l'EBP.

Nous nous sommes ensuite intéressés aux ligands peptidiques de l'EBP et avons focalisé notre attention sur le comportement structural de peptides contenant la séquence **GXXP** (X étant un résidu quelconque) et dérivant de protéines de la MEC, essentiellement de l'élastine, mais également de la fibrilline-1. Par simulations de dynamique moléculaire, de Monte-Carlo et par calcul de leur carte adiabatique, nous avons pu mettre en évidence qu'un coude β de type VIII sur **GXXP** semble être le repliement préférentiel de ces peptides. Tous les peptides étudiés, à l'origine d'effets biologiques et dont la fixation à l'EBP est prouvée expérimentalement, se replient seulement en coude β de type VIII. En revanche, les peptides biologiquement inactifs, se replient en d'autres types de coudes β même si ils peuvent adopter également et de manière transitoire un coude β de type VIII.

Finalement, nous avons démontré expérimentalement que le peptide Sgal peut directement se lier au peptide **VGVAPG** avec une grande affinité (constante de dissociation de 26,7 nM). Nous avons donc étudié, par amarrage moléculaire, l'interaction entre certains peptides dérivant de l'élastine et l'Elastase Pancréatique Porcine (PPE), prise comme modèle car elle possède un domaine d'interaction similaire à celui de l'EBP. La meilleure solution proposée pour chacun des peptides est un coude β déformé de type VIII sur la séquence **GXXP**, qui se positionne autour du résidu $\mathbf{Q_8}$ de la PPE. Le complexe peptide-PPE est stabilisé par 3 liaisons hydrogène. Nous avons ensuite testé la stabilité du complexe peptide-peptide Sgal dans la même configuration spatiale que précédemment. Pendant 20 ns, les 2 peptides restent en contact grâce aux 3 liaisons hydrogène qui sont observables durant toute la trajectoire de dynamique moléculaire. Nous avons donc probablement réussi à trouver la conformation nécessaire au peptide pour se lier à l'EBP (un coude β de type VIII sur **GXXP**), mais également les résidus ($\mathbf{Q_8}$) et la structure du domaine de fixation (coude β de type I sur $\mathbf{Q_8DEA_{11}}$) de l'EBP, impliqués dans l'interaction.

Mots-clés : EBP, Sgal, élastine, tropoélastine, peptide dérivé d'élastine (EDP), fibrilline-1, élastase pancréatique porcine (PPE), matrice extracellulaire (MEC); biophysique, biospectroscopie, bioinformatique; dichroïsme circulaire électronique, RMN, fluorescence, diffusion Raman; coudes β, polyproline II; modélisation moléculaire, mécanique et dynamique moléculaires, carte adiabatique, Monte-Carlo, amarrage moléculaire (docking).

ABSTRACT

This thesis deals with the theoretical and experimental study of the interaction between peptides derived from ExtraCellular Matrix (ECM) proteins and one of the elastin receptor complex sub-units: the Elastin Binding Protein (EBP).

In the first part of this work, we characterised the structure of the EBP binding domain, called the Sgal peptide (**VVGSPSAQDEASPLS**). Experimental results, obtained from Circular Dichroism and NMR spectroscopies, have shown that the Sgal peptide structure presents an equilibrium between a Polyproline II helix structure and an unordered conformation. The molecular dynamic simulations results are in good agreement with those obtained experimentally showing unordered conformations, many of them exhibiting however a type I β-turn spanning the QDEA sequence. On the basis of this turn stability, we proposed that a type I β-turn spanning the QDEA sequence is crucial and necessary for the EBP receptor activity.

We then focused on the structural behaviour of EBP peptides. The studied peptides were all containing a GXXPG sequence (X being any residue) and derived from ECM proteins (elastin and fibrillin-1). By molecular dynamics simulations, Monte-Carlo and adiabatic map calculations, we have demonstrated that a type VIII β-turn on the GXXP motif seemed to be their preferential fold. All the peptides, whose biological activity and anchoring to the EBP had already been proved, were presenting this king of folding. In spite of the fact that the inactive peptides could fold transitorily into a type VIII β-turn, they were the only peptides whose conformation was characterised by another β-turn type, mainly type II' or I.

Finally, we have demonstrated experimentally that the Sgal peptide was able to bind directly to the VGVAPG peptide with a high affinity (the dissociation constant was equal to 26,7 nM). We have then studied, using molecular docking, the interaction between several Elastin Derived Peptides (EDP) and the Pancreatic Porcine Elastase (PPE), which shows a sequence homology of 46% with the Sgal peptide. The best solution proposed for each EDP were a distorted type VIII β-turn on the GXXP sequence, located around the Q_8 residue of the PPE. The EDP-PPE complex was moreover stabilised by 3 H-bonds. We have studied the VGVAPG peptide – Sgal peptide complex stability with the same spatial configuration than in the previous complex: during 20 ns of molecular dynamic simulation, these 2 peptides were in contact and the 3 H-bonds were observable during all the molecular dynamic trajectory. We can then conclude that a type VIII β-turn on GXXP for the EDP and a type I β-turn on QDEA for the elastin binding domain seem to be essential for the interaction between EDP and EBP.

Keywords:
EBP, Sgal peptide, Elastin Derived Peptide (EDP), fibrillin-1, Pancreatic Porcine Elastase (PPE), ExtraCellular Matrix (ECM); biophysic, bioinformatic, NMR, fluorescence, Raman diffusion; β-turns, polyproline II; molecular modelling, molecular mechanic, molecular dynamic, adiabatic map, Monte-Carlo simulation, molecular docking.

TABLE DES MATIERES

TABLE DES MATIERES i

Liste des Figures vi

Liste des Tableaux xii

Liste des Abréviations xiv

Partie I : INTRODUCTION 01

A. INTRODUCTION GENERALE 01

B. STRUCTURE DES PROTEINES 04

B.1. Structure primaire 04

B.2. Structure secondaire 05
a. Les hélices 08
b. Les brins β 09
c. Les coudes 10
d. Les boucles 12

B.3. Structure tertiaire 12

B.4. Structure quaternaire 13

C. CONTEXTE BIOLOGIQUE 14

C.1. L'élastine 15
a. La tropoélastine 15
α. Structure primaire *15*
β. Structures secondaire et tertiaire *18*
b. Les peptides d'élastine 20

C.2. Le récepteur de l'élastine 23
a. Structure 23
b. Fonction 26

Partie II : MATERIELS ET METHODES 29

A. METHODES D'ETUDES EXPERIMENTALES 29

A.1. Synthèse peptidique 29

A.2. Spectroscopie de Dichroïsme Circulaire Electronique (DCE) 31
a. Principe du pouvoir rotatoire 31
b. Principes du dichroïsme circulaire 33
c. Aspect pratique du dichroïsme circulaire 36
d. Spectres caractéristiques des structures secondaires des protéines 36
e. Analyse de la structure secondaire par dichroïsme circulaire 39

A.3. Spectroscopie de Résonance Magnétique Nucléaire (RMN) **40**
a. Aspect théorique 40
b. Instrumentation, mesure et interprétation 42

A.4. Spectroscopie de diffusion Raman **43**
a. Aspect théorique 43
b. Instrumentation et interprétation 44

A.5. Spectroscopie de fluorescence **46**
a. Aspect théorique 46
b. Instrumentation et interprétation 47

B. METHODES D'ETUDES THEORIQUES : MODELISATION MOLECULAIRE **49**

B.1. Mécanique Moléculaire **51**
a. Fonction d'énergie potentielle 51
b. Traitement des interactions non liantes 53
c. Traitement du solvant 55
α. Solvatation implicite *56*
β. Solvatation explicite *58*
d. Méthodes de minimalisation 59
α. Méthode Steepest Descent SD *60*
β. Méthode du Gradient Conjugué GC *61*
γ. Méthode ABNR *61*

B.2. Exploration de l'hypersurface d'énergie potentielle **61**
a. Carte adiabatique 62
b. Simulation de Monte Carlo 65

B.3. Dynamique Moléculaire **66**
a. Principe 66
b. Description d'un protocole 68
c. Analyse 70
α. Séries temporelles *70*
β. RMSD et carte de RMSD croisés *70*
γ. Analyse des structures secondaires *72*
δ. Analyse des coudes β *72*
ε. Analyse de la carte de Ramachandran *72*
φ. Analyse par amas de conformations (clusters) *73*
χ. Calcul de variation d'énergie libre ΔG *74*

B.4. Amarrage moléculaire **74**

B.5. Moyens informatiques **77**
a. Logiciels utilisés 77
b. Moyens de calcul 78

Partie III : RESULTATS **79**

A. ETUDE DU RECEPTEUR **79**

A.1. Prédictions de structures secondaires **80**

A.2. Recherche d'épitopes linéaires **82**

B. ETUDE DU PEPTIDE SGAL **84**

B.1. Résultats expérimentaux **84**

a. Spectroscopie de dichroïsme circulaire électronique 84
α. DCE de Sgal dans l'eau *84*
β. DCE de Sgal dans le TFE *85*
b. Spectroscopie de résonance magnétique nucléaire 86

B.2. Résultats théoriques **87**

a. Construction d'un modèle par homologie du peptide Sgal 87
b. Simulations de Monte-Carlo du peptide Sgal 91
α. Simulation MC dans l'eau *93*
β. Simulation MC dans le cyclohexane *95*
γ. Conclusions *98*
c. Simulations de dynamique moléculaire du peptide Sgal 100
α. Simulations de dynamique moléculaire du peptide Sgal dans l'eau *101*
β. Simulations de dynamique moléculaire du peptide Sgal dans d'autres solvants *124*
d. Conclusions 128

C. ETUDE DU PEPTIDE DE 32 RESIDUS DONT LA SEQUENCE EST SPECIFIQUE A L'EBP **129**

C.1. Etude en spectroscopie de Dichroïsme Circulaire Electronique **129**

C.2. Spectroscopie de Résonance Magnétique Nucléaire **131**

C.3. Conclusion / perspectives **132**

Partie IV : ETUDE des LIGANDS PEPTIDIQUES de L'EBP **133**

A. PEPTIDES DERIVANT DE L'ELASTINE HUMAINE **134**

A.1. Etude structurale de l'ensemble des EDP contenant GXXP **134**

a. Peptides étudiés 134
b. Protocole de simulation de dynamique moléculaire 135
c. Résultats 136

α. Coudes β sur $X_1GX_2X_3$ *137*
β. Coudes β sur $X_2X_3PX_4$ *138*
γ. Coudes β sur GX_2X_3P *139*
δ. Présence d'hélice polyproline II *144*
ε.. Relations structure - activité biologique *146*
φ. Conclusions *148*
A.2. Etude structurale des EDPs d'activité biologique connue **149**
a. Etude par simulation de Monte-Carlo 149
α. Protocole *149*
β. Résultats *149*
γ. Conclusions *152*
b. Etude par carte adiabatique 153
α. Peptide VGVAPG *154*
β. Peptide PGAIPG *155*
γ. Peptide GVAPG *155*
δ. Peptide GLVPG *156*
ε. Peptide GGVPG *156*
φ. Conclusions *157*
c. Etude par dynamique moléculaire **157**
α. Peptide VGVAPG *158*
β. Peptide PGAIPG *166*
γ. Peptide GGVPG *171*
δ. Conclusions *177*
A.3. Etude d'un peptide inactif non naturel **178**

B. AUTRES LIGANDS PEPTIDIQUES DE L'EBP : LES PEPTIDES DERIVANT DE LA FIBRILLINE-1 **181**

B.1. Etudes expérimentales **182**
a. Peptides étudiés 182
b. Protocole d'étude 183
c. Résultats 183
B.2. Etudes par simulations de dynamique moléculaire **185**
a. Protocole des simulations de dynamique moléculaire 185
b. Résultats des simulations de dynamique moléculaire 186
B.3. Conclusions **191**

C. CONCLUSIONS GENERALES sur les LIGANDS de L'EBP **193**

Partie V : INTERACTIONS LIGANDS - RECEPTEUR 195

A. INTERACTION PPE / EDP 195

A.1. Etude par amarrage moléculaire 195

a. Vérification du protocole d'amarrage moléculaire utilisé 195
b. Amarrage moléculaire des EDP sur la PPE 198

A.2. Dynamique moléculaire de la PPE avec le peptide VGVAPG 200

a. Dynamique moléculaire de la PPE seule 200
b. Dynamique moléculaire de la PPE en interaction avec VGVAPG 203

A.3. Conclusions 208

B. INTERACTION PEPTIDE Sgal / EDP : 209

B.1. Résultats expérimentaux : 210

a. Dichroïsme circulaire électronique 210
b. Fluorescence 211
c. Conclusions 212

B.2. Résultats théoriques 213

a. Amarrage moléculaire 213
b. Simulations de dynamique moléculaire : 214

α. En partant des modèles proposés par l'amarrage moléculaire sur le peptide Sgal 214

β. En partant des modèles proposés par l'amarrage moléculaire sur la PPE 215

C. CONCLUSIONS 220

Partie VI : CONCLUSIONS ET PERSPECTIVES 221

Partie VII : BIBLIOGRAPHIE 225

Liste des Figures

Figure 01 : Structure des 20 acides aminés composant les protéines. 05

Figure 02 : Angles dièdres ω, ϕ et ψ. 06

Figure 03 : Liaison peptidique *trans* (ω = 180°) et *cis* (ω = 0°). 06

Figure 04 : Cartes de Ramachandran. 07

Figure 05 : Structures secondaires en hélice. 08

Figure 06 : Feuillets β. 10

Figure 07 : Coude β. 11

Figure 08 : Structure tertiaire de la porine. 12

Figure 09 : Structure quaternaire de l'hémoglobine. 13

Figure 10 : Séquences de la tropoélastine humaine et de ses exons associés. 17

Figure 11 : Séquences protéiques de l'EBP et de la β-galactosidase humaine. 24

Figure 12 : Homologies de séquences du peptide Sgal. 24

Figure 13 : Représentation du récepteur de l'élastine. 25

Figure 14 : Représentation de la protéine protectrice humaine. 25

Figure 15 : Rôle de l'EBP durant l'élastogenèse. 26

Figure 16 : Rôle extracellulaire de l'EBP. 27

Figure 17 : Chromatogramme du peptide Sgal. 31

Figure 18 : Schéma d'une onde plane monochromatique. 32

Figure 19 : Directions du vecteur champ électrique E. 32

Figure 20 : Décomposition du vecteur champ électrique E. 33

Figure 21 : Représentation de l'ellipticité de la lumière résultant du phénomène dichroïque. 34

Figure 22 : Coefficients d'extinction molaire et dichroïsme circulaire. 34

Figure 23 : Spectres DCE caractéristiques des structures secondaires. 37

Figure 24 : Spectres DCE de référence des coudes de type I et II et du random coil. 38

Figure 25 : Spectres DCE du peptide GDNP (coude β de type VIII). 38

Figure 26 : Energies des états de spin en corrélation avec le champ B_0. 40

Figure 27 : Niveaux énergétiques vibrationnels et transitions moléculaires impliqués dans les processus de diffusion. 44

Figure 28 : Niveaux énergétiques et transitions moléculaires impliqués dans le phénomène de fluorescence. 46

Figure 29 : Détermination de la constante de dissociation K_D. 48

Figure 30 : Eléments de la fonction énergie potentielle de CHARMM pour les interactions liantes. 52

Figure 31 : Traitement des interactions non liantes. 54

Figure 32 : Distribution des constantes diélectriques dans un modèle de solvant implicite basé sur le continuum électrostatique. 57

Figure 33 : Création des 26 images autour du système initial pour appliquer les conditions périodiques aux limites. 59

Figure 34 : Représentation schématique d'une surface d'énergie. 60

Figure 35 : Carte adiabatique d'un disaccharide. 62

Figure 36 : Suivi de l'énergie totale et de la température du système au cours des différentes phases d'une simulation de dynamique moléculaire. 69

Figure 37 : Carte de RMSD croisés. 71

Figure 38 : Représentation des différentes zones de la carte de Ramachandran prises en compte pour le calcul d'une carte adiabatique. 73

Figure 39 : Schéma d'une grille calculée par AutoGrid. 76

Figure 40 : Principe de l'algorithme génétique de Lamarck. 77

Figure 41 : Comparatif de la puissance de calcul des différentes machines utilisées. 78

Figure 42 : Prédictions des structures secondaires de l'EBP 82

Figure 43 : Spectres de DCE de Sgal dans l'eau. 84

Figure 44 : Spectres de DCE de Sgal dans le TFE. 85

Figure 45 : Homologies de séquences entre les EBP humaine et de mouton et des élastases. 88

Figure 46 : Homologie entre l'EBP humaine et l'Elastase Pancréatique de Porc (PPE). 88

Figure 47 : Structure 3D de l'Elastase Pancréatique de Porc (1BOE). 89

Figure 48 : Structure du peptide Sgal construit par homologie. 90

Figure 49 : Structures secondaires du modèle par homologie du peptide Sgal. 90

Figure 50 : Homologies de séquences entre le peptide Sgal et l'intégrine $\alpha_V\beta_3$. 90

Figure 51 : Protocole d'adaptation de la température en simulation Monte-Carlo. 91

Figure 52 : Protocole de choix des structures de départ des différents cycles Monte-Carlo 92

Figure 53 : Structures résultant des calculs de Monte Carlo dans l'eau. 93

Figure 54 : Les 5 structures Monte-Carlo de plus basses énergies (en kcal.mol^{-1}). 94

Figure 55 : Détails de la partie N-terminale des conformations 1, 2, 4 et 5. 95

Figure 56 : Structures obtenues par Monte Carlo dans le cyclohexane. 96

Figure 57 : Structures obtenues par Monte Carlo dans l'eau. 97

Figure 58 : Types de coudes β dans les structures de plus basses énergies obtenues par Monte-Carlo dans l'eau et le cyclohexane. 98

Figure 59 : Présence des différents types de coude β au cours du temps pour la simulation de 25 ns de Sgal dans l'eau. 102

Figure 60 : Zones de la carte de Ramachandran occupées par les résidus de Sgal pour la simulation de 25 ns dans l'eau. 103

Figure 61 : Carte des RMSD croisés de Sgal pour la DM de 25 ns dans l'eau. 104

Figure 62 : Amas de conformations de Sgal durant la DM de 25 ns dans l'eau, calculés par l'algorithme ART-2. 105

Figure 63 : Structure moyenne de chaque amas de conformations calculés par l'algorithme ART-2 au cours de la DM de 25 ns de Sgal dans l'eau. 106

Figure 64 : Présence de liaison H durant la simulation de 25 ns de Sgal dans l'eau. 106

Figure 65 : Présence des différents types de coudes β au cours du temps pour la simulation de 30 ns de Sgal dans l'eau. 108

Figure 66 : Zones de la carte de Ramachandran occupées par les résidus du peptide Sgal pour la simulation de 30 ns dans l'eau. 108

Figure 67 : Carte des RMSD croisés du peptide Sgal pour la simulation de 30 ns dans l'eau. 109

Figure 68 : Amas de conformations de Sgal durant la DM de 30 ns dans l'eau. 110

Figure 69 : Structure moyenne de chaque amas de conformations calculés par ART-2 au cours de la trajectoire de 30 ns de Sgal dans l'eau. 110

Figure 70 : Présence des différents types de coudes β au cours du temps pour la simulation de 15 ns de Sgal dans l'eau. 112

Figure 71 : Zones de la carte de Ramachandran occupées par les résidus de Sgal pour la simulation de 15 ns dans l'eau. 112

Figure 72 : Carte des RMSD croisés de Sgal pour la simulation de 15 ns dans l'eau. 113

Figure 73 : Amas de conformations de Sgal calculés par ART-2 durant la simulation de 15 ns dans l'eau. 114

Figure 74 : Structure moyenne de chaque amas de conformations calculés par ART-2 au cours de la trajectoire de 15 ns de Sgal dans l'eau. 114

Figure75 : Carte des RMSD croisés de Sgal pour les 3 simulations dans l'eau. 115

Figure 76 : Carte des RMSD croisés des résidus VVGSPSAQD de Sgal pour les 3 DM dans l'eau et structure moyenne des amas de conformations. 116

Figure 77 : RMSD des conformations de Sgal par rapport à la structure MC5 obtenue lors des simulations de DM de a) 15 ns et b) 25 ns. 117

Figure 78 : Carte de RMSD croisés des résidus AQDEAS de Sgal pour les 3 DM dans l'eau. 117

Figure 79 : Présence des différents types de coudes β de Sgal dans l'eau, au cours du temps pour la simulation sous contrainte de distance. 119

Figure 80 : Zones de la carte de Ramachandran occupées par les résidus de Sgal dans l'eau et sous contrainte de distance. 120

Figure 81 : Carte de RMSD croisés de Sgal dans l'eau et sous contrainte de distance.121

Figure 82 : Amas de conformations au cours du temps calculés par ART-2. 121

Figure 83 : Structures représentatives de chaque amas de conformation de la trajectoire de Sgal dans l'eau et sous contrainte de distance. 122

Figure 84 : Coudes β présents au cours du temps pour la simulation de Sgal dans le TFE. 125

Figure 85 : Zones de la carte de Ramachandran occupées par les résidus de Sgal dans le TFE. 125

Figure 86 : Carte de RMSD croisés de Sgal dans le TFE. 126

Figure 87 : Amas de conformations au cours du temps calculés par ART-2. 127

Figure 88 : Structures représentatives de chaque amas de conformations calculés par ART-2 au cours de la trajectoire de Sgal dans le TFE. 127

Figure 89 : Spectres de DCE du peptide de 32 résidus dans l'eau. 129

Figure 90 : Spectres COSY et TOCSY du peptide de 32 résidus. 131

Figure 91 : Séquence de la tropoélastine humaine. 134

Figure 92 : Pourcentages moyens de conformations contenant un coude β sur GX_2X_3P. 142

Figure 93 : Structures MC dans l'eau d'hexapeptides actifs. 150

Figure 94 : Structures MC dans le cyclohexane d'hexapeptides actifs. 150

Figure 95 : Structures MC dans l'eau et le cyclohexane des GLVPG et GVAPG. 151

Figure 96 : Structures MC dans l'eau et le cyclohexane du peptide inactif GGVPG. 151

Figure 97 : Structures représentatives des amas de conformations de VGVAPG. 154

Figure 98 : Structures représentatives des amas de conformations de PGAIPG. 155

Figure 99 : Structures représentatives des amas de conformations de GVAPG. 155

Figure 100 : Structures représentatives des amas de conformations de GLVPG. 156

Figure 101 : Structures représentatives des amas de conformations de GGVPG. 156

Figure 102 : Types de coudes β au cours de la DM de VGVAPG dans l'eau. 159

Figure 103 : Hélice polyproline II (sur VAP) de VGVAPG dans l'eau. 159

Figure 104 : Carte des RMSD croisés de VGVAPG dans l'eau. 160

Figure 105 : Angles dièdres ϕ et ψ de V_3 dans la DM de VGVAPG. 160

Figure 106 : ΔG du coude β-VIII sur GVAP dans la DM de VGVAPG dans l'eau. 161

Figure 107 : Types de coudes β au cours de la DM de VGVAPG dans le TFE. 162

Figure 108 : Carte des RMSD croisés du peptide VGVAPG dans le TFE. 162

Figure 109 : Liaison hydrogène au cours de la DM de VGVAPG dans le TFE. 163

Figure 110 : ΔG du coude β-VIII sur GVAP dans la DM de VGVAPG dans le TFE. 163

Figure 111 : Histogrammes des valeurs de ϕ durant la DM de VGV_3A_4PG dans l'eau. 165

Figure 112 : Histogrammes des valeurs de ϕ durant la DM de VGV_3A_4PG dans le TFE. 165

Figure 113 : Types de coudes β au cours de la DM du peptide PGAIPG dans l'eau. 167
Figure 114 : Hélice PP II sur AIP au cours de la DM de PGAIPG dans l'eau. 167
Figure 115 : Carte des RMSD croisés du peptide PGAIPG dans l'eau. 168
Figure 116 : ΔG du coude β -VIII sur GAIP dans la DM de PGAIPG dans l'eau. 168
Figure 117 : Types de coude β au cours de la DM de PGAIPG dans le TFE. 169
Figure 118 : Carte de RMSD croisés du peptide PGAIPG dans le TFE. 170
Figure 119 : Liaison hydrogène au cours de la DM de PGAIPG dans le TFE. 170
Figure 120 : ΔG du coude β-VIII sur GAIP dans la DM de PGAIPG dans le TFE. 171
Figure 121 : Types de coudes β au cours de la DM de GGVPG dans l'eau. 172
Figure 122 : Carte des RMSD croisés du peptide GGVPG dans l'eau. 172
Figure 123 : Distance $C_\alpha(G_1)$-$C_\alpha(P_4)$ dans la DM du peptide GGVPG dans l'eau. 173
Figure 124 : ΔG du coude GGVP dans la DM de GGVPG dans l'eau. 173
Figure 125 : Types de coude β au cours de la DM de GGVPG dans le TFE. 174
Figure 126 : Carte des RMSD croisés du peptide GGVPG dans le TFE. 175
Figure 127 : Distance $C\alpha(G_1)$-$C\alpha(P_4)$ dans la DM du peptide GGVPG dans le TFE. 175
Figure 128 : ΔG du coude GGVP dans la DM de GGVPG dans le TFE. 176
Figure 129 : Histogramme de ϕ (V_3) durant la DM de GGVPG dans le TFE 177
Figure 130 : Spectre Raman normal du peptide cyclique C-VGVAPG-C. 179
Figure 131 : Spectre Raman normal du peptide cyclique C-VGVAPG-C dans le domaine caractéristique des ponts disulfures. 180
Figure 132 : Domaines constituant le peptide de fibrilline 1. 183
Figure 133 : Influence des peptides recombinants sur la MMP-1. 184
Figure 134 : Western blot de la MMP-1 obtenus après stimulation. 184
Figure 135 : Domaine cbEGF de la fibrilline 1. 185
Figure 136: Types de coudes β présents dans les DM T1 et T2. 188
Figure 137 : Structures des peptides EGFEPE et EGFESG 189
Figure 138 : ΔG en fonction du temps : de EGFEPG et EGEFSG. 190
Figure 139 : Différence de positionnement de la liaison peptidique entre E et G. 192
Figure 140 : Différence de position de la liaison peptidique située entre G et X. 194
Figure 141 : Coude β de type VIII. 194
Figure 142 : Représentation de la PPE (1GVK). 196
Figure 143 : Site actif de la PPE (1GVK) 197
Figure 144 : Localisation du meilleur candidat proposé par Autodock 199
Figure 145 : Suivi des structures secondaires en fonction du temps 201
Figure 146 : Présence des différents types de coudes β 202

Figure 147 : Zones explorées par les 15 premiers résidus de la DM de PPE seule. 203

Figure 148 : Suivi des structures secondaires au cours de la DM de la PPE en interaction avec le peptide VGVAPG. 204

Figure 149 : Présence des coudes β au cours de la DM de la PPE en interaction avec le peptide VGVAPG. 205

Figure 150 : Zones explorées par les 15 premiers résidus de la PPE au cours de la DM de la PPE en interaction avec le peptide VGVAPG 206

Figure 151 : Carte des RMSD croisés durant la DM de VGVAPG en interaction avec la PPE. 206

Figure 152 : Zones explorées par les résidus au cours de la DM de VGVAPG en interaction avec PPE. 207

Figure 153 : Distance C_α (G_2) - C_α (P_5) du peptide VGVAPG, au cours de sa DM en interaction avec la PPE. 208

Figure 154 : Liaison hydrogène entre VGVAPG et la PPE au cours de la DM. 208

Figure 155 : Spectre DCE de l'interaction entre le peptide Sgal et $(VGVAPG)_3$. 210

Figure 156 : Spectre DCE de l'interaction entre le peptide Sgal et VGVAPG. 210

Figure 157 : Graphique de Scatchard pour le complexe VGVAPG - peptide Sgal. 211

Figure 158 : Graphique de Scatchard pour le complexe YGVG - peptide Sgal. 212

Figure 159 : Meilleurs candidats proposés par Autodock pour l'interaction entre le peptide Sgal et a) VGVAPG ou b) YGVG. 213

Figure 160 : Présence des liaisons hydrogène entre VGVAPG et le peptide Sgal au cours de la DM de 20 ns dans l'eau. 215

Figure 161 : Présence des différents types de coude β du peptide Sgal en interaction avec le peptide VGVAPG au cours de la DM de 20 ns dans l'eau. 216

Figure 162 : Carte de RMSD croisés du peptide Sgal durant sa DM en interaction avec VGVAPG. 217

Figure 163 : Présence des différents types de coudes β du peptide VGVAPG en interaction avec le peptide Sgal au cours de la DM de 20 ns dans l'eau. 218

Figure 164 : Carte de RMSD croisés du peptide VGVAPG durant sa DM en interaction avec le peptide Sgal. 219

Liste des Tableaux

Tableau 01 : Valeurs canoniques des angles dièdres des différents types de coudes β. 11

Tableau 02 : Longueur d'excitation et d'émission des fluorophores utilisés. 47

Tableau 03 : Angles dièdres des 10 régions définies pour la carte adiabatique. 63

Tableau 04 : Qualité de prédiction des structures secondaires. 80

Tableau 05 : Epitopes de l'EBP prédits par PEOPLE. 83

Tableau 06 : Déplacements chimiques des protons de Sgal dans l'eau. 86

Tableau 07 : Présence des coudes β dans les différentes conformations. 95

Tableau 08 : Présence de coudes β dans les différentes conformations obtenues après simulations de Monte-Carlo dans le cyclohexane. 97

Tableau 09 : Conditions de simulation des dynamiques moléculaires de Sgal. 101

Tableau 10 : Pourcentages de coudes β observés dans la structure de Sgal, lors de la simulation de 25 ns dans l'eau. 102

Tableau 11 : Pourcentages des zones de la carte de Ramachandran occupées par les résidus de Sgal pour la simulation de 25 ns dans l'eau. 103

Tableau 12 : Coudes β des amas de conformations proches au cours de la trajectoire de 25 ns de Sgal dans l'eau. 105

Tableau 13 : Pourcentages de coudes β observés dans la structure de Sgal lors de la simulation de 30 ns dans l'eau. 107

Tableau 14 : Pourcentages des zones de la carte de Ramachandran occupées par les résidus de Sgal pour la simulation de 30 ns dans l'eau. 109

Tableau 15 : Coudes β des amas de conformations proches, au cours de la trajectoire de 30 ns de Sgal dans l'eau. 111

Tableau 16 : Pourcentages de coudes β observés dans la structure de Sgal lors de la simulation de 15 ns dans l'eau.

Tableau 17 : Pourcentages des zones de la carte de Ramachandran occupées par les résidus de Sgal pour la simulation de 15 ns dans l'eau. 113

Tableau 18 : Coudes β des amas de conformations proches au cours de la trajectoire de 15 ns de Sgal dans l'eau. 115

Tableau 19 : Pourcentages des types de coude β de Sgal dans l'eau avec une contrainte de distance. 118

Tableau 20 : Pourcentages des zones de la carte de Ramachandran occupées par les résidus de Sgal dans l'eau et sous contrainte de distance. 120

Tableau 21 : Coudes β des amas de conformations proches au cours de la trajectoire de Sgal dans l'eau et sous contrainte de distance. 122

Tableau 22 : Pourcentages des types de coudes β de Sgal dans le TFE. 124

Tableau 23 : Pourcentages des zones de la carte de Ramachandran occupées par les résidus de Sgal dans le TFE. 126

Tableau 24 : Coudes β des amas de conformations au cours de la trajectoire de Sgal dans le TFE. 128

Tableau 25 : Pourcentages des structures secondaires estimées par CDPRO du peptide de 32 résidus dans l'eau pour différentes températures. 130

Tableau 26 : Conditions des simulations de dynamique moléculaire. 136

Tableau 27 : Pourcentages des différents coudes β sur la séquence $X_1GX_2X_3$. 137

Tableau 28 : Pourcentages des différents coudes β sur la séquence $X_2X_3PX_4$. 138

Tableau 29 : Pourcentages des différents coudes β sur la séquence GX_2X_3P dans 139

Tableau 30 : Pourcentages de conformations en hélice polyproline II. 144

Tableau 31 : Pourcentages de coudes β sur le motif GX_2X_3P et de polyproline II. 147

Tableau 32 : Types de coudes β des structures MC des peptides contenant GXXP. 152

Tableau 33 : Conditions de simulation en solvant explicite des EDPs. 158

Tableau 34 : Pourcentages des types de coudes β de VGVAPG dans l'eau. 158

Tableau 35 : Pourcentages des types de coudes β de VGVAPG dans le TFE. 161

Tableau 36 : Résultats obtenus par RMN pour VGVAPG suivant le solvant utilisé. 164

Tableau 37 : Pourcentages des types de coudes β du peptide PGAIPG dans l'eau. 166

Tableau 38 : Pourcentages de coudes β du peptide PGAIPG dans le TFE. 169

Tableau 39 : Pourcentages des types de coudes β du peptide GGVPG dans l'eau. 172

Tableau 40 : Pourcentages des types de coudes β du peptide GGVPG dans le TFE. 174

Tableau 41 : Protéines de la MEC contenant au moins 3 motifs GXXPG. 181

Tableau 42 : Types de coudes β observés pour les peptides EGFEPG et EGFESG. 187

Tableau 43 : ΔG à convergence pour les quatre simulations de DM. 191

Tableau 44 : Distance entre certains atomes de l'inhibiteur et de la PPE 198

Tableau 45 : Caractéristiques des meilleures solutions proposées par Autodock 199

Tableau 46 : Pourcentages de coudes β sur les séquences QRNS, NSWP et WPSQ au cours de la DM de la PPE seule. 202

Tableau 47 : Pourcentages de coudes β sur les séquences VVGG, QRNS, NSWP et WPSQ au cours de la DM de la PPE en interaction avec le peptide VGVAPG. 205

Tableau 48 : Pourcentages de coudes β du peptide Sgal en interaction avec VGVAPG au cours de la DM de 20 ns dans l'eau. 216

Tableau 49 : Pourcentages de coudes β observés dans la structure du peptide VGVAPG en interaction avec le peptide Sgal au cours de la DM de 20 ns dans l'eau. 217

Liste des Abréviations

ABNR	Adopted Basis Newton-Raphson
ADT	AutoDock Tools
CHARMM	Chemistry at HARvard of Macromolecular Mechanics
CLHP	Chromatographie Liquide Haute Pression
COSY	Correlation SpectroscopY
DCE	Dichroïsme Circulaire Electronique
DCC	DiCyclohexylCarbodiimide
DM	Dynamique Moléculaire
DMAP	4-DiMéthylAminoPyridine
DSSP	Define Secondary Structure of Proteins
EBP	Elastin Binding Protein
EDP	Elastin Derived Peptide
FANTOM	Fast Newton-Raphson Torsion Minimizer
FID	Free Induction Decay
GBMV	Generalized Born using Molecular Volume
GC	Gradient Conjugué
GOR	Garnier Ogusthorpe Robson
HOBt	1-HydrOxylBenzotriazole
HSQC	Heteronuclear Single Quantum SpectroscopY
K_D	Constante de dissociation
MAGP	Microfibrille Associated GlycoProtein
MC	Monte-Carlo
MEC	Marice ExtraCellulaire
MMP	Matrix MetalloProteinases
MOLMOL	MOLecule analysis and MOLecule display
NOESY	Nuclear Overhauser Effect SpectroscopY
PDB	Protein Data Bank
PEOPLE	Predictive Estimation Of Protein Linear Epitope
PP II	PolyProline II
PPE	Porcine Pancreatic Elastase
RMSD	Root Mean Square Deviation

ROESY	Rotating Overhauser Effect SpectroscopY
SASA	Solvant Accessible Surface Area
SD	Steepest Descent
SPPS	Solid-Phase Peptide Synthesis
TFA	TriFluoroAcetic acid.
TFE	2,2,2-TriFluoroEthanol
TOCSY	Total Correlation SpectroscopY
VMD	Visual Molecular Dynamics

Partie I

INTRODUCTION

A. INTRODUCTION GENERALE

Les êtres vivants sont principalement constitués de quatre classes de biomolécules. Les acides nucléiques, tout d'abord, contiennent l'information génétique des organismes permettant leur reproduction. Les glucides, ensuite, servent de réservoir d'énergie et d'intermédiaires métaboliques. Les lipides sont également des réserves d'énergie mais forment aussi la majorité des constituants des membranes biologiques. Enfin, les protéines, constituées d'acides aminés, jouent des rôles cruciaux dans pratiquement tous les processus biologiques : la catalyse enzymatique, le transport et la mise en réserve des petites molécules et d'ions, la protection immune, le contrôle de la croissance et de la différenciation des cellules et peuvent également être un support mécanique dans la peau et les os.

Les acides aminés, éléments constitutifs des peptides et protéines, ne sont qu'au nombre de vingt, mais leur enchaînement spécifique (séquence = structure primaire) et leur nombre confèrent une structure tridimensionnelle essentielle à la fonction biologique. L'étude de cette relation entre la structure et la fonction est donc primordiale pour la compréhension des phénomènes biologiques mettant en jeu les protéines. Les peptides, molécules de petite taille d'au maximum 50 acides aminés, sont extrêmement flexibles. Cette grande flexibilité leur permet donc d'adopter plusieurs conformations, c'est pour cela qu'il est obligatoire de réaliser l'étude de leur relation structure / fonction / dynamique.

L'espace qui entoure les cellules, appelé espace extracellulaire, contient un ensemble de macromolécules, polysaccharides ou glycosaminoglycanes, protéines, sels et eau, que l'on désigne sous le nom de matrice extracellulaire. Cette matrice extracellulaire joue un rôle essentiel dans la constitution, le maintien et le remodelage de l'architecture tissulaire. Le collagène et l'élastine sont les principales protéines qui la composent.

L'élastine est l'une des protéines majoritairement présente dans les tissus soumis à de fortes contraintes mécaniques comme le sont les poumons, la peau ou les vaisseaux sanguins. Elle possède des propriétés élastiques indispensables au bon fonctionnement de ces tissus. D'une manière étonnante, l'élastine, est totalement insoluble dans l'eau alors que la présence d'eau est indispensable pour que les propriétés élastiques de l'élastine s'expriment.

La structure tridimensionnelle de la tropoélastine, monomère précurseur de l'élastine tout comme la manière dont elle s'agence en fibres élastiques restent pour l'instant très peu connues.

Sous l'action de certaines élastases qui sont des enzymes protéolytiques, l'élastine est dégradée et donne ainsi lieu à la libération de peptides de taille variable au sein de la matrice extracellulaire.

Ces peptides dérivant de l'élastine, sont la cause de nombreux processus biologiques d'importance, leur donnant ainsi un rôle majeur dans les phénomènes d'inflammation, de vieillissement tissulaire, tout comme dans la migration tumorale. Parmi les peptides dérivant de l'élastine, ceux qui contiennent une séquence **GXXP** (X étant un acide aminé quelconque), ont l'origine de leurs activités biologiques provenant de leur interaction avec un récepteur spécifique appelé récepteur de l'élastine.

Le récepteur de l'élastine est localisé à la surface des membranes plasmiques des cellules. Il est constitué de trois sous-unités : une neuraminidase, qui permet la transmission du signal intracellulaire, une protéine protectrice, qui stabilise le complexe, et une dernière protéine responsable de la fixation des peptides d'élastine qui est appelée EBP, pour Elastin Binding Protein. L'EBP permet également au cours de l'embryogenèse, par son interaction avec la tropoélastine, de l'amener à la surface des cellules pour être incorporée dans les fibres élastiques durant l'élastogenèse.

La structure tridimensionnelle du récepteur de l'élastine, tout comme l'agencement de ses sous-unités ou encore la nature de son interaction avec la membrane plasmique est pour l'instant totalement inconnue. En effet, sur les trois sous-unités qui composent le récepteur de l'élastine, seule la protéine protectrice a vu sa structure tridimensionnelle résolue, les deux autres n'ont que leur séquence polypeptidique connue.

Il a cependant été démontré que la séquence responsable de l'interaction entre l'EBP, qui a en fait le véritable rôle de récepteur du complexe, et ses ligands est constituée de quinze résidus. Ces quinze résidus sous forme de peptide, appelé peptide Sgal, conservent les propriétés de reconnaissance et de fixation vis-à-vis de la tropoélastine.

Le but de cette étude est :

- dans un premier temps, d'étudier la relation structure-fonction qui peut exister chez le peptide Sgal ainsi que chez les ligands de l'EBP, c'est-à-dire les peptides dérivant de l'élastine et possédant en particulier la séquence **GXXP**.
- dans un deuxième temps, de comprendre la nature et les conformations nécessaires à l'interaction entre le peptide Sgal ou l'EBP et les peptides dérivant de l'élastine.

L'objectif principal de ce travail est donc de comprendre et de caractériser structuralement l'interaction entre ces protéines, pour ensuite tenter d'inhiber de façon spécifique le rôle délétère du récepteur de l'élastine dans certains processus biologiques.

En combinant les approches expérimentales de biophysiques (spectroscopies de résonance magnétique nucléaire, de dichroïsme circulaire, de diffusion Raman) et les approches théoriques de bioinformatiques (stratégie 1D - 4D, modélisations moléculaires : Monte-Carlo, simulations de mécanique et dynamique moléculaire), nous avons essayé de caractériser la structure que peut adopter le peptide Sgal, libre en solution.

Nous avons ensuite étudié *in silico* les conformations échantillonnées par les peptides contenant la séquence **GXXP** de la tropoélastine humaine et les conséquences structurales qu'apportent les acides aminés placés de part et d'autre de ce motif.

Nous nous sommes particulièrement intéressés à certain de ces peptides, contenant la séquence **GXXP**, dont l'activité ou l'inactivité biologique est expérimentalement connue, c'est-à-dire ayant la capacité de se lier ou non à l'EBP.

Enfin, nous avons vérifié expérimentalement, ce qui n'avait jamais été montré auparavant, que le peptide Sgal est également suffisant pour se lier à un peptide contenant **GXXP**.

A partir des connaissances structurales acquises sur le peptide Sgal et des peptides dérivant de l'élastine (contenant **GXXP)**, nous avons tenté d'appréhender et de comprendre les mécanismes d'interaction en utilisant successivement des méthodes d'amarrage moléculaire et des méthodes de simulation de dynamique moléculaire afin de tester la stabilité des complexes proposés.

Le manuscrit s'organise ainsi en sept parties principales.

Partie I, de nature introductive : donne une rapide description des structures que peuvent prendre les peptides et les protéines, puis rappelle le contexte biologique qui entoure ce travail.

Partie II, de nature descriptive : aborde les diverses méthodes expérimentales et théoriques qui ont été utilisées durant cette étude.

Parties III - IV - V et VI, de nature descriptive et analytique : consacrées à la présentation et la discussion de l'ensemble des résultats obtenus.

- Partie III : présente l'étude du récepteur EBP et concerne essentiellement la structure du peptide Sgal, site de liaison du récepteur EBP, à l'élastine.

- Partie IV : porte sur ligands peptidiques de l'EBP, que sont essentiellement les peptides dérivés de l'élastine (EDP).

- Partie V : aborde la manière dont peuvent se faire les interactions ligands-Sgal.

- Partie VI : conclut sur les principaux résultats et expose les perspectives que l'on peut envisager de donner à ces travaux de thèse.

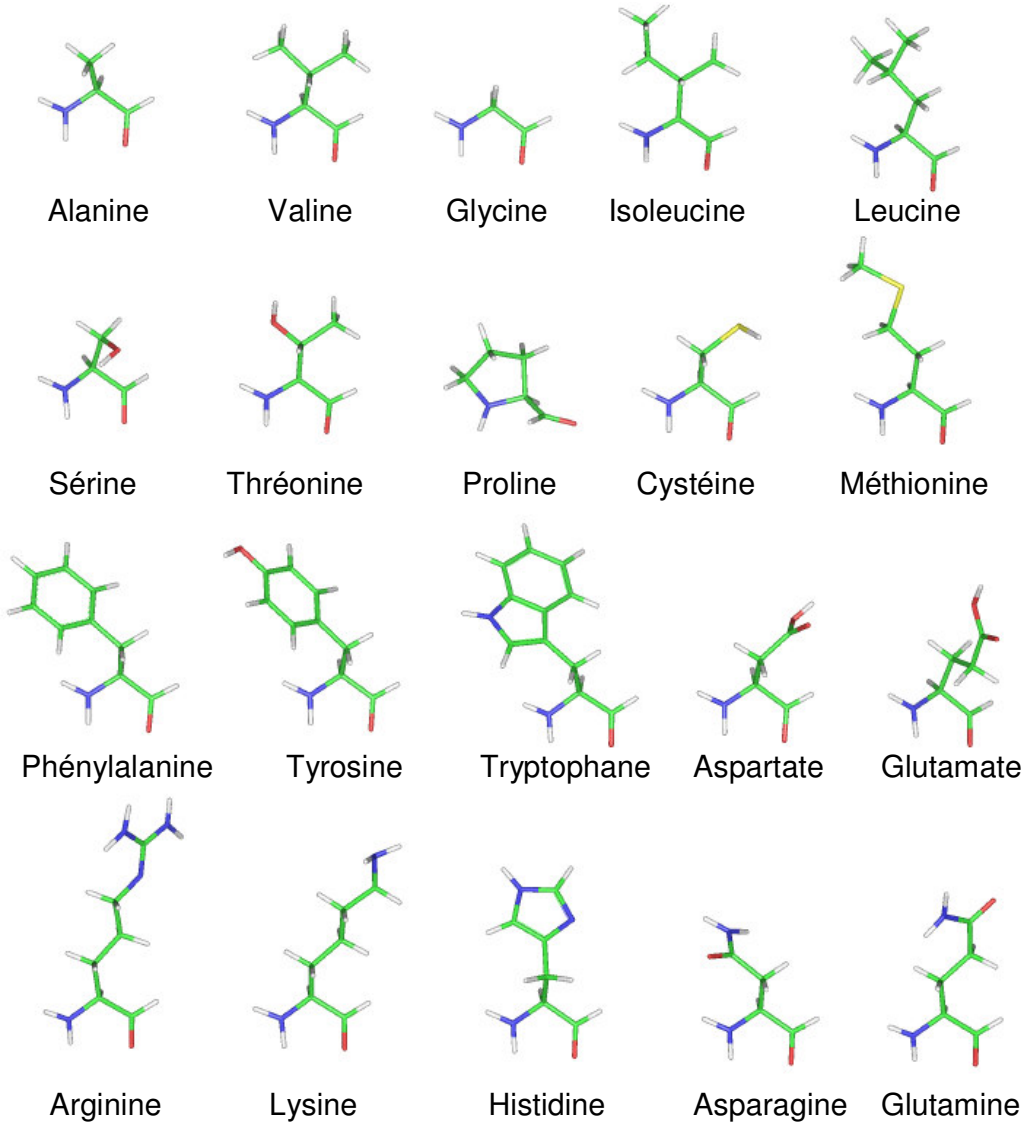
Partie VII : donne une bibliographie extensive pour l'ensemble du manuscrit.

B. STRUCTURE DES PROTEINES

Les molécules constituant les êtres vivants peuvent être classées en quatre grandes familles: les lipides, les carbohydrates, les acides nucléiques et les protéines. D'un point de vue chimique, une protéine est un hétéropolymère linéaire, c'est-à-dire qu'elle est le résultat de l'enchaînement linéaire de monomères : les acides aminés ou résidus. La structure d'une protéine peut être décrite à plusieurs niveaux de structure.

B.1. Structure primaire

La structure primaire correspond à la séquence proprement dite de la protéine, c'est-à-dire à l'enchaînement des acides aminés par la liaison peptidique. On dénombre pas moins de 20 acides aminés qui ne diffèrent les uns des autres que par leur chaîne latérale (groupement R) portée par le C_α (**Figure 01**). On parle de peptide lorsque le nombre de résidus formant la molécule est inférieur à 50 et de protéine au delà de 50 résidus.



<u>Figure 01</u> : Structure des 20 acides aminés composant les protéines.

B.2. Structure secondaire

La structure secondaire d'une protéine correspond aux repliements qu'adoptent les portions partielles de la séquence protéique, caractérisées par les angles dièdres ω, ϕ, ψ de leurs résidus (**Figure 02**) et par la présence de certaines liaisons hydrogène.

Dans les protéines, les acides aminés sont reliés entre eux par une liaison peptidique **O=C–N–H** caractérisée par 2 formes imposant ainsi à l'angle de torsion de la liaison **C-N** (angle dièdres ω) deux valeurs particulières (**Figure 03**) :

0° : liaison peptidique *cis* : son taux d'occurrence est très faible. C'est avec la proline que l'on a le plus de chance d'observer une liaison peptidique *cis* : jusqu'à 10 % de l'ensemble des prolines (Fischer et al. 1994).

180° : liaison peptidique *trans* : c'est la forme la plus stable correspondant à la forme quasiment exclusivement observée dans les protéines.

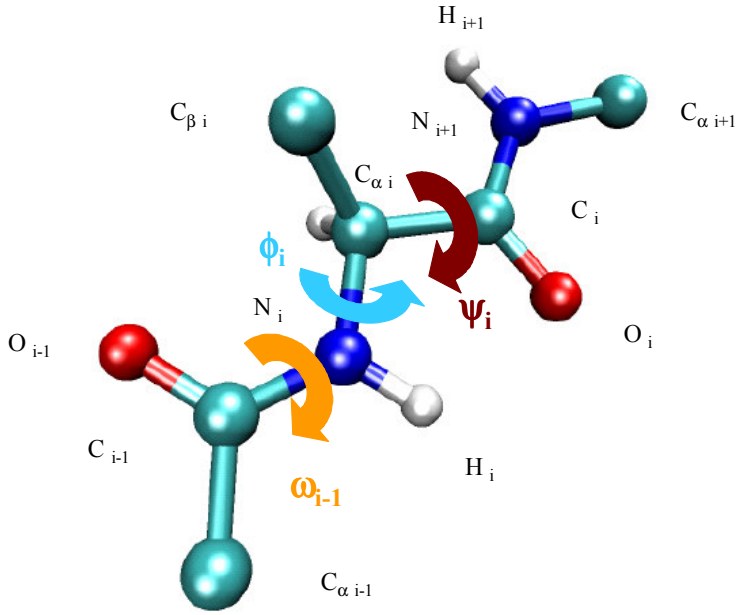


Figure 02 : Définition des Angles dièdres ω, ϕ et ψ.

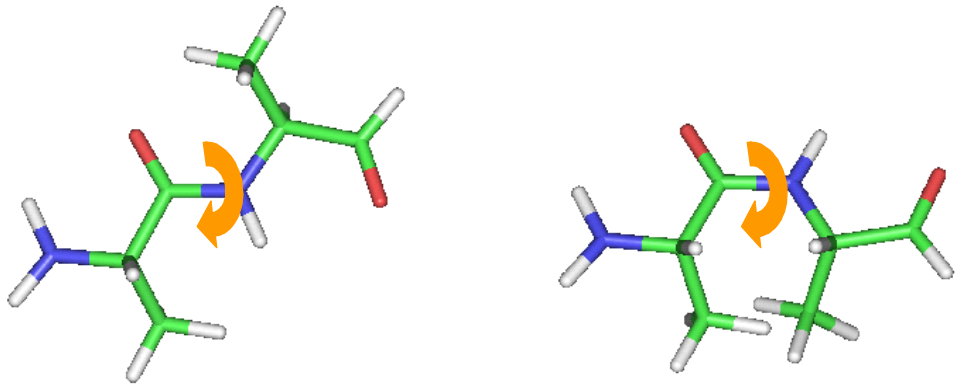
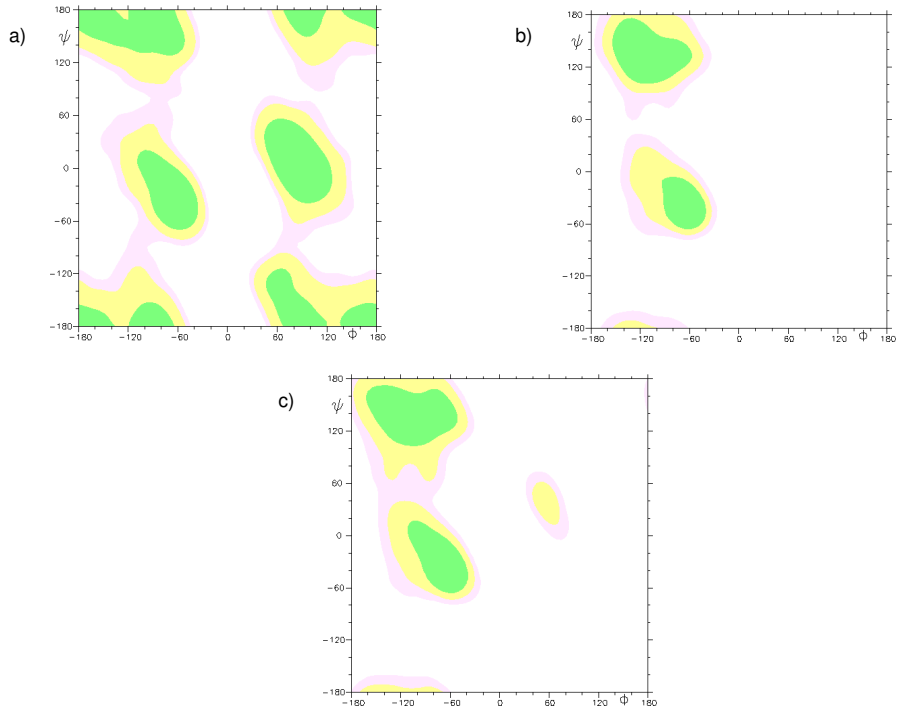


Figure 03 : Liaisons peptidiques *trans* (ω = 180°) et *cis* (ω = 0°).

Puisque la grande majorité des liaisons peptidiques sont en conformation trans (ω = 180°), il est donc possible de décrire la structure d'une protéine à partir de ses seuls angles dièdres ϕ et ψ. L'angle dièdre ϕ correspond à l'angle de torsion de la liaison **N-C_α** et ψ, à l'angle de torsion de la liaison **C_α-C** (**Figure 02**). Ramachandran (Ramachandran et al. 1963) a été le premier à

proposer d'utiliser la valeur de ϕ en abscisse et celle de ψ en ordonnée pour tracer une carte 2D, depuis appelée carte de Ramachandran (**Figure 04**), sur laquelle chacun des résidus de la protéine est représenté par un point.



<u>Figure 04</u> : Cartes de Ramachandran de la glycine (a), de la valine, de l'isoleucine et de la thréonine (b) et des autres résidus (c), d'après le logiciel MOLMOL (Koradi et al. 1996).

Ce type de représentation permet ainsi de voir le comportement différent que peut adopter un résidu. La glycine dont la chaîne latérale est un simple atome d'hydrogène, ne connaît pratiquement pas de restriction de valeurs d'angles dièdres : une très grande partie de la carte de Ramachandran est explorée (**Figure 04a**). A l'opposé, la proline, seul acide aminé cyclique, ne possède que deux régions véritablement favorables du point de vue énergétique. Des analyses, effectuées au sein d'une banque de structures de protéines, sur la conformation des prolines nous apprend que 44 % des prolines se situent dans la zone ϕ = - 61°, ψ = - 35° et 56 % dans la zone ϕ = - 65°, ψ = + 150° (MacArthur et al. 1991).

Les structures secondaires ordonnées des protéines peuvent être classées en trois catégories : les hélices, les brins et les coudes. Lorsqu'une portion de protéine ne satisfait à aucun des critères caractérisant les structures précédentes, on est en présence de structure secondaire non ordonnée appelée boucle ou pelote statistique (random coil).

B.2.a. Les hélices

Une structure en hélice est définie comme étant une portion de la chaîne peptidique s'enroulant en spirale. Il en existe plusieurs types selon le nombre de résidus par tour d'hélice **(Figure 05**).

L'hélice α est l'élément de structure secondaire le plus courant. C'est une hélice enroulée à droite constituée de 3,6 résidus par tour. C'est une structure très stable grâce à la présence de liaisons hydrogène entre le **C=O** du résidu *i* et le **N-H** du résidu *i+4*. Ses valeurs d'angles dièdres sont : $\phi = -57°$ et $\psi = -47°$.

L'hélice α_G (gauche) est stabilisée par les mêmes types de liaisons hydrogène, mais elle possède un sens d'enroulement gauche, donnant ainsi des valeurs d'angles dièdres inversées : $\phi = +57°$ et $\psi = +47°$. Cette zone de la carte de Ramachandran n'est ainsi accessible qu'aux résidus glycines, ce qui explique la rareté de cette structure au sein des protéines.

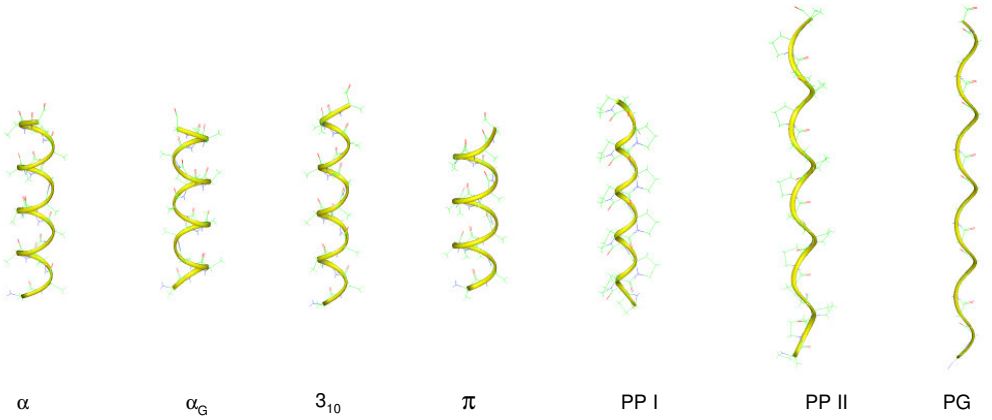
L'hélice 3_{10} est moins compacte car elle ne possède que 3 résidus par tour. Elle est moins stable et donc beaucoup moins fréquemment observée que l'hélice α. Elle est stabilisée par des liaisons hydrogène entre les résidus *i* et *i+3*, et correspond en fait à une succession de coudes $\beta 3_{10}$ (3 résidus et 10 atomes impliqués).

Les angles dièdres des résidus la constituant ont pour valeur : $\phi = -49°$ et $\psi = -29°$.

L'hélice π est, elle, plus compacte que l'hélice **α** avec 4,4 résidus par tour, des liaisons hydrogène entre les résidus *i* et *i+5* pouvant être présentes.

Elle possède des angles dièdres de valeur : $\phi = -57°$ et $\psi = -70°$.

Il existe également d'autres types de structures hélicoïdales qui s'organisent en structures régulières mais qui, selon différents auteurs, sont ou ne sont pas considérées comme des structures secondaires.



<u>Figure 05</u> : Structures secondaires en hélices (PP : PolyProline; PG : PolyGlycine).

Il a été observé que la succession de résidus prolines pouvait prendre une forme hélicoïdale particulière qui a naturellement été nommée polyproline, bien qu'elle puisse être constituée par d'autres résidus que la proline.

Il existe deux types de structure polyproline :

La structure polyproline I, nécessite une liaison peptidique *cis*. C'est une hélice enroulée à droite constituée de 3,3 résidus par tour.

Les valeurs des angles dièdres sont : ϕ = - 83° et ψ = + 149°.

La structure polyproline II nécessite une liaison peptidique *trans*. C'est une hélice enroulée à gauche constituée de 3 résidus par tour.

Les valeurs des angles dièdres sont ϕ = - 78° et ψ = + 149°.

Sous forme de peptide, il a été montré que la polyproline II était stable dans l'eau (Sreerama et al. 1999), peut-être grâce à des liaisons hydrogène intra-chaînes formées par l'intermédiaire de molécules d'eau.

Sa présence a notamment été démontrée dans certaines protéines fibreuses telles que le collagène, qui est l'un des composants majoritaires de la matrice extracellulaire, et qui résulte de l'association de trois chaînes peptidiques en structure polyproline II.

De même, il a été montré que cette structure est présente dans l'élastine d'après des études spectroscopiques en RMN et dichroïsme circulaire (Tamburro et al. 2003, Bochicchio et al. 2004b), en diffusion Raman (Bochicchio et al. 2004b) de certains de ses peptides.

Enfin la structure **polyglycine**, essentiellement retrouvée dans des domaines riches en glycine, est une structure très proche de la structure polyproline. Elle est caractérisée par des valeurs d'angles dièdres suivant : ϕ = - 80° et ψ = + 150°.

B.2.b. Les brins β

Le brin β est le deuxième type de structure secondaire ordonnée retrouvé dans les protéines. C'est une structure proche des structures étendues avec des valeurs standards d'angles dièdres ϕ = - 120° et ψ = + 120°.

Deux ou plusieurs brins β peuvent s'associer en feuillet β grâce à des liaisons hydrogènes caractéristiques ; il en existe, entre autres, deux types caractéristiques (**Figure 06**) :

- le feuillet β anti-parallèle dont l'orientation des brins successifs est inversée avec des angles dièdres ϕ = - 139° et ψ = + 135°.

- le feuillet β parallèle dont l'orientation des brins successifs est la même, avec des angles dièdres $\phi = -119°$ et $\psi = +113°$.

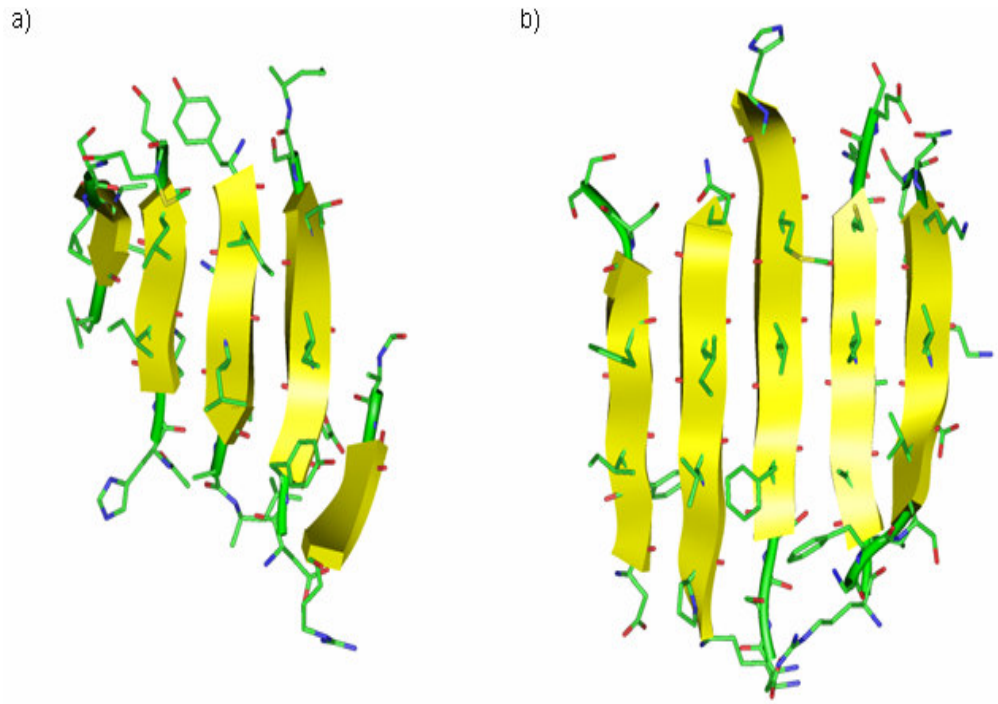


Figure 06 : Feuillets β anti-parallèles (a)et parallèles (b).

B.2.c. Les coudes

Les coudes permettent à la chaîne polypeptidique d'effectuer un changement de direction à 180°. Ils sont classés selon le nombre de résidus impliqués dans ce changement de direction : on parle de coude δ pour 2 résidus, de coude γ pour 3 résidus (le coude $\beta 3_{10}$ est un cas très particulier de coude à 3 résidus, dont la répétition peut engendrer une hélice $\alpha 3_{10}$), de coude β pour 4 résidus, de coude α pour 5 résidus et de coude π pour 6 résidus (Chou 2000).

Les coudes γ et β sont les plus couramment mis en évidence dans les structures des protéines. Ils sont en particulier souvent observés dans la portion peptidique reliant deux brins β formant un feuillet β anti-parallèle.

Différents types de coudes ont été définis à partir de valeurs précises des angles dièdres φ et ψ. Ainsi un coude γ est dit de type **classique** pour des valeurs d'angles dièdres du résidu central $\phi = +75°$ et $\psi = -64°$ et de type **inverse** pour $\phi = -79°$ et $\psi = +69°$ (Milner-White et al. 1988).

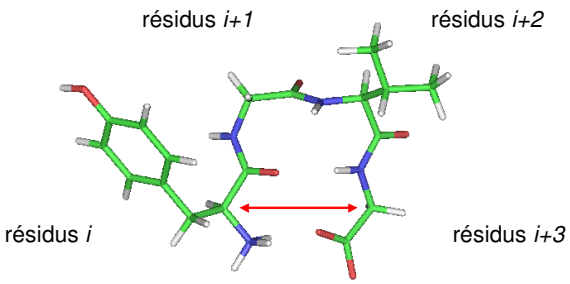


Figure 07 : Coude β.

Les coudes β, caractérisés par une distance inférieure à 7 Å entre les C_α des résidus terminaux (i et i+3) (**Figure 07**), sont le plus souvent stabilisés par une liaison hydrogène entre le **C=O** et le **N-H** de ces résidus terminaux.

Ils sont classés en sept types réguliers : I, I', II, II', VI (a1, a2) VIb et VIII, selon les valeurs des angles dièdres ϕ et ψ des résidus centraux *i+1* et *i+2* (**Tableau 01**).

Types	ϕ_{i+1}	ψ_{i+1}	ϕ_{i+2}	ψ_{i+2}	ω_{i+2}
I	– 60°	– 30°	– 90°	0°	180°
I'	60°	30°	90°	0°	180°
II	– 60°	120°	80°	0°	180°
II'	60°	– 120°	– 80°	0°	180°
VIII	– 60°	– 30°	– 120°	120°	180°
VIa1	– 60°	120°	– 90°	0°	0°
VIa2	– 120°	120°	– 60°	0°	0°
VIb	– 135°	135°	– 75°	160°	0°
IV	–	–	–	–	–

Tableau 01 : Valeurs canoniques des angles dièdres des différents types de coudes β.

Il est à noter que les coudes β de type VI ne sont présents que lorsque le résidu i+2 possède une liaison peptidique *cis* (ω = 0°) donc essentiellement lorsque ce résidu est une proline.

En pratique, pour caractériser un type de coude β, on accorde une tolérance de ± 30° sur trois des angles et une tolérance de ± 45° sur le quatrième (Hutchinson et al. 1994). Quand seul le critère de distance inférieure à 7 Å entre le $C\alpha_i$ et $C\alpha_{i+3}$ est respecté, c'est-à-dire quand la valeur des angles dièdres ne permet pas, malgré la tolérance accordée, de pouvoir le classer dans un des types précédemment définis, alors ce coude β sera classé type IV.

B.2.d. Les boucles

Le terme boucle ou pelote statistique (random coil) est utilisé lorsque aucune autre structure secondaire ordonnée ne peut être affectée à une conformation locale. Mais il faut cependant bien garder à l'esprit que cela ne signifie pas l'absence d'une structure stable ou d'une structure particulière.

B.3. Structure tertiaire

La structure tertiaire des protéines correspond au repliement et à l'assemblage des différents éléments de structures secondaires. La structure tertiaire correspond en fait à la structure tridimensionnelle (structure 3D) de la protéine (**Figure 08**). Ce sont des interactions non liantes de type électrostatique et de van der Waals, ainsi que les ponts salins, les liaisons hydrogènes et les ponts disulfures qui permettent de stabiliser ce type de structure.

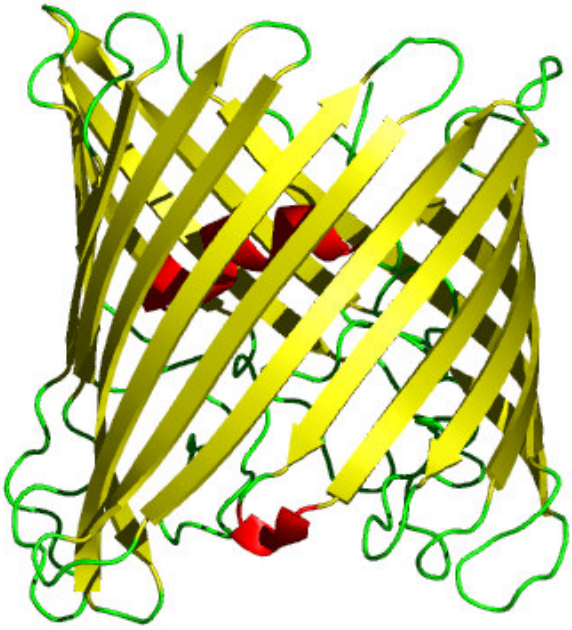
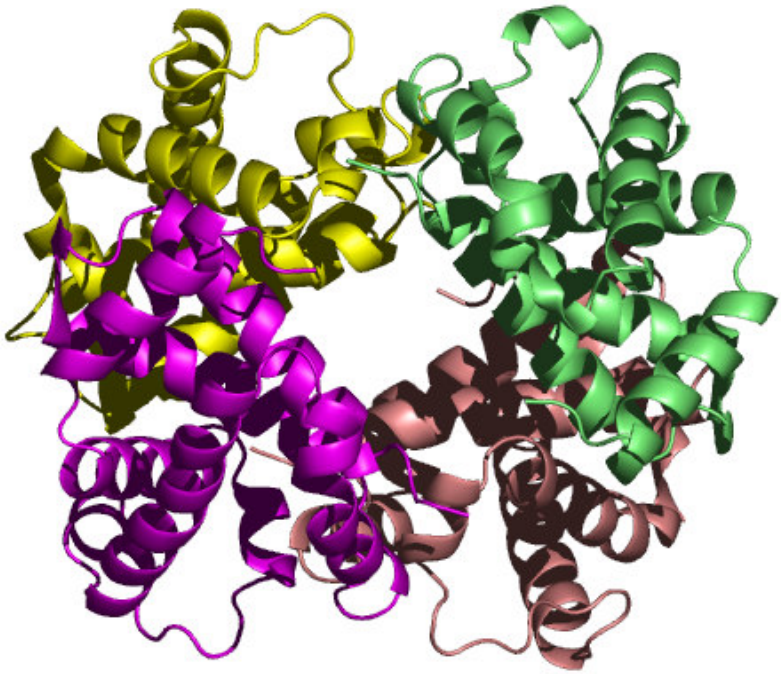


Figure 08 : Structure tertiaire de la porine (code PDB : 1E54) : association de brins β anti-parallèles en tonneau β (β-barrel).

B.4. Structure quaternaire

La structure quaternaire d'une protéine résulte de l'association de plusieurs chaînes polypeptidiques ou sous-unités. Ces sous-unités peuvent être identiques comme c'est le cas pour l'hémoglobine (**Figure 09**) ou bien radicalement différente, comme pour le récepteur de

l'élastine. L'association peut être nécessaire à l'activité biologique qui, dans ce cas, est le plus souvent extrêmement diminuée, voir inexistante au niveau tertiaire.



<u>Figure 09</u> : Structure quaternaire de l'hémoglobine humaine (code PDB : 1a3N), constituée de 4 sous-unités identiques représentées en différentes couleurs.

C. CONTEXTE BIOLOGIQUE :

L'ensemble de macromolécules, polysaccharides ou glycosaminoglycanes, protéines fibreuses, sels et eau qui entoure les cellules de l'organisme est appelé tissu conjonctif ou matrice extracellulaire (MEC).

La matrice extracellulaire a deux grands rôles principaux. Elle a tout d'abord un rôle de structure ou d'architecture. Elle permet en effet l'intégration des cellules dans les tissus et, suivant sa composition, elle peut donner un tissu extrêmement solide et rigide, comme l'os, ou bien autoriser une grande élasticité comme celle nécessaire aux tissus soumis à de fortes contraintes mécaniques comme la peau, les poumons ou les artères. Le deuxième rôle est celui de régulateur du métabolisme cellulaire : les molécules la composant, ou bien leurs produits de dégradation, peuvent influencer le comportement des cellules environnantes en interagissant avec leurs récepteurs membranaires.

La matrice extracellulaire est principalement composée de quatre types de molécules : les protéoglycanes, les glycoprotéines et les protéines fibreuses que sont les collagènes et l'élastine :

- **les protéoglycanes** : résultent de la fixation covalente de glycosaminoglycanes (GAG) sur une chaîne protéique. Les GAG sont des polymères non ramifiés de disaccharides dont l'un des sucres est aminé; ils peuvent être sulfatés (chondroïtine sulfate, dermatane sulfate, héparane sulfate) ou non (acide hyaluronique).

- **les glycoprotéines** : sont des protéines glycosylées qui assurent la cohérence de la matrice extracellulaire et l'adhésion des cellules à celle-ci. Ainsi, la fibronectine est capable de se fixer au collagène ou à l'héparane sulfate. De même, les intégrines membranaires interagissent avec la fibronectine par l'intermédiaire des séquences **RGD** (Arginine-Glycine-Aspartate).

- **les collagènes :** sont très abondants au sein de la MEC : jusqu'à 25 % des protéines de l'organisme. Son unité élémentaire est le procollagène qui est principalement composé de répétitions de type glycine-proline-hydroxyproline, imposant ainsi une structure en hélice de type polyproline II. Les procollagènes s'associent en trimères pour composer une super-hélice : la molécule de collagène. Les molécules de collagène sont sécrétées vers l'extérieur de la cellule, dans l'espace extracellulaire. Elles vont ensuite s'associer à leur tour par des liaisons croisées intermoléculaires, avec un décalage régulier donnant ainsi un aspect de striation régulière périodique aux fibrilles de collagène lorsque l'on observe ces dernières en

microscopie électronique. Les fibrilles de collagène vont enfin elles-mêmes s'assembler en fibres de collagène. Ce sont les fibres de collagène qui confèrent la rigidité de la MEC.

- l'**élastine** : constitue le dernier composant d'importance de la MEC.

C.1. L'élastine

Il est nécessaire à certains tissus d'un organisme, en réponse à un stress mécanique, de posséder des propriétés d'élasticité et/ou de flexibilité. Les protéines capables d'exercer ce type de rôle au sein d'un organisme sont les protéines dites élastomériques, c'est-à-dire dotées de propriétés élastiques. Ces protéines sont caractérisées par une alternance de domaines riches en acides aminés hydrophobes, auxquels sont attribuées les propriétés élastiques, et de domaines de pontages permettant des liaisons croisées intra- et/ou inter-chaînes (Tatham et al. 2000).

Plusieurs de ces protéines ont ainsi été caractérisées dans le monde animal : la résiline chez les arthropodes, l'abductine chez les mollusques et **l'élastine** chez les vertébrés.

L'élastine est, avec les microfibrilles, le composant majoritaire des fibres élastiques (Ross et Bornstein 1969). Plusieurs revues lui ont été dédiées de manière détaillée ainsi qu'à son porécurseur monomère, la tropoélastine (Vrhovski et al. 1998; Debelle et Alix 1999; Debelle et Tamburro 1999).

C.1.a. La tropoélastine

La tropoélastine conduit, par polymérisation, à la formation du réseau tridimensionnel élastine au sein des fibres élastiques de la MEC.

α. Structure primaire

Le gène de la tropoélastine est situé sur le chromosome 7 (Fazio et al. 1991) dont les ARN messagers issus de la transcription peuvent subir des épissages alternatifs (Bashir et al. 1990). En effet, lors de la synthèse protéique, les régions de l'ARN messager appelées introns sont éliminées, tandis que d'autres appelées exons sont conservées et permettent la construction de la protéine. Cependant de nombreux exons codant pour des unités structurales et fonctionnelles définies, peuvent eux mêmes, sous certaines contraintes environnementales, être éliminés : c'est ce que l'on appelle un épissage alternatif. Ainsi, la séquence de résidus correspondant à l'exon 33, n'est présente que dans 50 % des molécules de tropoélastine.

La tropoélastine humaine est constituée de 760 résidus (786 avec le peptide signal) pour un poids moléculaire d'environ 70 kDa (Indik et al. 1987; Bashir et al. 1990). Elle est présente sous la référence P15502 dans la banque de donnée Swissprot (Bairoch et Apweiler 1997). Une analyse statistique montre que la séquence contient plus de 80 % de résidus à caractère hydrophobe : glycines, alanines, valines, leucines et prolines (**Figure 10**). L'alanine est principalement retrouvée au niveau des domaines de pontages à l'intérieur desquels sont intercalées les lysines qui permettent de créer des liaisons inter et intra-chaînes appelées desmosines et isodesmosines (Bressan et Prockop 1977).

Une autre caractéristique remarquable de cette protéine est le fait que les exons du gène codent pour des domaines particuliers :

- les exons 1, 2, 3, 5, 7, 9, 11, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, et 33, principalement constitués de prolines, glycines, alanines, valines et leucines, sont les domaines hydrophobes supposés être responsables des propriétés élastiques.

- les exons 4, 6, 8, 12, 13, 15, 17, 19, 21, 23, 25, 27, 29, et 31, principalement constitués de prolines, alanines et de lysines, forment les domaines de pontages intra- et inter-chaînes.

- l'exon 26A est associé à plusieurs pathologies dont l'hypertension artérielle (Liptay et al. 1993; Bisaccia et al. 1998; Ostuni et al. 2002).

- l'exon 36 est caractérisé par les deux seules cystéines présentes dans la tropoélastine qui forment un pont disulfure intra-chaîne (Brown et al. 1992). Ce domaine est exposé, car il peut être reconnu par des anticorps spécifiques (Wei et al. 1998) et interagit directement avec d'autres composants des fibres élastiques lors de l'élastogenèse, notamment avec les MAGP-1 (Microfibril-Associated GlycoProtein-1) (Brown-Augsburger et al. 1994). On a longtemps cru que seul ce domaine avait cette propriété, mais une étude récente a démontré que d'autres régions de la tropoélastine pouvaient également interagir avec les MAGP-1 (Bashir et al. 1990; Clarke et Weiss 2004). Ce domaine permet également la fixation de la tropoélastine à une protéine d'importance de la MEC, l'intégrine $\alpha_v\beta_3$ (Rodgers et Weiss 2004).

Une étude récente par méthodes biospectroscopiques (CD, RMN) et bioinformatiques (homologie 1D, 3D, mécanique et dynamique moléculaires..) a été réalisée sur le domaine C-terminal de la tropoélastine humaine. Les structures des formes ouverte et fermée (par le pont SS) du peptide **GGACLGKACGRKRK** correspondant à l'exon36 ont été analysées en détail avec leurs possibilités d'interactions avec d'autres protéines de la MEC. Bien que les structures analysées soient retrouvées dans d'autres protéines (homologies 1D et/ou 3D), ceci mettant en évidence des séquences consensus, aucune conclusion définitive ne peut encore être énoncée en terme de relation structure / fonction. (Floquet et al. 2005).

```
  1 : MAGLTAAAPR PGVLLLLLSI LHPSRPGGVP GAIPGGVPGG VFYPGAGLGA
 51 : LGGGALGPGG KPLKPVPGGL AGAGLGAGLG AFPAVTFPGA LVPGGVADAA
101 : AAYKAAKAGA GLGGVPGVGG LGVSAGAVVP QPGAGVKPGK VPGVGLPGVY
151 : PGGVLPGARF PGVGVLPGVP TGAGVKPKAP GVGGAFAGIP GVGPFGGPQP
201 : GVPLGYPIKA PKLPGGYGLP YTTGKLPYGY GPGGVAGAAG KAGYPTGTGV
251 : GPQAAAAAAA KAAAKFGAGA AGVLPGVGGA GVPGVPGAIP GIGGIAGVGT
301 : PAAAAAAAAA AKAAKYGAAA GLVPGGPGFG PGVVGVPGAG VPGVGVPGAG
351 : IPVVPGAGIP GAAVPGVVSP EAAAKAAAKA AKYGARPGVG VGGIPTYGVG
401 : AGGFPGFGVG VGGIPGVAGV PSVGGVPGVG GVPGVGISPE AQAAAAAKAA
451 : KYGAAGAGVL GGLVPGPQAA VPGVPGTGGV PGVGTPAAAA AKAAAKAAQF
501 : GLVPGVGVAP GVGVAPGVGV APGVGLAPGV GVAPGVGVAP GVGVAPGIGP
551 : GGVAAAAKSA AKVAAKAQLR AAAGLGAGIP GLGVGVGVPG LGVGAGVPGL
601 : GVGAGVPGFG AGADEGVRRS LSPELREGDP SSSQHLPSTP SSPRVPGALA
651 : AAKAAKYGAA VPGVLGGLGA LGGVGIPGGV VGAGPAAAAA AAKAAAKAAQ
701 : FGLVGAAGLG GLGVGGLGVP GVGGLGGIPP AAAAKAAKYG AAGLGGVLGG
751 : AGQFPLGGVA ARPGFGLSPI FPGGACLGKA CGRKRK

peptide signal : MAGLTAAAPRPGVLLLLLSILHPSRP

exon  1  :  G
exon  2  :  GVPGAIPGGVPGGVFYP
exon  3  :  GAGLGALGG
exon  4  :  GALGPGGKPLKPV
exon  5  :  PGGLAGAGLGA
exon  6  :  GLGAFPAVTFPGALVPGGVADAAAAYKAAKA
exon  7  :  GAGLGGVPGVGGLGVSA
exon  8  :  GAVVPQPGAGVKPGKVP
exon  9  :  GVGLPGVYPGGVLPGA
exon 10  :  RFPGVGVLPGVPTGAGVKPKAP
exon 11  :  GVGGAFAGIP
exon 12  :  GVGPFGGPQPGVPLGYPIKAPKLP
exon 13  :  GGYGLPYTTGKLPY
exon 14  :  GYGPGGVAGAAGKAGYPTGT
exon 15  :  GVGPQAAAAAAA KAAAKF
exon 16  :  GAGAAGVLPGVGGAGVPGVPGAIPGIGGIA
exon 17  :  GVGTPAAAAAAAAAAKAAKY
exon 18  :  GAAAGLVPGGPGFGPGVVGVPGAGVPGVGVPGAGIPVVPGAGIPGAAVP
exon 19  :  GVVSPEAAAKAAAKAAKY
exon 20  :  GARPGVGVGGIPTYGVGAGGFPGFGVGVGGIPGVAGVPSVGGVPGVGGVPGVGIS
exon 21  :  PEAQAAAAAKAAKY
exon 22  :  GAAGAGVLGGLVPGPQAAVPGVPGTGGVP
exon 23  :  GVGTPAAAAAKAAAKAAQF
exon 24  :  GLVPGVGVAPGVGVAPGVGVAPGVGLAPGVGVAPGVGVAPGVGVAPGI
exon 25  :  GPGGVAAAAKSAAKVAAKAQL
exon 26  :  RAAAGLGAGIPGLGVGVGVPGLGVGAGVPGLGVGAGVPGFGA
exon 26A :  GADEGVRRSLSPELREGDPSSSQHLPSTPSSPRV
exon 27  :  PGALAAAKAAKY
exon 28  :  GAAVPGVLGGLGALGGVGIPGGVV
exon 29  :  GAGPAAAAAAAKAAAKAAQF
exon 30  :  GLVGAAGLGGLGVGGLGVPGVGGLG
exon 31  :  GIPPAAAAKAAKY
exon 32  :  GAAGLGGVLGGAGQFPLG
exon 33  :  GVAARPGFGLSPIFP
exon 36  :  GGACLGKACGRKRK
```

<u>Figure 10</u> : Séquences de la tropoélastine humaine et de ses exons associés.
gris : peptide signal; bleu : domaines hydrophobes; rouge : domaines de pontage;
violet : exon 26A; vert : exon 36.

β. Structure secondaire et tertiaire

A l'heure actuelle, la structure tridimensionnelle de la tropoélastine reste inconnue. A cause de ses propriétés physicochimiques inhabituelles, les méthodes expérimentales se sont révélées incapables de permettre la détermination de la structure tridimensionnelle de la tropoélastine. En effet, sa grande flexibilité la rend non cristallisable et interdit donc le recours à la diffraction des rayons X. De même, son insolubilité dans les solvants habituels rend extrêmement difficile l'utilisation des techniques de spectroscopies classiques. Outre ces problèmes, le fait qu'un faible nombre de résidus (alanine, glycine, proline, valine) représentent plus de 80 % de sa composition, ainsi que la grande longueur de sa séquence, rendent impossible sa détermination structurale par spectroscopie de RMN.

Des études de prédiction de structures secondaires basées sur les méthodes type Chou-Fasman (Argos et al. 1978) et GOR (Garnier et al. 1978; Busetta 1987), réalisées sur les séquences humaine et bovine de la tropoélastine, ont permis de proposer une première distribution locale des structures secondaires où les exons de pontage, qui contiennent majoritairement de l'alanine, sont proposés en hélice α et les autres exons présentent alternativement des brins et des coudes. La structure secondaire globale est constituée de 18 ± 5 % d'hélices α, 63 ± 17 % de brins β, 13 ± 13 % de coudes β et 6 ± 6 % de structures non régulières (Debelle et al. 1992). Dans cette même étude, grâce à l'utilisation de la méthode de prédiction d'antigénicité PEOPLE (Alix 1999), plusieurs épitopes ont été proposés dont les séquences correspondant à l'exon 26A et à l'exon 30 que l'on sait être effectivement des cibles d'anticorps (Wei et al. 1998).

Des expériences de spectroscopie de Dichroïsme Circulaire (DC) et de spectroscopie Raman ont été effectuées sur la tropoélastine bovine et ont permis d'évaluer la présence d'environ 10 % d'hélices α, 45 % de brins β et environ 45 %, de structures non régulières (Debelle et Alix 1995; Debelle et al. 1995). Plus récemment, une méthode d'affinage des prédictions de structures secondaires basée sur des données obtenues par spectroscopie de DC, appelée LINK (Alix 1997; Gouliaev 1998), a permis d'estimer approximativement les pourcentages de structures secondaires de la tropoélastine humaine à 10 % d'hélices α, 35 % de brins β et 55 % de structures non régulières (Debelle et al. 1998). D'autres études du même type ont été effectuées (Vrhovski et al. 1997; Muiznieks et al. 2003) et ont démontré des pourcentages de structures secondaires proches de ceux précédemment obtenus. Ainsi, bien que les méthodes de prédictions de structures secondaires soient basées sur des banques de données de structures de protéines globulaires, contrairement à l'élastine qui est fibrillaire, et en dépit de

la faible hétérogénéité des résidus la constituant (principalement les "briseurs de structures secondaires" que sont la glycine et la proline), ces techniques ont tout de même prouvé leur intérêt surtout en association avec des résultats expérimentaux. Ces travaux ont ainsi permis de proposer un premier modèle "pseudo-3D" de la structure de la tropoélastine humaine (Debelle et Alix 1999; Alix 2001; Alix et al. 2003).

Ne pouvant ainsi être étudiée dans sa totalité et à cause de l'organisation particulière de ses domaines issus des différents exons, la structure de la tropoélastine a été étudiée par le biais de séquences caractéristiques sous forme de peptides.

L'une des séquences caractéristiques de la tropoélastine humaine est **VPGVG** pour laquelle Urry a proposé que, sous forme peptidique, ce motif formerait un coude β de type II sur **VPGV**, stabilisé par une liaison hydrogène entre le **C=O** de la première valine et le **N-H** de la deuxième valine (Urry et al. 1974a; Urry et al. 1974b; Urry et al. 1975a; Khaled et Urry 1976). Des études portant sur les répétitions d'autres motifs ont également été effectuées : $(\mathbf{APGVGV})_n$ (Urry et al. 1975b), $(\mathbf{VAPG})_n$ et $(\mathbf{VPGG})_n$ (Rapaka et Urry 1978; Long et al. 1980), $(\mathbf{APGG})_n$ et $(\mathbf{APGVG})_n$ (Rapaka et al. 1978)… Toutes ces études ont eu pour conséquence de proposer un modèle de la structure de la tropoélastine, basé sur la répétition de coudes β de type II donnant naissance à une spirale, appelée spirale β (Urry et al. 1975b; Chang et al. 1989; Urry et Parker 2002). L'étude de ce modèle de spirale β a été effectuée par plusieurs méthodes de modélisation moléculaire (Li et al. 2001a; Li et al. 2001b; Li et al. 2002; Li et Dagget 2002; Li et Dagatt 2003). A la lecture de ces articles, la spirale β semble ne pas être stable au cours du temps, ce qui remet donc totalement en cause la validité de ce modèle.

De la même manière, d'autres peptides de séquence **GXGGX** (**X** étant un résidu quelconque) caractéristiques de la tropoélastine ou des peptides de séquences dérivées de l'élastine (Castiglione-Morelli et al. 1990; Guantieri et al. 1990; Tamburro et al. 1990; Tamburro et al. 1991; Lelj et al. 1992; Tamburro et al. 1992; Villani et Tamburro 1995; Martino et al. 1998; Villani et Tamburro 1999; Martino et al. 2000; Martino et al. 2001; Martino et al. 2002), mais également de la lamprine, une autre protéine élastomérique provenant de la Lamproie (Bochicchio et al. 2001) ont été étudiés. Un modèle de structure de la tropoélastine a ainsi pu être proposé, d'après lequel, les séquences répétées **GXGGX** ont une conformation en coude β de type II sur **GXGG**. Le motif **GXGG** étant très labile, celui-ci peut "glisser" dynamiquement en coude β sur **XGGX** (Tamburro et Guantieri 1991; Lelj et al. 1992; Martino et al. 2000). Parallèlement à cela, à partir de ce modèle de structure, un modèle

d'élasticité a été proposé selon lequel l'état relaxé de l'élastine est caractérisé par des structures en coudes très labiles pouvant se convertir facilement d'un type à un autre, tout en glissant le long de la chaîne peptidique. La structure de l'état relaxé possèderait donc une grande entropie, fortement favorisée par une hydratation importante qui augmenterait les mouvements de la chaîne peptidique. A l'inverse, soumise à des forces d'étirement, l'élastine adopterait une conformation plus structurée de type étendue ou quasi étendue (hélice de polyproline II, brin β), peu hydratée et donc possédant une faible entropie. Lorsque les forces d'étirement cessent, l'élastine devient moins structurée, son hydratation augmente, tout comme son entropie, ce qui conduit au retour de l'état relaxé. L'un des atouts de ce modèle d'élasticité, est qu'il permet d'expliquer l'un des paradoxes de l'élastine : bien qu'elle soit complètement insoluble dans l'eau, la présence d'eau est indispensable au phénomène d'élasticité de l'élastine (Debelle et Tamburro 1999).

A cause de l'organisation particulière "en cassette" du gène de la tropoélastine, où chaque exon du gène code pour un domaine particulier, il est possible d'étudier séparément sous forme de peptide chacun des exons de la tropoélastine. La synthèse chimique de tous les exons de la tropoélastine humaine et leur analyse structurale systématique a été entrepris récemment (Bisaccia et al. 2000; Tamburro et al. 2003; Bochicchio et al. 2004a; Bochicchio et al. 2004b). Ces études ont permis de montrer, sur des molécules de taille supérieure à celle des peptides provenant de séquences caractéristiques, la présence de coudes β non seulement de type II, mais aussi de type I et VIII, ainsi que la présence d'hélices de type polyproline II. Ces différentes conformations en équilibre pourraient expliquer les propriétés élastiques des domaines hydrophobes en confirmant le modèle d'élasticité décrit précédemment, mais ne peuvent permettre en aucun cas de prédire la structure globale de la tropoélastine.

C.1.b. Les peptides d'élastine

Malgré son indéniable importance structurale et architecturale au sein de la MEC, l'élastine n'est pas seulement cantonnée à cette fonction. Il a été en effet démontré que des peptides dérivant de l'élastine étaient présents dans la circulation sanguine, aussi bien dans des conditions pathologiques que normales (Wei et al. 1993). L'origine de ces peptides n'a pas été clairement démontrée *in vivo*, mais il semble vraisemblable qu'ils proviennent essentiellement de la dégradation des fibres d'élastine par des élastases (de type métalloprotéinase matricielle ou sérine protéinase), bien que des dégradations post-traductionnelles de la tropoélastine puissent également être responsables de la présence de ces peptides (Robert et al. 1995). Quelles que soient leurs origines, l'identification de la séquence, de la structure et des activités

biologiques de ces peptides dérivant de l'élastine semble être primordiale pour la compréhension de nombreux mécanismes cellulaires.

Les peptides qui ont servi à ces études peuvent provenir d'une hydrolyse chimique à base d'acide oxalique, on est alors en présence d'α-élastine (Partridge et al. 1955), ou bien à base d'hydroxyde de potassium ce qui donne de la κ-élastine (Jacob et Hornebeck 1985). Cette hydrolyse chimique conduit ainsi à l'obtention d'un mélange indéterminé de différents peptides. A l'inverse, il est possible, par synthèse peptidique chimique, de n'avoir qu'un seul type déterminé de peptide de manière à connaître spécifiquement son rôle biologique ou sa structure.

Le premier et le plus intensivement étudié des peptides d'élastine est **VGVAPG**, dont la séquence provient de l'exon 24 de la tropoélastine humaine, dans lequel il est répété six fois (voir **Figure 01**). Il a été démontré que ce peptide **VGVAPG** possédait des propriétés chimioattractantes vis-à-vis des monocytes, des fibroblastes et des cellules tumorales (Senior et al. 1984; Bisaccia et al. 1994), une activité vasorelaxante sur les cellules de l'épithélium vasculaire (Faury et al. 1998; Lograno et al. 1998) et qu'il induit chez les leucocytes humains une modification des flux d'ions, la libération d'élastases et la production d'anions superoxydes (Hauck et al. 1995) (Fulop et al. 1998). Une autre activité biologique d'importance de **VGVAPG** est d'induire la production de MMPs (pour Matrix MetalloProteases) par les cellules humaines de fibrosarcome (Brassart et al. 1998). Toutes ces activités ont pour point commun d'être médiées par la fixation du peptide sur une protéine de 67 kDa appelée EBP (pour Elastin Binding-Protein) qui est une des trois sous-unités du récepteur de l'élastine. De même, le peptide **PGAIPG**, dont la séquence est répétée deux fois dans la tropoélastine humaine, possède une activité chimioattractante vis-à-vis des fibroblastes, des neutrophiles et des cellules de carcinome humain (Grosso et Scott 1993b) et est également un ligand de l'EBP (Grosso et Scott 1993c). De nombreux autres peptides dérivant de l'élastine ont eu leurs activités biologiques étudiées de la même manière, démontrant ainsi qu'ils sont susceptibles de modifier grandement le comportement des cellules au contact de la MEC. Pour une description détaillée, il est conseillé de se reporter à la revue de Bissacia et collaborateurs (Bisaccia et al. 1999).

Il a été aussi démontré que les peptides contenant la séquence **GXXP** étaient reconnus par le même anticorps, prouvant ainsi qu'ils avaient des structures tridimensionnelles proches (Grosso et Scott 1993a).

L'étude de la relation structure-activité, notamment des peptides **YGVG**, **GLVPG**, **VGVAPG**, **GVAPG** et **GGVPG** (Bisaccia et al. 1994; Castiglione Morelli et al. 1997), n'a pas permis de mettre en évidence par spectroscopies de RMN ou de DCE une conformation particulière à cause de la grande flexibilité de ces peptides.

Plus récemment, il a été montré au laboratoire, par spectroscopie de DCE, que les peptides capables d'induire une production de MMP-1 et MMP-3 chez les fibroblastes humains (**VGVAPG**, **GVAPGV**, **PGAIPG** et **LGTIPG** qui est un peptide dérivant de la laminine, une autre protéine de la MEC) présentaient une structure différente de celle des peptides n'ayant pas cette propriété biologique (Brassart et al. 2001; Fuchs et al. 2001). Cette étude a également mis en évidence qu'une glycine située après le motif **GXXP** était indispensable pour provoquer la production de MMPs.

De même, l'utilisation d'une nouvelle méthode de prédiction des types de coudes β, basée sur une banque de 205 protéines de structure connue, a permis de proposer que la structure du motif **GXXP** est un coude β de type VIII chez les peptides actifs (Alix 2001; Fuchs et al. 2001; Alix et al. 2003; Fuchs et Alix 2005).

Une étude utilisant différentes méthodes de modélisation moléculaire a été réalisée au laboratoire sur le peptide **VGVAPG**, et a permis de montrer qu'un coude β de type VIII était effectivement observable sur le motif GVAP (Floquet et al. 2004).

Il est à noter que les peptides d'élastine ne sont pas les seuls à induire une activité biologique.

De nombreux autres peptides ont également cette capacité de modifier le comportement cellulaire tout en provenant de la protéolyse de constituants de la MEC, comme par exemple le domaine N-terminal de la fibronectine qui stimule l'expression de MMP-1 par les fibroblastes (Xie et al. 1994) ou encore le tripeptide **RFK**, dérivant de la thrombospondine-1, qui participe à l'activation du TGF β (Ribeiro et al. 1999). Ces peptides ont été baptisés "matrikines" (Maquart et al. 1999; Duca et al. 2004; Maquart et al. 2004) ou encore "matricryptines" (Davis et al. 2000).

Par analogie, les peptides d'élastine impliqués dans les interactions avec une structuration particulière (par exemple coude β de type VIII) ont été baptisés "morphoélastokines" (Alix et al. 2003, Duca et al. 2004).

C.2. Le récepteur de l'élastine :

C.2.a. Structure

Le récepteur de l'élastine est constitué de 3 sous-unités (Wrenn et al. 1988). Les deux premières, de respectivement 61 et 55 kDa, sont ancrées dans la membrane plasmique de la cellule (Mecham et al. 1989). La troisième sous-unité, qui a un poids moléculaire de 67 kDa, se situe à la périphérie de la membrane plasmique. Elle est appelée EBP, pour Elastin Binding-Protein, car c'est elle qui est capable d'interagir avec la tropoélastine et ses peptides avec une grande affinité puisque la constante de dissociation est de 8 nM (Wrenn et al. 1988). L'EBP possède également un autre site ayant des propriétés galactolectines, c'est-à-dire qui permet de fixer les sucres de type β-galactose ou possédant un β-galactose, comme le lactose, ou certains glycosaminoglycanes de la MEC, comme le dermatane sulfate ou chondroitine sulfate (Hinek et al. 1988). L'interaction entre un β-galactose et le site galactolectine entraîne un changement de conformation de l'EBP qui provoque une diminution de son affinité vis-à-vis des deux autres sous-unités, ce qui provoque une inhibition des propriétés biologiques du récepteur de l'élastine (Mecham et al. 1991). L'EBP est en fait un variant d'épissage de la β-galactosidase qui ne possède pas d'activité β-galactosidique (Hinek et al. 1993). Sa séquence protéique (Morreau et al. 1989; Privitera et al. 1998) et celle de la β-galactosidase sont représentées dans la **Figure 11**.

Une partie de la séquence spécifique de l'EBP humaine possède des homologies de séquence (**Figure 12**) avec certaines élastases de la famille des sérine-protéinases comme l'élastase pancréatique porcine, l'élastase leucocytaire humaine et le facteur IX humain, ainsi que l'EBP de mouton ce qui laisse à penser qu'il s'agit peut-être d'une séquence consensus permettant de lier l'élastine (Hinek et al. 1993; Hinek et Rabinovich 1994). En plus de leurs homologies de séquence, les structures de ces protéines sont extrêmement proches car le même anticorps est capable de se fixer sur le site de l'EBP humaine ou sur le site de l'élastase pancréatique porcine. De manière encore plus intéressante, cette même séquence sous forme de peptide, appelé peptide Sgal, est capable de se lier directement à l'élastine, par l'intermédiaire des répétitions **VGVAPG** de l'exon 24, ce qui confirme que nous sommes bien en présence du site de fixation de l'EBP à l'élastine (Hinek et Rabinovich 1994).

EBP humaine : 546 résidus

```
  1 : MPGFLVRILL LLLVLLLLGP TRGLRNATQR MFEIDYSRDS FLKDGQPFRY ISGSIHYSRV
 61 : PRFYWKDRLL KMKMAGLNAI QTLPGSCGQV VGSPSAQDEA SPLSEWRASY NSAGSNITDA
121 : FLSQRKCEPK GPLINSEFYT GWLDHWGQPH STIKTEAVAS SLYDILARGA SVNLYMFIGG
181 : TNFAYWNGAN SPYAAQPTSY DYDAPLSEAG DLTEKYFALR NIIQKFEKVP EGPIPPSTPK
241 : FAYGKVTLEK LKTVGAALDI LCPSGPIKSL YPLTFIQVKQ HYGFVLYRTT LPQDCSNPAP
301 : LSSPLNGVHD RAYVAVDGIP QGVLERNNVI TLNITGKAGA TLDLLVENMG RVNYGAYIND
361 : FKGLVSNLTL SSNILTDWTI FPLDTEDAVR SHLGGWGHRD SGHHDEAWAH NSSNYTLPAF
421 : YMGNFSIPSG IPDLPQDTFI QFPGWTKGQV WINGFNLGRY WPARGPQLTL FVPQHILMTS
481 : APNTITVLEL EWAPCSSDDP ELCAVTFVDR PVIGSSVTYD HPSKPVEKRL MPPPPQKNKD
541 : SWLDHV
```

β galactosidase humaine : 677 résidus

```
  1 : MPGFLVRILL LLLVLLLLGP TRGLRNATQR MFEIDYSRDS FLKDGQPFRY ISGSIHYSRV
 61 : PRFYWKDRLL KMKMAGLNAI QTYVPWNFHE PWPGQYQFSE DHDVEYFLRL AHELGLLVIL
121 : RPGPYICAEW EMGGLPAWLL EKESILLRSS DPDYLAAVDK WLGVLLPKMK PLLYQNGGPV
181 : ITVQVENEYG SYFACDFDYL RFLQKRFRHH LGDDVVLFTT DGAHKTFLKC GALQGLYTTV
241 : DFGTGSNITD AFLSQRKCEP KGPLINSEFY TGWLDHWGQP HSTIKTEAVA SSLYDILARG
301 : ASVNLYMFIG GTNFAYWNGA NSPYAAQPTS YDYDAPLSEA GDLTEKYFAL RNIIQKFEKV
361 : PEGPIPPSTP KFAYGKVTLE KLKTVGAALD ILCPSGPIKS LYPLTFIQVK QHYGFVLYRT
421 : TLPQDCSNPA PLSSPLNGVH DRAYVAVDGI PQGVLERNNV ITLNITGKAG ATLDLLVENM
481 : GRVNYGAYIN DFKGLVSNLT LSSNILTDWT IFPLDTEDAV RSHLGGWGHR DSGHHDEAWA
541 : HNSSNYTLPA FYMGNFSIPS GIPDLPQDTF IQFPGWTKGQ VWINGFNLGR YWPARGPQLT
601 : LFVPQHILMT SAPNTITVLE LEWAPCSSDD PELCAVTFVD RPVIGSSVTY DHPSKPVEKR
661 : LMPPPPQKNK DSWLDHV
```

Figure 11 : Séquences protéiques de l'EBP et de la β-galactosidase humaine : en gras les séquences spécifiques des protéines.

EBP humaine (peptide Sgal)	**VVGSP SAQDE ASPLS**
EBP ovine	**VVGGT EAQRN SWPLQ**
Elastase pancréatique porcine	**VVGGT EAQRN SWPSQ**
Elastase leucocytaire humaine	**IVGGR RARPH AWPFM**
Facteur IX humain	**VVGGE DAKPG QFPWQ**

Figure 12 : Homologies de séquences du peptide Sgal (Hinek et Rabinovich 1994). en rouge : les résidus conservés dans chacune des séquences.

La structure exacte du récepteur de l'élastine reste encore pour l'instant inconnue. Mais puisque l'EBP est le variant inactif d'épissage de la β-galactose, qui est elle-même associée dans la membrane des lysosomes à une neuraminidase (Nmase) de 61 kDa et à une protéine protectrice (PP) de 55 kDa, il a été proposé que l'architecture moléculaire des deux complexes était similaire (**Figure 13**) (Hinek 1996). Il a d'ailleurs été démontré que l'association entre la β-galactosidase, la Nmase et la PP est nécessaire au niveau de la membrane lysosomiale pour la stabilité et l'activité de la β-galactosidase (D'Agrosa et Callahan 1988; Galjart et al. 1988; Potier et al. 1990). Aucune structure tridimensionnelle n'est encore connue pour la β-galactosidase humaine ou pour la Nmase humaine.

La structure tridimensionnelle de la PP est représentée ci-dessous **Figure 14**.

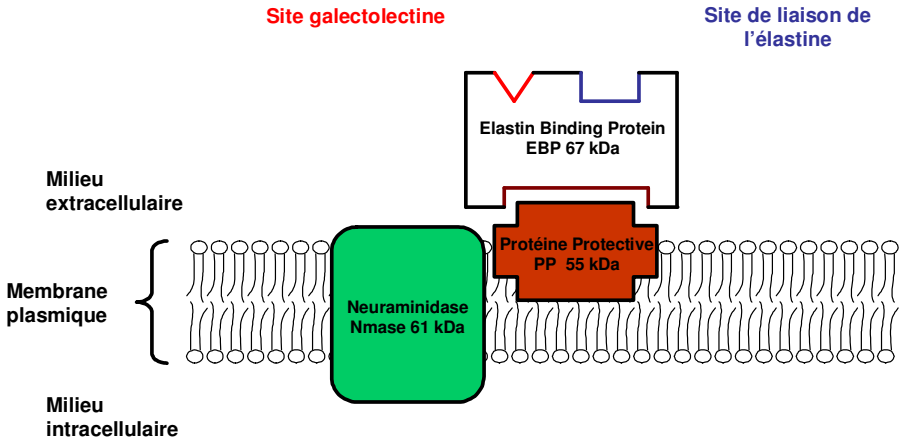


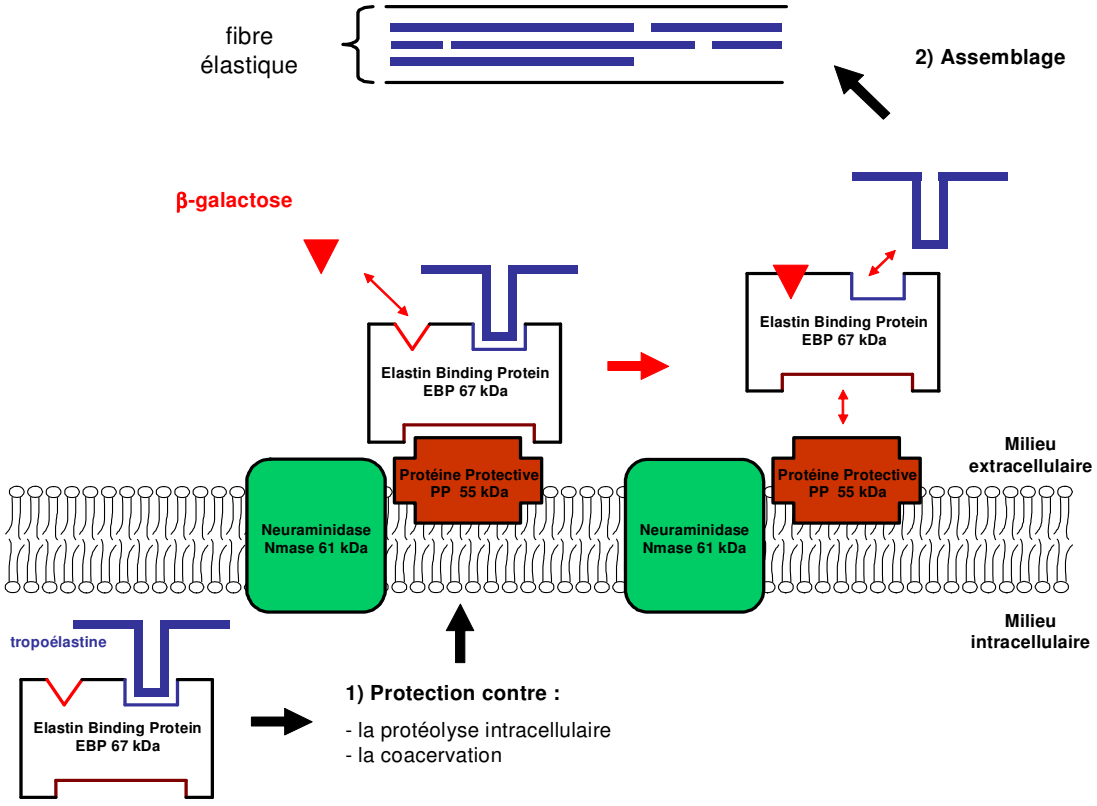
Figure 13 : Représentation schématique du récepteur de l'élastine (Hinek 1996).

Figure 14 : Représentation de la protéine protectrice humaine (code PDB : 1IVY).

La forme mature de la protéine protectrice humaine est formée de deux peptides de 32 et 20 kDa reliés par des ponts disulfures (D'Agrosa et al. 1992; Morreau et al. 1992) et sa structure a été résolue par cristallographie de rayons X (code PDB : 1IVY) (Rudenko et al. 1995).

C.2.b. Fonction

Le récepteur de l'élastine et plus précisément sa sous-unité EBP, possède deux rôles biologiques d'importance : un rôle de protéine chaperonne de la tropoélastine durant l'élastogenèse et un rôle de récepteur proprement dit. Son premier rôle est primordial dans la régulation de l'assemblage de la tropoélastine en fibre élastique (**Figure 15**).

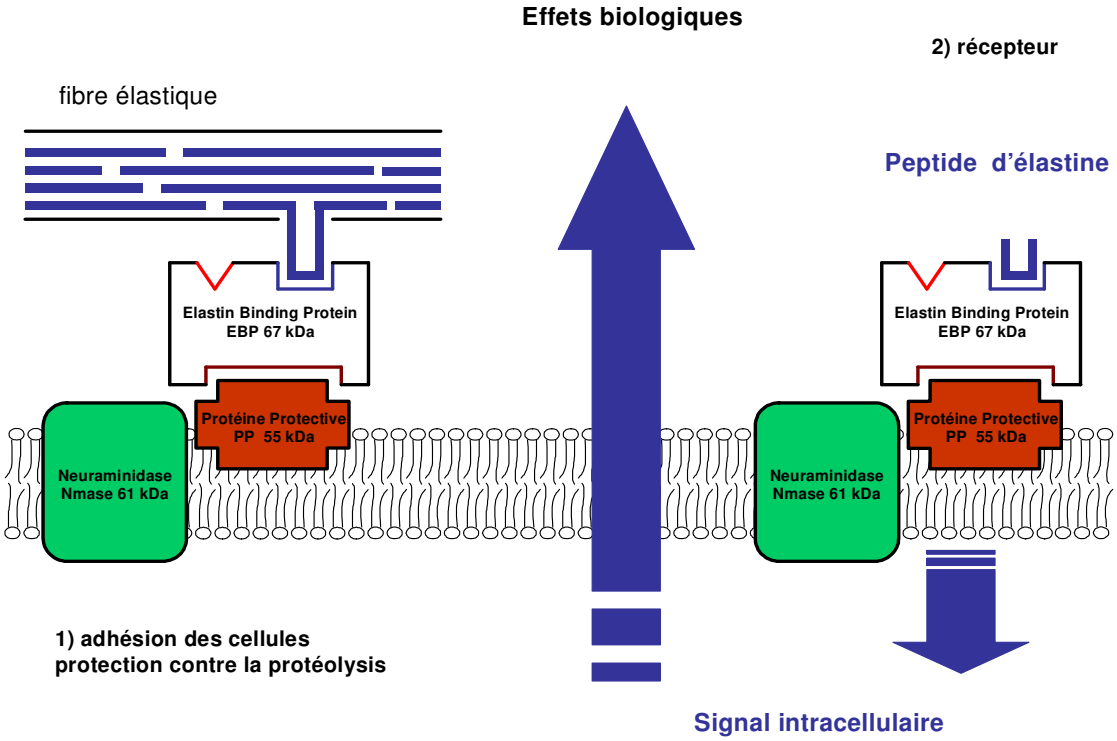


<u>Figure 15</u> : Rôle de l'EBP durant l'élastogenèse :
1) l'EBP protège la tropoélastine contre la protéolyse intracellulaire et la coacervation,
2) après la fixation d'un β-galactose sur l'EBP, la tropoélastine est libérée pour pouvoir être assemblée dans la fibre élastique.

Durant la maturation intracellulaire et le transport de la tropoélastine vers le milieu extracellulaire, l'EBP protège la tropoélastine contre les élastases intracellulaires et son agrégation spontanée appelée "coacervation" (Hinek et al. 1993; Hinek et Rabinovich 1994). En effet, l'EBP se lie à la tropoélastine par l'intermédiaire des répétitions **VGVAPG** de l'exon 24 de la tropoélastine qui sont aussi les sites de fixation des élastases sérine-protéinases, les

empêchant ainsi d'interagir avec la tropoélastine et de la dégrader. De même la coacervation résulte de l'interaction entre les domaines hydrophobes contenant des répétitions **GVGVP**, **GGVP** et **GVGVAP** (Vrhovski et Weiss 1998). L'EBP en se liant à la tropoélastine interdit ces interactions en masquant les répétitions **VGVAPG**. Il a été également montré que la séquence **PGAIPG**, répétée deux fois dans la tropoélastine, peut également être un site de fixation de la tropoélastine sur l'EBP (Grosso et Scott 1993c).

Durant l'élastogenèse, l'EBP et la tropoélastine sont synthétisées dans le réticulum endoplasmique rugueux où elles y forment un complexe qui est conduit dans des vésicules de transport à travers la cellule jusqu'à l'appareil de Golgi. Le complexe est sécrété par endocytose à l'extérieur de la cellule où il interagit avec la Nmase et la PP.



<u>Figure 16</u> : Rôle extracellulaire de l'EBP :
1) permet l'adhésion aux fibres élastiques tout en la protégeant de la protéolyse,
2) rôle de récepteur des peptides d'élastine et induction d'effets biologiques.

Au voisinage des fibres élastiques en formation, sont présentes des glycoprotéines et des protéoglycanes dont les chaînes glycosidiques contenant du β-galactose vont se lier à l'EBP par l'intermédiaire du domaine galactolectine. L'EBP voit son affinité pour les deux autres sous-unités diminuer, provoquant leur séparation. Dans le même temps, il y a perte d'affinité de l'EBP pour la tropoélastine qui va pouvoir être incorporée dans les fibres élastiques (Hinek 1996). L'EBP va ensuite être endocytée et recyclée directement dans les vésicules de transport

de la cellule pour pouvoir jouer à nouveau son rôle de protéine chaperonne de la tropoélastine et continuer l'assemblage des fibres élastiques (Hinek et al. 1995).

C'est à la surface des membranes plasmiques des cellules, que l'EBP permet au récepteur de l'élastine d'acquérir ses propriétés de fixation vis-à-vis de l'élastine et de ses dérivés (voir ci-dessus **Figure 16**). On retrouve le récepteur de l'élastine à la surface de nombreux types cellulaires comme les fibroblastes, les cellules lisses musculaires, les chondroblastes, les myocytes vasculaires, les monocytes, les neutrophiles, les leucocytes et certaines cellules tumorales issues de mélanomes, de cancers des poumons ou des seins (Hinek 1996).

De la même manière que durant l'élastogenèse, en association avec les deux autres sous-unités du récepteur ancrées dans la membrane plasmique, l'EBP peut fixer la tropoélastine des fibres élastiques. Non seulement, il y a une protection contre la protéolyse extracellulaire de la tropoélastine, mais cela permet aussi l'adhésion de la cellule aux fibres élastiques pour permettre sa migration au sein de la MEC (Hinek 1996). L'EBP permet aussi l'adhésion des cellules au domaine V de la chaîne B1 de la laminine, *via* la séquence **LGTIPG** (Liotta et al. 1986), ainsi que sur le collagène de type IV (Senior et al. 1989).

A travers la fixation à l'EBP, de nombreux peptides d'élastines, parmi lesquels **VGVAPG**, **PGAIPG**, **GVAPG**, **GLVPG** et le peptide dérivant de la laminine **LGTIPG**, sont responsables des modifications du comportement cellulaire. Ils induisent des propriétés de chimioattraction, des modifications intracellulaires d'ions, ou encore la libération de MMPs … Ce rôle du récepteur de l'élastine a déjà été décrit précédemment. Mais, la manière dont sont déclenchés les effets biologiques après la fixation des ligands sur le récepteur de l'élastine reste inconnue. Il a été en revanche démontré que la production de pro-MMP1 par les fibroblastes humains impliquait les voies de signalisation intracellulaire MEK/ERK (Duca et al. 2002).

Il est cependant important de rappeler que la conformation et non pas seulement la séquence des peptides d'élastine semble être importante puisqu'une différence de structure est observable par spectroscopie de Dichroïsme Circulaire entre des peptides actifs et inactifs (Brassart et al. 2001; Fuchs et al. 2001). Il est donc vraisemblable que l'interaction entre les peptides d'élastine et l'EBP se fasse sur le modèle très courant "serrure-clé", c'est-à-dire que chacune des molécules en présence doit avoir une conformation particulière pour que l'interaction ait lieu. La structure active proposée pour les peptides dérivant de l'élastine est le coude β de type VIII (Brassart et al. 2001; Alix 2001; Fuchs et al. 2001; Floquet et al. 2004), mais la structure de l'EBP tout comme celle de son domaine de réception appelé Sgal reste totalement inconnue.

Partie II
MATERIELS et METHODES

A. METHODES D'ETUDE EXPERIMENTALE

A.1. Synthèse peptidique

La plupart des peptides sont difficiles, voire impossibles, à extraire à partir d'échantillons biologiques ou à produire par génie génétique. La synthèse totale chimique permet l'obtention d'à peu près n'importe quel peptide si le nombre de résidus n'est pas trop important.

Les synthèses peptidiques des peptides Sgal, **YGVG**, **VGVAPG** et (**VGVAPG**)$_3$ ont été réalisées à l'aide d'un synthétiseur Applied Biosystem 431A utilisant la méthode SPPS (Solid-Phase Peptide Synthesis) dans le Laboratoire du Professeur A.M. TAMBURRO (Département de Chimie, Université de Potenza, Italie) dans le cadre d'un séjour doctoral ERASMUS de 3 mois.

La méthode SPPS, qui est actuellement l'une des plus utilisées et des plus performantes, est basée sur l'utilisation d'une matrice solide insoluble, appelée résine, à laquelle la chaîne peptidique est attachée de manière covalente par l'intermédiaire de la fonction carboxylique de l'acide aminé C-terminal.

Elle implique 3 phases successives :

(i) l'assemblage

Les acides aminés, qui vont être assemblés, sont protégés par un groupement protecteur appelé Fmoc (9-fluorénylméthyloxydecarbonyl) au niveau de la fonction amine terminale (NH_2). Ce groupement permet d'éviter que les acides aminés ne se lient spontanément dans un ordre différent de celui du peptide désiré. Le premier résidu qui sera finalement le résidu C-terminal voit son groupement acide carboxylique **(-COOH**) activé par du DCC (dicyclohexylcarbodiimide) et de la DMAP (4-diméthylaminopyridine). Le premier résidu va ainsi pouvoir se lier à une résine (la matrice solide). On va ensuite déprotéger le résidu en provoquant le détachement du groupement Fmoc par de la pipéridine. Le second résidu peut être ajouté et va être à son tour activé en ajoutant du DCC et désormais de l'HOBt (1-hydroxybenzotriazole) (Konig et Geiger 1970). Le résidu va alors pouvoir se coupler au résidu précédent en créant une liaison peptidique avec la fonction amine. Après le couplage,

le résidu va ensuite être à son tour déprotégé pour permettre à la synthèse peptidique de se poursuivre. En fait, l'assemblage est une série de déprotections et de couplages des différents résidus constituant le peptide à synthétiser.

(ii) la séparation de la résine et des groupes protecteurs des chaînes latérales

Du TFA (acide trifluoroacétique) est utilisé pour permettre le détachement du peptide de la résine. Comme certains acides aminés sont susceptibles d'altérer le bon déroulement de la synthèse peptidique, ils ont leur chaîne latérale elle aussi protégée par un groupement protecteur. C'est donc après le détachement du peptide de la résine que l'on procède à l'élimination des groupements protecteurs.

(iii) la purification

Une fois le peptide obtenu, il est indispensable de le purifier, c'est-à-dire de le débarrasser de tous les produits secondaires d'une synthèse imparfaite dûe à un mauvais assemblage des résidus. Pour réaliser la purification, nous avons eu recours à la CLHP (Chromatographie Liquide Haute Performance). La chromatographie permet de séparer des molécules selon leur affinité différentielle pour une phase stationnaire et une phase mobile. Dans notre cas, la phase stationnaire est un support solide de type silice greffé par des chaînes alkyles en C_{18} et elle est associée à une élution en gradient multi étape à l'aide des mélanges de solvants **H_2O** / TFA et **CH_3CN** / TFA. La particularité de la CLHP est que la pression utilisée permet d'avoir une meilleure discrimination des molécules à purifier. Pour ce travail, nous avons utilisé une CLHP Spectra System avec une colonne semi-préparative Phenomex de type Jupiter C18, pilotée par ordinateur pour permettre de suivre graphiquement en temps réel l'absorbance et de permettre de recueillir les différentes molécules visualisées par un pic d'absorbance (**Figure 17**).

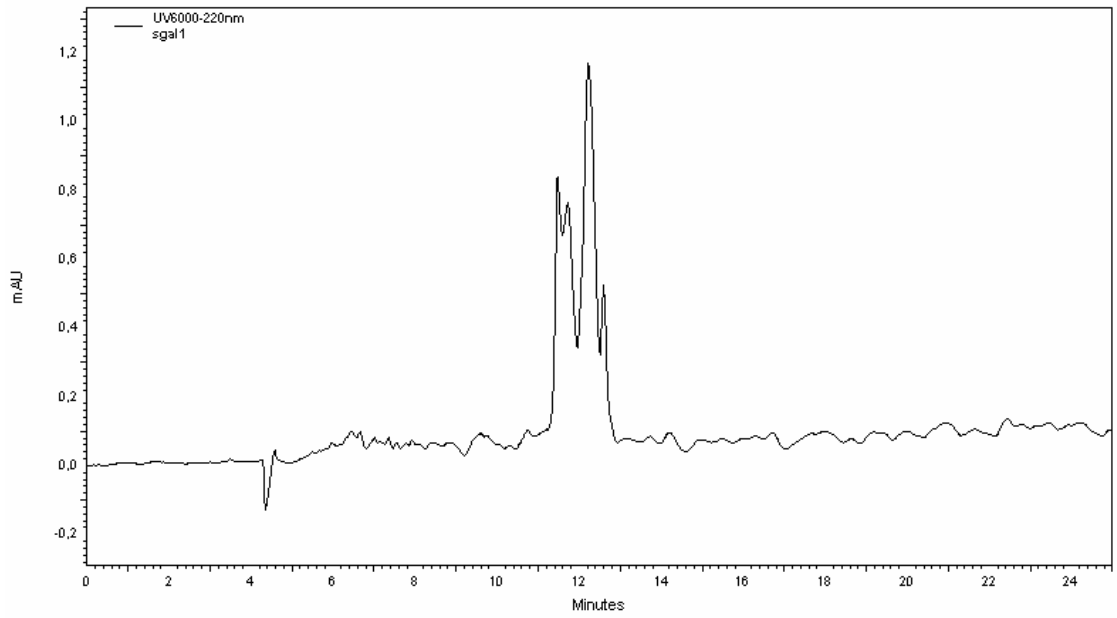


Figure 17 : Chromatogramme du peptide Sgal,
la flèche indique le pic correspondant à la séquence synthétisée.

Une petite quantité de chacun des pics est envoyée au service de spectroscopie de masse du CNR (Centro Nazionale di Ricerca) de Naples pour identifier les produits et leur degré de pureté. Lorsque la synthèse s'est avérée exacte, on peut purifier l'ensemble des produits provenant de la synthèse.

Le peptide **CVGVAPGC** cyclique et le peptide de 32 résidus dérivant de la partie spécifique de l'EBP ont été synthétisés respectivement par le laboratoire du Professeur Jean WALLACH (Lyon, France) et par l'entreprise Ansynth Service B.V. (Roosendaal, Pays-Bas) selon la même méthode de synthèse peptidique.

A.2. Spectroscopie de Dichroïsme Circulaire Electronique (DCE)

A.2.a. Principe du pouvoir rotatoire

D'après sa définition ondulatoire, la lumière décrite comme une onde électromagnétique se déplaçant dans le vide à la vitesse de la lumière *c*, est caractérisée par un champ électrique ($\vec{E}$), et un champ magnétique ondulatoire ($\vec{B}$) perpendiculaires à la direction de la propagation. Une lumière est polarisée rectilignement quand $\vec{E}$ oscille toujours dans un même plan appelé **plan de vibration**. Le plan perpendiculaire à $\vec{E}$ dans lequel oscille le champ magnétique $\vec{B}$ est appelé **plan de polarisation** (**Figure 18**).

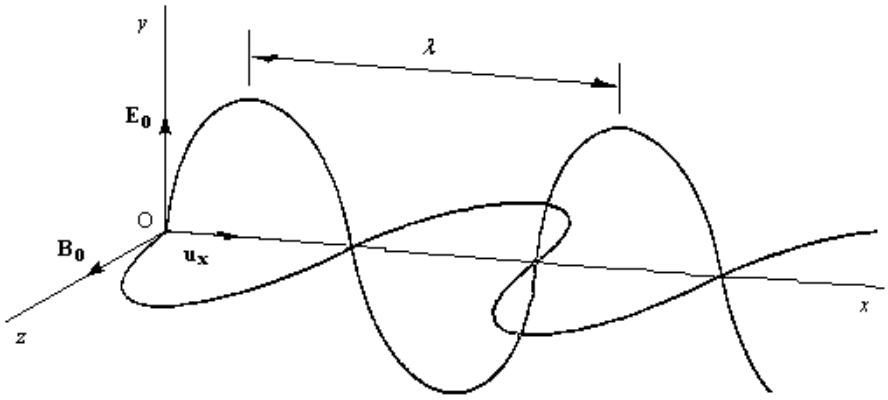


Figure 18 : Schéma d'une onde plane monochromatique définie par son champ électrique ($\vec{E}$) et son champ magnétique ondulatoire ($\vec{B}$).

Dans la lumière naturelle, le vecteur champ électrique $\vec{E}$ peut prendre toutes les directions perpendiculaires à la direction de propagation; la lumière est dite non polarisée (**Figure 19a**). Il est cependant possible de la polariser artificiellement, c'est-à-dire de fixer le vecteur champ électrique $\vec{E}$ pour tous les photons (**Figure 19b**).

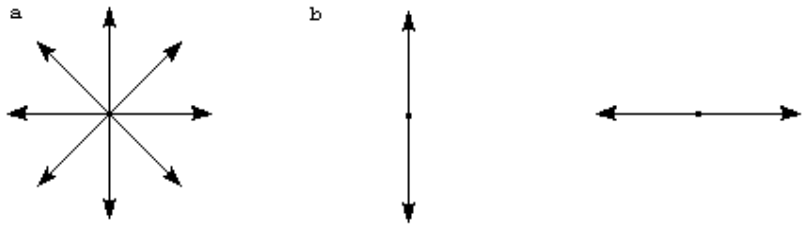


Figure 19 : Directions du vecteur champ électrique $\vec{E}$
a) lumière naturelle non polarisée,
b) lumière polarisée verticalement ou horizontalement.

Lorsqu'une lumière polarisée interagit avec une molécule non symétrique, c'est-à-dire chirale, cela donne lieu au phénomène appelé **activité optique**. Concrètement, l'activité optique se manifeste par une modification des propriétés du faisceau de lumière incident après son passage au travers d'une molécule chirale. La théorie cinématique de Fresnel permet d'expliquer l'origine du pouvoir rotatoire (**Figure 20**). Le vecteur champ électrique $\vec{E}$ peut être décomposé en deux composantes E_G et E_D, de même amplitude et de même fréquence mais tournant en phase en sens inverse.

Dans un milieu inactif (**Figure 20a**) les champs se déplacent à la même vitesse. A la sortie du milieu, le plan dans lequel vibre $\vec{E}$ n'a pas changé de direction. Un milieu actif possède des

propriétés physiques dissymétriques (non isotropes) pour les champs E_G et E_D (**Figure 20b**). Cette dissymétrie se traduit, à l'échelle macroscopique, par des indices de réfraction n_G et n_D différents. La vitesse de E_G est alors différente de celle de E_D :

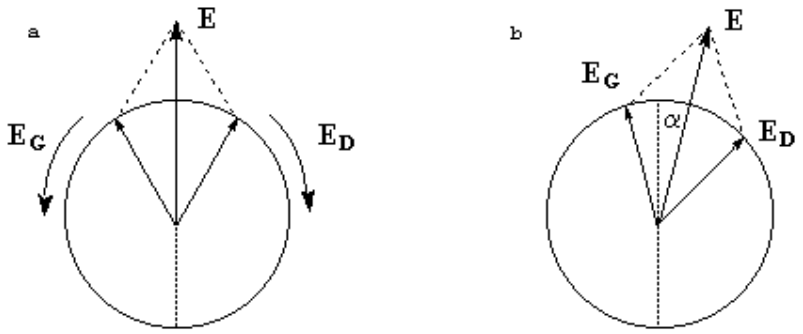
$$V_G = \frac{c}{n_G} \quad \text{et} \quad V_D = \frac{c}{n_D}$$

avec c : vitesse de la lumière dans le vide.

Après le passage dans milieu actif, c'est-à-dire possédant des propriétés de chiralité, les champs présentent un déphasage l'un par rapport à l'autre. Tout en étant encore polarisé rectilignement, le plan de vibration a tourné d'un angle α dont la valeur est donnée par la formule de Fresnel :

$$\alpha = \frac{\pi \, l \, (n_G - n_D)}{\lambda}$$

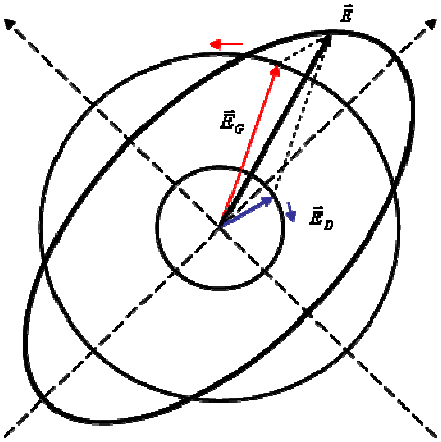
avec α : pouvoir rotatoire, l : épaisseur du milieu actif traversée et λ : longueur d'onde de la lumière.



<u>Figure 20</u> : Décompositions du vecteur champ électrique $\vec{E}$
a) milieu optiquement inactif b) milieu optiquement actif.

A.2.b. Principes du dichroïsme circulaire

Le phénomène de dichroïsme circulaire a lieu lorsque les deux composantes E_G et E_D diffèrent par leur absorption. La lumière résultante, qui est la somme des deux composantes, n'est plus polarisée linéairement, mais elliptiquement (**Figure 21**).



<u>Figure 21</u> : Représentation de l'ellipticité de la lumière résultant du phénomène dichroïque.

Le dichroïsme circulaire est donc défini par la différence d'absorption ΔA du champ électrique $\vec{E}$ polarisé alternativement circulairement à gauche et à droite :

$$\Delta A = A_G - A_D$$

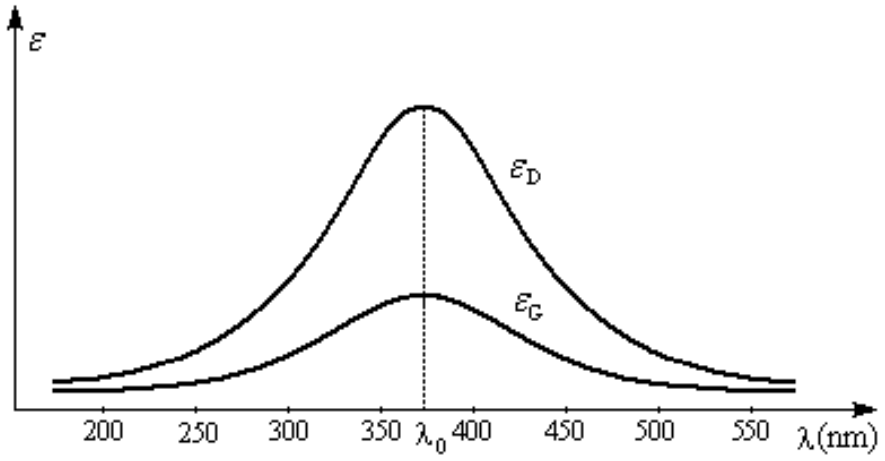
D'après la loi de Beer-Lambert, on a :

$$A = \varepsilon \, l \, C$$

avec ε : coefficient d'extinction molaire; l : parcours optique et C : concentration molaire.

Le pouvoir dichroïque ou **dichroïsme circulaire** Δε (**Figure 22**) est donc :

$$\Delta\varepsilon = \varepsilon_G - \varepsilon_D = \frac{\Delta A}{C \, l}$$



<u>Figure 22</u> : Coefficients d'extinction molaire ε_G, ε_D et dichroïsme circulaire Δε en fonction de la longueur d'onde λ.

En pratique, on parle peu de valeur de dichroïsme circulaire, mais plutôt d'ellipticité molaire (en degré.cm^2.dmol^{-1}) :

$$[\theta] = \frac{100M}{C\,l} = 3300\ \Delta\varepsilon$$

avec M : masse moléculaire; C : concentration molaire et l : le parcours optique.

Afin de bien pouvoir comparer les résultats obtenus, on utilise plutôt l'ellipticité molaire par résidu (en degré.cm^2.dmol^{-1}) :

$$[\theta]_{MRW} = \frac{[\theta]}{N}$$

avec N : nombre de résidus.

Les applications du dichroïsme circulaire en biologie structurale sont multiples. A condition d'avoir une molécule (acide nucléique, polysaccharide ou protéine) chirale et qui absorbe dans les longueurs d'onde de la lumière excitatrice (UV, visible, IR), on peut :

- étudier l'effet de l'environnement (solvant, température, salinité …) sur sa structure,
- réaliser des études dynamiques de dénaturation et de renaturation,
- étudier les effets avec d'éventuels ligands,
- effectuer des études d'interaction acides nucléiques – protéine ou protéine – protéine,
- déterminer les pourcentages de structures secondaires globales d'une protéine.

La zone d'absorption des protéines provient en grande partie de la fonction amide de la liaison peptidique. Deux transitions électroniques principales, observables dans l'UV lointain, sont ainsi à l'origine du signal de dichroïsme circulaire électronique : la transition n-π* autour de 230 nm et la transition π-π* aux environs de 200 nm. Par contre, bien que le C_α des acides aminés soit asymétrique, le signal qu'il génère est négligeable devant ceux du chromophore peptidique. En fait, le spectre de dichroïsme circulaire ne dépend que de la géométrie relative des liaisons peptidiques, c'est-à-dire des structures secondaires de la protéine.

Il est à noter que les résidus aromatiques comme la phénylalanine, la tyrosine ou l'histidine, ainsi que les ponts disulfures présentent eux aussi des signaux de dichroïsme circulaire qui, si ils peuvent être négligés dans le cas de protéines, ne peuvent l'être dans le cas de petits peptides.

A.2.c. Aspect pratique du dichroïsme circulaire électronique

L'un des avantages du dichroïsme circulaire électronique est que l'on peut travailler en solution diluée avec de faibles quantités de produits. Nous avons donc utilisé 2 types de solvants qui, évidemment, ne possèdent pas d'asymétrie, ni une absorbance trop importante dans le domaine UV des longueurs d'onde étudiées : l'eau et le TFE (2,2,2-trifluoroéthanol) qui est un solvant connu pour favoriser les structures secondaires ordonnées des protéines. La concentration des peptides utilisée est de 0,1 $mg.mL^{-1}$. Les longueurs d'ondes UV balayées vont de 190 nm à 250 nm, avec un pas de 0,1 nm. L'ensemble des mesures a été effectué dans une cuve en quartz de parcours optique de 0,1 cm.

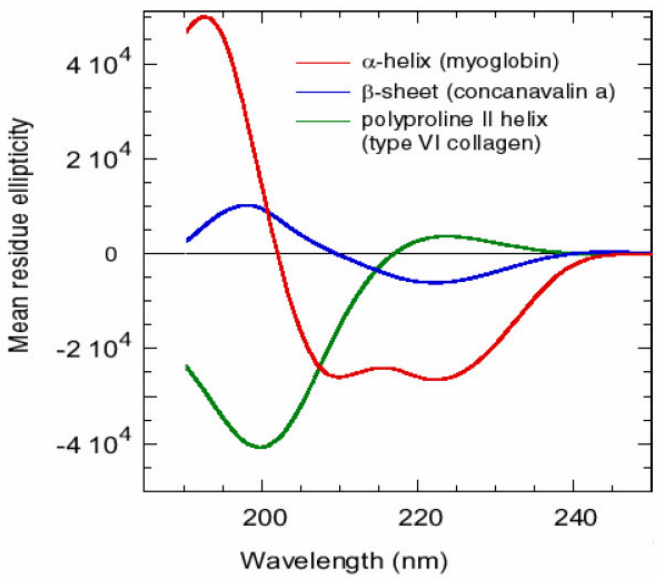
Les spectres du peptide Sgal et des peptides dérivant de l'élastine ont été enregistrés lors de mon stage en Italie sur un dichrographe JASCO 600 équipé d'un bain thermostaté HAAKE. Le peptide de 32 résidus dérivant de la partie spécifique de l'EBP a été étudié sur le dichrographe JASCO 810 du laboratoire, dont la température de l'échantillon est contrôlée par un thermostat à effet Peltier. Pour diminuer le rapport signal sur bruit, 16 accumulations ont été effectuées pour chacune des expériences.

A.2.d. Spectres caractéristiques des structures secondaires des protéines

L'hélice α possède le spectre le plus caractéristique car il présente 2 bandes négatives d'intensité équivalente à 222 et 208 nm correspondant respectivement aux transitions n-π^* et π-π^*, ainsi qu'un maximum positif à 190 nm (**Figure 23**).

Il est à noter que l'allure du spectre de l'hélice α n'est que très peu modifiée par les conditions expérimentales ou par la nature des résidus formant l'hélice α.

Le spectre des brins β est moins intense que celui de l'hélice α. Il est caractérisé par une bande négative autour de 216 nm due à la transition n-π^* et une bande positive autour de 195 nm (transition π-π^*) qui est difficilement observable car elle est dans une zone de longueurs d'ondes où la détection du signal est souvent très bruitée (**Figure 23**).



<u>Figure 23</u> : Spectres DCE caractéristiques des structures secondaires : hélice α, feuillet β et hélice polyproline II (Wallace et al. 2001).

L'allure du spectre du feuillet β dépend énormément de l'environnement et de la nature des résidus qui influent sur l'agencement et la structure des brins constituant le feuillet.

Une structure polyproline II est plus difficile à mettre en évidence car comme pour le spectre d'une protéine ne possédant pas de structure secondaire (**Figure 24**), le spectre obtenu présente une bande négative vers 195 - 200 nm (**Figure 23**). Une étude en température est nécessaire pour pouvoir distinguer ce type de structure : plus la température est abaissée, plus une bande positive apparaît autour de 217 nm. Cet effet de la température sur le spectre d'une protéine ne comportant pas de structure secondaire n'est en revanche pas observable.

A l'heure actuelle, seuls les coudes β de type I et II sont connus dans la littérature comme étant bien caractérisés en dichroïsme circulaire, leurs spectres étant reproduits en Figure 24 (Perczel et al. 1993). Leur étude n'est possible que sous forme de peptide, car dans une protéine, l'effet dichroïque des coudes β est noyé dans ceux des structures secondaires correspondant aux hélices α et aux brins β. Le type I a naturellement une allure de spectre similaire à celui de l'hélice α puisque l'hélice α peut être considérée en fait comme un enchaînement de coudes β voisins du type I. Le coude β de type II est bien caractérisé par la présence d'une bande positive vers 200 nm.

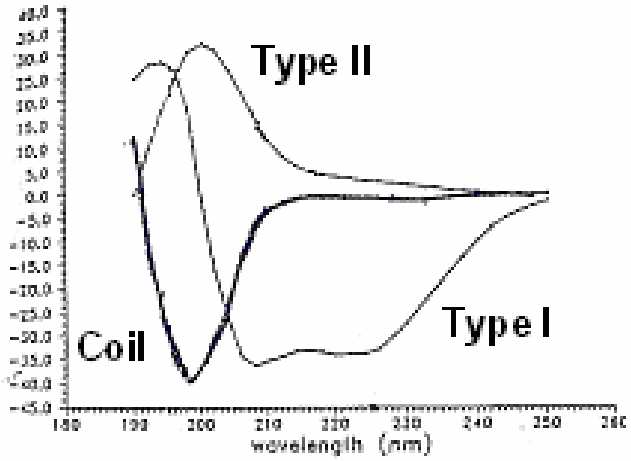


Figure 24 : Spectres DCE de référence des coudes β de type I et II et des boucles (random coil) (Perczel et al. 1993).

Le coude β de type VIII est en cours de caractérisation complète par notre laboratoire (Fuchs et al. 2002; Fuchs et al. 2005) ; il présente un signal proche de celui d'une structure non ordonnée avec une bande négative autour de 200 nm associée à un épaulement vers 220 nm (**Figure 25**).

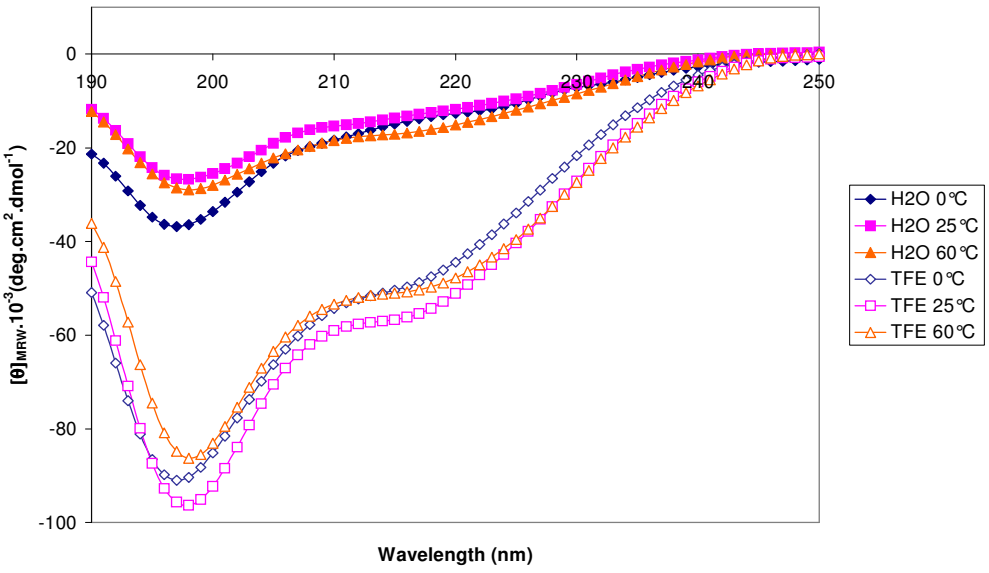


Figure 25 : Spectres DCE dans l'eau et le TFE du peptide GDNP (structure en coude β de type VIII, Fuchs et al. 2005).

A.2.e. Analyse de la structure secondaire par dichroïsme circulaire

L'analyse de la structure secondaire d'une protéine par dichroïsme circulaire repose sur 3 hypothèses :

- la fonction amide du squelette protéique est la seule à contribuer au spectre ; on néglige les éventuelles contributions des résidus aromatiques ou des ponts disulfure,
- les contributions des diverses structures secondaires sont additives,
- chaque élément de structure secondaire ne produit qu'un seul type de spectre.

Plusieurs types de méthode d'analyse sont disponibles suivant le choix de l'algorithme, des spectres de référence et des contraintes prises en compte. Dans ce travail, nous avons choisi d'utiliser le logiciel CDPRO (Sreerama et Woody. 2000), rassemblant les programmes ContinLL (Provencher et Glockner 1981), Selcon (Sreerama et Woody 1993) et CDsstr (Johnson 1999) qui effectuent des décompositions du spectre suivant les bandes caractéristiques des différentes structures secondaires à partir d'un jeu de protéines globulaires de référence. Pour chacune des méthodes, un spectre théorique est calculé puis comparé au spectre expérimental précédemment obtenu, ce qui permet de choisir la méthode la plus adéquate. Le taux d'hélice α est bien estimé quelle que soit la méthode utilisée. Pour les brins β, la qualité de la détermination est plus variable suivant la méthode. Pour les autres types de structures, les résultats sont plus aléatoires. Ce type de méthode reposant sur une banque de protéines globulaires, dont la structure est connue, n'est certainement pas strictement applicable aux peptides de moins de vingt résidus.

L'étude de la structure d'un peptide se fait encore le plus souvent par comparaison (allure générale, bandes caractéristiques …) du spectre obtenu avec les spectres de références des différents types de coude β ou des différentes structures secondaires précédemment décrits (**Figures 23 - 25**).

Une étude en température est souvent nécessaire pour affiner les conclusions structurales.

Enfin, il est possible d'utiliser les résultats de dichroïsme circulaire pour étudier l'interaction existant entre une protéine ou un peptide et son ligand. En effet, puisque la contribution de chaque liaison peptidique est additive, deux molécules ensemble devraient donner un spectre égal à la somme de leur spectre individuel si leur conformation n'est pas modifiée par la présence de l'autre. Si ce n'est pas le cas, cela signifie qu'un changement de structure est observable et donc qu'il y a interaction de l'une avec l'autre.

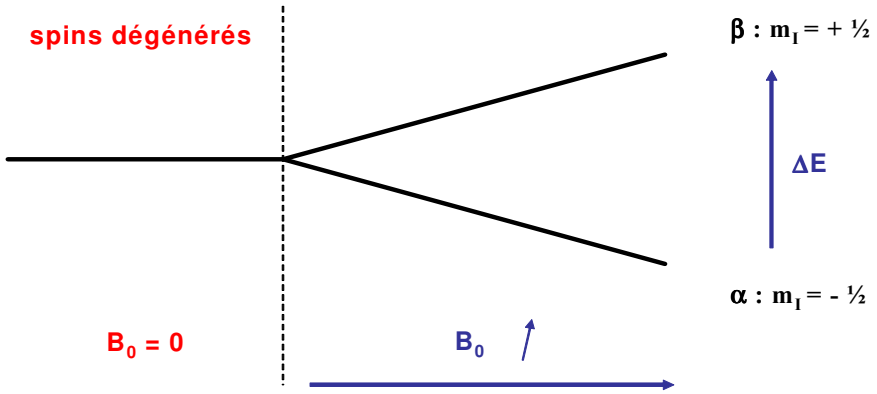
A.3. Spectroscopie de Résonance Magnétique Nucléaire (RMN)

La spectroscopie de Résonance Magnétique Nucléaire est une méthode de choix pour l'étude structurale des peptides et des protéines en solution puisqu'elle permet d'obtenir des informations à l'échelle du résidu. La majorité des structures obtenues sont calculées sur la base des distances interprotons, issues des mesures de type NOE, et des angles dièdres de torsion, provenant des constantes de couplages J ^{1}H - ^{1}H.

A.3.a. Aspect théorique

La RMN est une technique basée sur l'absorption du rayonnement électromagnétique par les noyaux de certains atomes (**^{1}H, ^{13}C, ^{15}N, ^{31}P** …). Pour être observables en RMN, les noyaux des atomes doivent être constitués d'un nombre impairs de protons et/ou de neutrons de manière à avoir un nombre de spin non nul.

Un noyau possédant un spin non nul se comporte magnétiquement comme un petit aimant. Placé dans un champ magnétique (B_0), il peut y prendre deux orientations : dans l'une, son spin est parallèle au champ et de même sens, dans l'autre, il lui est parallèle, mais de sens opposé (orientation antiparallèle). A ces deux orientations sont associés deux niveaux quantifiés d'énergie : le plus bas correspondant à l'orientation parallèle, qui est donc la plus stable. Hors champ magnétique ($B_0 = 0$), les spins n'auront pas d'orientation particulière, ils ont la même énergie, on dit qu'ils sont dégénérés.



<u>Figure 26</u> : Energies des états de spin en corrélation avec le champ B_0.

Dans le cas de la RMN de l'atome d'hydrogène, celui ci n'étant constitué que d'un seul proton, son spin I est égal à celui du proton I = ½ ce qui correspond à deux valeurs du nombre de spin $m_I = \pm ½$ (voir ci-dessus **Figure 26**) : $m_I = + ½$ correspond à l'état de spin parallèle au champ magnétique et $m_I = - ½$, à l'état de spin antiparallèle au champ magnétique (ces 2

valeurs correspondent respectivement aux fonctions d'onde propres α et β , provenant de la résolution de l'équation de Schrödinger appliquée au proton).

La séparation de l'énergie ΔE est proportionnelle à l'intensité de champ magnétique B_0 et crée la condition nécessaire à l'existence d'une transition spectroscopique, constituant ainsi la base de la spectroscopie RMN.

Le proton occupera l'état α qui est l'état fondamental (l'état de plus faible énergie) avec une plus grande probabilité que l'état β qui est l'état excité (d'énergie plus importante). Plus le champ B_0 augmente, plus la différence d'énergie ΔE entre les deux états de spin augmente. Le peuplement des deux niveaux d'énergie est décrit par la distribution de Boltzmann :

$$\frac{N_\beta}{N_\alpha} = e^{-\frac{\Delta E}{k_b T}}$$

avec N_α : nombre de spins à l'état fondamental, N_β : nombre de spins à l'état excité, ΔE : énergie de transition entre les deux états, K_b : constante de Boltzmann et T : température absolue en K.

Ainsi, pour induire la transition vers l'état d'énergie plus élevé, il faut donner au système une énergie : $\Delta E = h \nu_0$,

$$h \upsilon_0 = \gamma \hbar B_0$$

avec υ_0 : fréquence de résonance du noyau, $\hbar$: constante de Planck divisée par 2π, et γ: rapport gyromagnétique.

Si un proton est donc excité avec une impulsion de fréquence suffisamment haute pour permettre la transition entre les niveaux d'énergie α et β, correspondant au passage du nombre de spins + ½ à - ½, il pourra en absorber l'énergie par le phénomène appelé résonance magnétique nucléaire.

A la fin de l'excitation, les protons se relaxent, c'est-à-dire reviennent à l'état initial en émettant un signal appelé interférogramme ou encore FID pour Free Induction Decay. Ce signal est enregistré puis est traité à l'aide d'une transformée de Fourier pour donner un spectre RMN exploitable et caractéristique de la nature de l'environnement du proton.

A.3.b. Instrumentation, mesure et interprétation

Les spectres des peptides ont été enregistrés sur des spectromètres de type Bruker DRX 500.

L'expérience RMN proprement dite repose sur une succession, appelée séquence, de trois voire quatre intervalles de temps :

- **la durée de préparation** : le système de spins à étudier est préparé pour l'expérience. Il est destiné à permettre le retour des noyaux excités à leur état d'équilibre entre deux exécutions successives de la séquence,
- **la durée d'évolution** : le système de spins évolue sous l'influence de divers facteurs,
- **la durée de mixage** : le système de spins subit une séquence d'impulsions,
- **la durée de détection** : c'est durant cette période que le signal modulé est acquis.

Les mesures des paramètres RMN permettent d'accéder à des informations structurales précises du système étudié :

- **expérience 2D J-Résolue Homonucléaire** : permet d'obtenir le déplacement chimique de chacun des protons et les constantes de couplages,
- **expérience COSY (COrrelation SpectroscopY)** : engendre un échange de la magnétisation entre deux noyaux fortement couplés entre eux. Elle permet de savoir si deux protons sont vicinaux,
- **expérience TOCSY (Total Correlation SpectroscopY)** : donne les mêmes informations que l'expérience COSY, mais étant basée sur une autre séquence, permet une attribution complémentaire des pics correspondant aux échanges de magnétisation notamment en cas de "relais", c'est-à-dire en cas d'échange de magnétisation de deux protons par l'intermédiaire d'un troisième situé entre eux,
- **expérience NOESY (Nuclear Overhauser Effect SpectroscopY)** : permet d'accéder à une information concernant la proximité spatiale de deux protons (effet nucléaire Overhauser ou NOE), d'une distance allant de 2 à 5 Å. Cette séquence est très utilisée pour la détermination structurale des grosses molécules comme les protéines,
- **séquence ROESY (Rotating Overhauser Effect SpectroscopY)** : apporte des informations similaires à la séquence NOESY, en permettant elle-aussi d'avoir des informations concernant les distances entre certains protons. La séquence ROESY est surtout utilisée en complément de la séquence NOESY ou bien pour déterminer la structure de petites molécules,
- **séquence HSQC (Heteronuclear Single Quantum Coherence)** : permet d'obtenir des paramètres RMN d'un noyau peu sensible, comme le carbone, en observant le signal RMN

d'un noyau beaucoup plus sensible, typiquement un proton. La séquence HSQC est spécialement utilisée pour la détermination structurale de petites protéines car elle permet d'aider à l'attribution des atomes d'hydrogène relié à un atome de carbone.

Grâce à ces expériences RMN, on peut reconnaître, non seulement les protons proches dans l'espace, mais également déterminer la valeur des angles de torsion en utilisant les constantes de couplages ^3J entre protons vicinaux grâce à la loi de Karplus (Karplus 1959) :

$$^3J = A \cos^2 \theta + B \cos \theta + C$$

avec A, B et C : coefficients, dépendant de la nature de la liaison considérée et $\theta = \left| \phi - 60 \right|$.

Cette relation permet de déterminer l'angle dièdre ϕ des résidus d'une protéine ou d'un peptide grâce à la constante de couplage $^3J_{NH\text{-}H\alpha}$ entre le proton amide et le proton du C_α (Pardi et al. 1984) :

$$^3J_{NH\text{-}H\alpha} = 6{,}4 \cos^2 \theta - 1{,}4 \cos \theta + 1{,}9$$

A.4. Spectroscopie de diffusion Raman

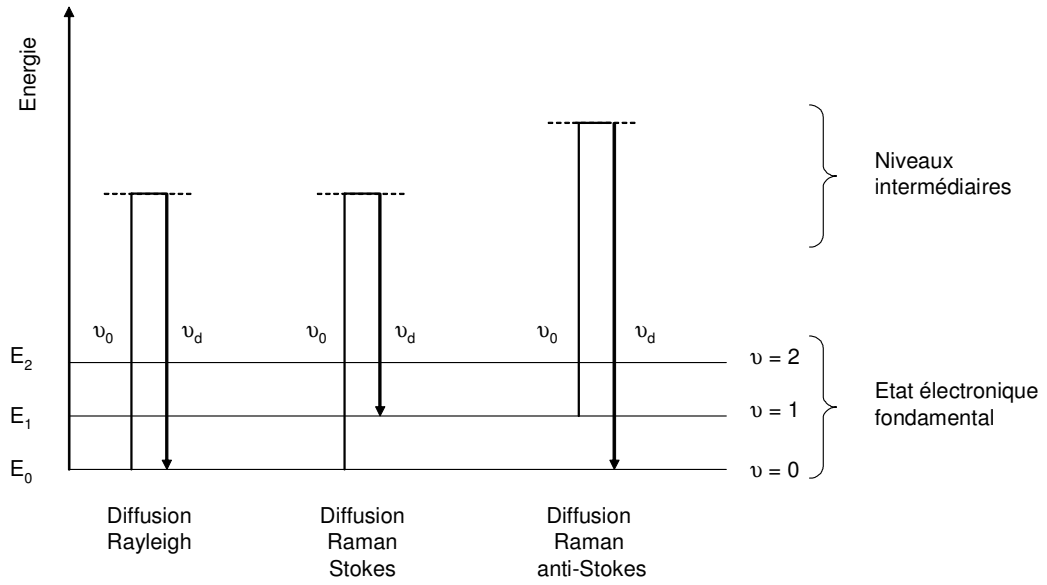
A.4.a. Aspect théorique

L'effet Raman est un phénomène de changement de longueur d'onde qui correspond à la diffusion inélastique de la lumière par un milieu matériel. Ce phénomène fut prédit par Smekal en 1923 (Smekal 1923), mais c'est Raman et Krishnan qui en 1928 l'ont mis en évidence expérimentalement (Raman et Khrisnan 1928).

Lorsqu'on éclaire un milieu (gaz, liquide ou solide) par un faisceau de lumière monochromatique, cette radiation excitatrice peut être transmise, réfléchie, absorbée ou diffusée par le milieu. Le phénomène de diffusion inélastique peut s'interpréter, de manière schématique, par des transitions moléculaires radiatives faisant intervenir des échanges d'énergie entre les photons et les molécules **(Figure 27**). Lorsqu'un photon de fréquence υ_0 irradie la molécule se trouvant au niveau d'énergie E_0, il peut communiquer son énergie $h\upsilon_0$ pour porter la molécule à un niveau énergétique intermédiaire instable. Cependant, si ce niveau ne correspond pas à un niveau énergétique quantifié stable, le système transite vers l'un des niveaux quantifiés stables en diffusant un photon de fréquence υ_d et d'énergie $h\upsilon_d$ (la diffusion peut s'effectuer dans toutes les directions de l'espace).

Le processus génère soit des photons décalés vers les basses fréquences ($\upsilon_d < \upsilon_0$), c'est le phénomène de diffusion Raman Stokes, soit des photons de plus hautes fréquences ($\upsilon_d > \upsilon_0$), il s'agit alors du phénomène de la diffusion Raman anti-Stokes.

Il existe également le phénomène de diffusion élastique (diffusion Rayleigh) pour laquelle le retour s'effectue au niveau initial et s'accompagne de la diffusion d'un photon de fréquence υ_d égale à celle du photon de la lumière incidente de fréquence υ_0.



<u>Figure 27</u> : Processus de diffusions élastique (Rayleigh) et inélastique (Raman).

La spectroscopie de vibration est liée aux transitions vibrationnelles de valeurs d'énergie ΔE (différence d'énergie entre le niveau initial et final de la transition). La spectroscopie Raman normale met en évidence les mouvements des vibrations fondamentales de la molécule associées à l'état électronique fondamental.

Une molécule comportant *N* atomes possède 3*N* degrés de liberté. En faisant abstraction des mouvements que sont les 3 translations et les 3 rotations d'ensemble (solide rigide), on dénombre (3*N* - 6) mouvements de déformations internes de types vibrations fondamentales. Dans le cas particulier d'une molécule linéaire, seules deux rotations d'ensemble sont discernables, il y a donc (3*N* - 5) vibrations fondamentales.

A.4.b. Instrumentation et interprétation

L'ensemble des radiations diffusées formant le spectre Raman-Stokes normal du peptide **C-VGVAPG-C**, cyclisé par un pont disulfure entre les deux cystéines, a été

enregistré au laboratoire sur un spectromètre MOLE S3000 (Jobin-Yvon) piloté par ordinateur *via* le logiciel SPECTRALINK. Un Laser Argon+ (SpectraPhysics) délivrant une longueur d'onde excitatrice de 514 nm est utilisé comme source monochromatique incidente.

Pour chacun des spectres obtenus, 100 acquisitions (moyenne temporelle) ont été effectuées pour augmenter le rapport signal sur bruit d'un facteur 10.

Un spectre Raman d'une protéine ou d'un peptide possède plusieurs régions caractéristiques donnant des informations structurales d'importance. Ainsi, la liaison peptidique plane **O=C-N-H** est à l'origine de diverses bandes amides dont en particulier :

(i) la bande amide I située dans le domaine de fréquences 1630 à 1700 cm^{-1} correspondant essentiellement à la vibration d'élongation de la double liaison covalente **C=O** (mouvement prépondérant de l'oxygène dans le plan),

(ii) la bande amide III située dans le domaine de fréquence 1230 à 1300 cm^{-1} correspondant en particulier à la vibration d'élongation de la liaison covalente C-N. et à la vibration de déformation angulaire **CNH** (mouvement prépondérant de l'hydrogène dans le plan.

Ces deux bandes amide I et III permettent de connaître par décomposition de leurs massifs de bandes, les pourcentages des structures secondaires (Alix et al. 1985; Berjot et al. 1987; Alix et al. 1988; Pedanou 1994; Alix 1997) comme les hélices α, les feuillets β , les coudes β et même ceux des hélices polyproline II (Bochicchio et al. 2004a).

L''utilisation des décompositions des bandes amide I et III pour connaître la structure d'un peptide de quelques résidus seulement doit être faite avec prudence car l'étude de ces bandes a surtout été réalisée dans le cas de protéines de plusieurs dizaines ou centaines de résidus.

Nous n'avons donc utilisé la spectroscopie Raman que dans le cas de l'étude du peptide **C-VGVAPG-C** cyclique, l'intérêt étant de permettre l'étude du pont disulfure (-liaison **-S-S-**) de ce peptide. Un pont disulfure est en effet caractérisé géométriquement par 3 angles dièdres appelés τ_1, τ_2 et τ_3. Lorsque l'un de ces angles à une valeur de $\pm 90°$, on dit qu'il est en conformation *gauche* (G), lorsque cette valeur d'angle dièdre est de $\pm 180°$, il est dit en conformation *trans* (T). A cause de la géométrie particulière d'un pont disulfure, l'angle τ_2 est toujours *gauche*, seuls τ_1 et τ_3 peuvent être soit gauche soit trans. On obtient ainsi 4 conformations locales possibles de ponts disulfures dont 3 réellement différentes : GGG, TGT et GGT, qui est équivalent à TGG. A chacun de ces types de structure correspond une bande caractéristique de la liaison **S-S**: à 510 cm^{-1}, 525 cm^{-1} et 540 cm^{-1} respectivement pour GGG, GGT (ou TGG) et TGT.

A.5. Spectroscopie de fluorescence

A.5.a. Aspect théorique

Le phénomène de fluorescence est caractérisé par le fait que les transitions mises en jeu ne sont pas toutes de type radiatif, contrairement à l'effet Raman normal ou de résonance (pour lequel l'excitation porte la molécule à un niveau intermédiaire instable supérieur au premier niveau électronique excité),

En fluorescence, l'énergie $h\upsilon_0$ apporté par les photons est suffisante pour dépasser le premier état électronique excité. Une partie de l'énergie est dissipée sous forme non radiative, le plus souvent sous forme de chaleur, puis la molécule retombe à son état électronique fondamental en émettant un photon de fréquence υ_f inférieure à la celle υ_0 du photon excitateur (**Figure 28**). La fluorescence se traduit ainsi par une augmentation de la longueur d'onde λ_f du photon émis (effet Stokes).

Pour observer le phénomène de fluorescence, il faut que l'échantillon possède un ou plusieurs fluorophores qui sont presque exclusivement des groupements aromatiques. Dans le cas des protéines, seuls les résidus tyrosine, phénylalanine, histidine et tryptophane possèdent des propriétés de fluorescence.

L'intensité de la fluorescence émise est directement liée à son environnement.

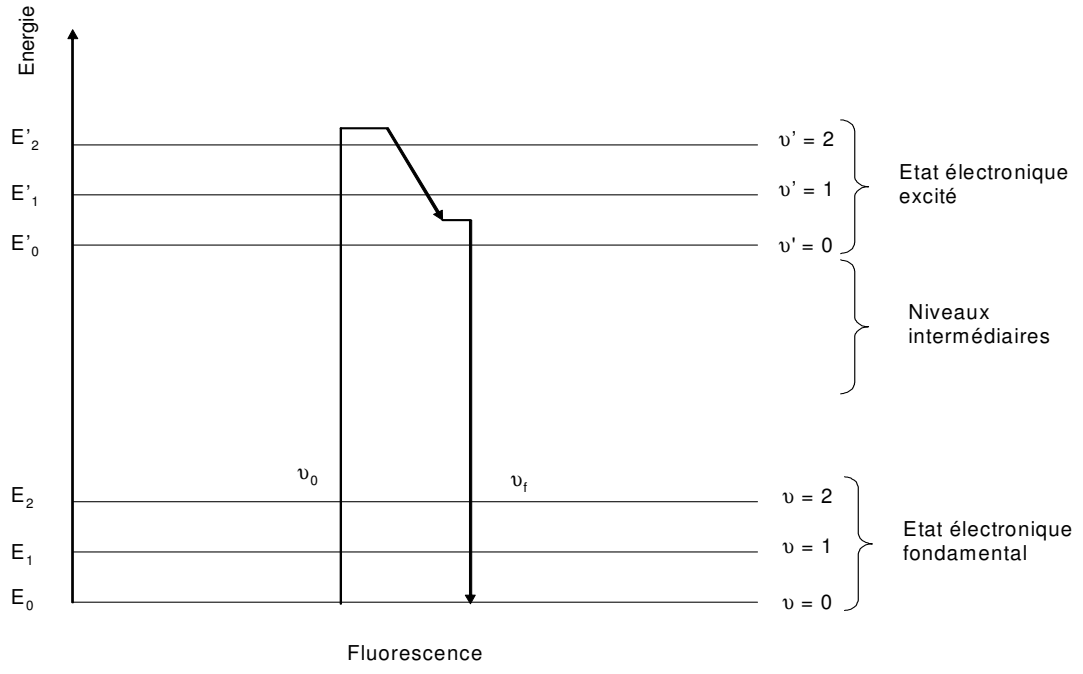


Figure 28 : Niveaux énergétiques et transitions moléculaires impliqués dans le phénomène de fluorescence.

A.5.b. Instrumentation et interprétation

Les mesures de fluorescence ont été effectuées, au Département de Chimie, Université de Potenza, (Italie) sur un spectrofluorimètre Fluorolog-3 (Jobin Yvon-Spex Instruments S.A. Inc) à la température ambiante et en accumulant 9 mesures.

La spectroscopie de fluorescence a été utilisée dans ce travail pour étudier l'interaction entre le peptide Sgal et les peptides dérivant de l'élastine **YGVG** et **VGVAPG**. Le peptide Sgal ne possédant pas de fluorophore, c'est le suivi de la fluorescence des ligands qui a été réalisé. Dans le cas du peptide **YGVG**, c'est la fluorescence de résidus tyrosines qui va être mesurée. Pour **VGVAPG**, nous avons laissé, durant la synthèse peptidique, le groupement Fmoc attaché à la valine qui constitue ainsi le fluorophore (**Tableau 02**).

Peptides	**Longueurs d'onde**	
	excitation (nm)	**émission (nm)**
YGVG	274	303
Fmoc-VGVAPG	264	304

Tableau 02 : Longueurs d'onde d'excitation et d'émission des fluorophores utilisés (en gras dans les peptides étudiés).

Pour ces deux peptides, nous avons fait varier la concentration du peptide Sgal tout en mesurant l'intensité de la fluorescence. Il a été soustrait à cette fluorescence l'intensité de la fluorescence intrinsèque F_0 due à la présence du peptide Sgal seul en solution :

$$\Delta F = F - F_0$$

La valeur ΔF_{max} est la valeur de fluorescence maximale obtenue lorsque tous les sites de fluorescence sont en interaction avec le peptide Sgal. ΔF_{max} est obtenue en traçant le graphique 1/ΔF en fonction de 1/[Sgal] (**Figure 29a**).

On peut connaître la valeur de la concentration de fluorophore, ce qui équivaut à la concentration en EDP lié (**YGVG** ou Fmoc-**VGVAPG**) en utilisant la relation suivante :

$$\frac{\Delta F_{max}}{\Delta F} = \frac{[EDP]_{lié\ \ max}}{[EDP]_{lié}}$$

la concentration maximale en EDP lié correspond en fait à la concentration en peptide Sgal introduite. La relation précédente devient alors:

$$\frac{\Delta F_{max}}{\Delta F} = \frac{[SGAL]_{introduit}}{[EDP]_{lié}}$$

soit :

$$[EDP]_{lié} = \frac{\Delta F \times [SGAL]_{introduite}}{\Delta F_{max}}$$

La concentration en EDP libre est calculée comme étant la différence entre la concentration introduite et la concentration en EDP lié :

$$[EDP]_{libre} = [EDP]_{introduite} - [EDP]_{lié}$$

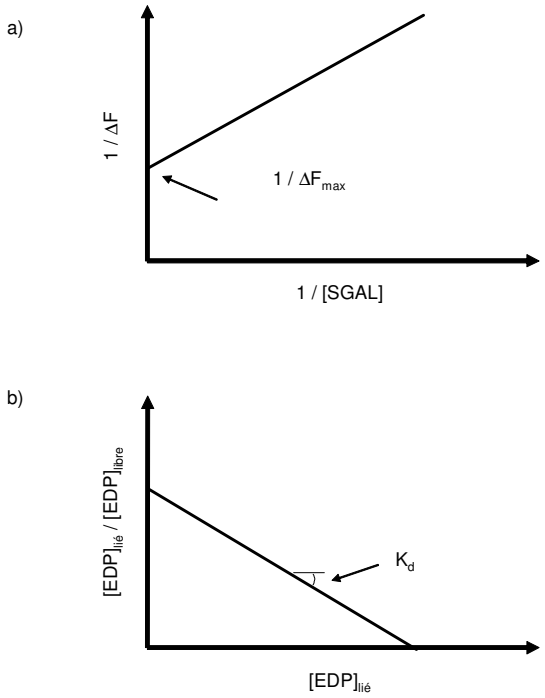


Figure 29 : Détermination de la constante de dissociation K_D :
a) 1 / ΔF en fonction de 1 / [Sgal],
b) grahique de Scatchard : $[EDP]_{lié}$ en fonction de $[EDP]_{lié} / [EPD]_{libre}$.

On peut, en accord avec l'équation suivante, tracer un graphique de Scatchard, c'est-à-dire, $[peptide]_{lié} / [peptide]_{libre}$ en fonction de $[peptide]_{lié}$:

$$\frac{[EDP]_{lié}}{[EDP]_{libre}} = -\frac{[EDP]_{lié}}{K_d} + e_0$$

La constante de dissociation du complexe K_d est calculée comme étant le coefficient directeur de pente de la droite ainsi tracée (**Figure 29b**).

B. METHODES D'ETUDES THEORIQUES
MODELISATION MOLECULAIRE

La modélisation moléculaire est le terme générique regroupant les méthodes qui permettent de simuler le comportement d'un système de particules. La taille du système étudié peut aller d'une simple molécule diatomique jusqu'à des macromolécules biologiques de plusieurs dizaines de milliers d'atomes.

La modélisation moléculaire peut ainsi permettre de décrire le comportement électronique des atomes et des molécules pour expliquer leurs réactivités, comprendre les processus de repliement d'une protéine ou encore expliquer l'importance de certains acides aminés d'un site catalytique enzymatique.

Pour réaliser ce type d'étude, il est nécessaire de déterminer une expression de l'énergie d'interaction des atomes du système moléculaire en fonction de leurs positions relatives. Pour ce faire, les méthodes de modélisation moléculaire reposent entre autres sur deux approches différentes appelées Mécanique Quantique et Mécanique Moléculaire, pour évaluer cette énergie d'interaction.

La mécanique quantique repose sur la résolution de l'équation de Schrödinger pour décrire les interactions au niveau atomique en tenant compte des électrons et des noyaux de chaque atome du système :

$$H \Psi = E \Psi$$

où Ψ décrit la fonction d'onde du système, H est l'opérateur Hamiltonien (contenant les énergies cinétique et potentielle du système) s'appliquant à Ψ et E l'énergie du système.

Cependant, la résolution analytique exacte de cette équation n'est possible que pour les atomes monoélectroniques tels que les hydrogénoïdes. Pour des systèmes de taille plus importante, on a recours en particulier à l'approximation de Born-Oppenheimer qui permet de découpler les termes de l'opérateur Hamiltonien. Comme la masse du noyau atomique est significativement plus grande que la masse des électrons, le mouvement des électrons peut ainsi être considéré comme découplé du mouvement des noyaux, ce qui signifie que les électrons s'adaptent instantanément aux changements de positions des noyaux. La fonction d'onde totale de la molécule peut désormais s'écrire :

$$\Psi_{tot} \text{ (noyaux, électrons)} = \Psi \text{ (électrons)} \, \Psi \text{ (noyaux)}$$

L'énergie totale d'une molécule est égale à la somme de l'énergie nucléaire (la répulsion électrostatique entre les noyaux) et de l'énergie électronique. L'énergie électronique comprend l'énergie cinétique et potentielle des électrons se déplaçant dans le champ électrostatique des noyaux.

$$E_{tot} = E \text{ (électrons)} + E \text{ (noyaux)}$$

Quand l'approximation de Born-Oppenheimer est utilisée, les noyaux sont supposés fixes quand ce sont les mouvements des électrons qui sont en jeu. Pour chaque réarrangement de la position des noyaux, l'équation de Schrödinger est résolue seulement pour les électrons.

C'est ce type de procédure qui est la base sur laquelle repose les méthodes de modélisation moléculaire quantique *ab initio*. Le terme *ab initio* provient du fait que les calculs sont effectués indépendamment de toute donnée expérimentale, mais ils permettent d'avoir accès à l'ensemble des propriétés qui découlent des distributions électroniques, en particulier pour les réactions chimiques dans lesquelles les liaisons sont brisées et/ou formées.

Cependant, les calculs qui en résultent sont encore extrêmement longs et limités aux petits systèmes de quelques dizaines d'atomes à cause de l'ensemble des électrons pris en compte et de la complexité des équations à résoudre pour l'évaluation de l'énergie d'interaction. C'est pourquoi ce type d'évaluation de l'énergie d'interaction est incompatible avec nos systèmes protéiques qui peuvent comprendre plusieurs dizaines de milliers d'atomes.

Nous avons donc eu recours à des calculs de Mécanique Moléculaire qui reposent sur les principes de la mécanique classique. Dans ce type de méthode, les électrons ne sont pas pris en compte, mais seulement les noyaux, ce qui permet de diminuer grandement les temps de calculs.

Il existe également des méthodes hybrides de Mécanique Moléculaire / Mécanique Quantique qui ont été développées pour essayer de profiter des avantages des deux méthodes. Le principe de base étant d'utiliser la Mécanique Quantique pour une zone restreinte d'intérêt comme le site actif d'une protéine et de traiter le reste du système avec la Mécanique Moléculaire pour gagner du temps de calcul. Bien que ces méthodes aient déjà prouvé leur efficacité, elles ne sont pas réellement adaptées à nos problématiques.

B.1. Mécanique Moléculaire

La Mécanique Moléculaire permet le calcul de l'énergie d'interaction d'un système en fonction des seules positions des noyaux, en ignorant ainsi le mouvement des électrons. L'approximation de Born-Oppenheimer est aussi utilisée, cependant, le mouvement des noyaux n'est plus décrit par un Hamiltonien quantique comme en Mécanique Quantique, mais par les descriptions de la mécanique classique où les atomes sont assimilés à des points massiques (éventuellement chargés) et les liaisons chimiques à des ressorts mécaniques.

B.1.a. Fonction d'énergie potentielle

La plupart des propositions de modèles moléculaires est basée sur l'utilisation d'une fonction d'énergie potentielle empirique V, fonction des coordonnées atomiques r_i , avec $i = 1, 2 \ldots N$ avec N : le nombre d'atomes constituant le système. La fonction d'énergie est donc une fonction du jeu de coordonnées cartésiennes R spécifiant les positions de tous les atomes à partir desquelles sont calculées les coordonnées internes du système.

La fonction énergie potentielle est représentable comme une hypersurface à 3N dimensions.

Elle est définie empiriquement comme étant la somme des différentes contributions traduisant les interactions physiques inter et intramoléculaires du système étudié.

La fonction d'énergie est généralement décrite en termes internes reliés directement aux liaisons covalentes des atomes (interactions liantes) et en termes externes traduisant les interactions entre atomes non liés de manière covalente (interactions non liantes).

Nous avons, durant cette étude, utilisé la fonction d'énergie potentielle (associée à un champ de forces) du programme de modélisation moléculaire CHARMM (Chemistry at HArdvard Molecular Mechanics) (Brooks et al. 1983), dont la forme générale est la suivante :

$$V(R) = V_{\text{liantes}} + V_{\text{non liantes}}$$

Avec :

$$V_{\text{liantes}} = V_{\text{liaisons}} + V_{\text{angles}} + V_{\text{impropres}} + V_{\text{dièdres}}$$

$$V_{\text{liantes}} = \sum_{liaisons} k_b (b - b_0)^2 + \sum_{angles} k_\theta (\theta - \theta_0)^2 + \sum_{impropres} k_\omega (\omega - \omega_0)^2 + \sum_{dièdres} k_\phi^n (1 + \cos(n\phi - \delta))$$

et :

$$V_{\text{non liantes}} = V_{\text{électrostatiques}} + V_{\text{van der Waals}}$$

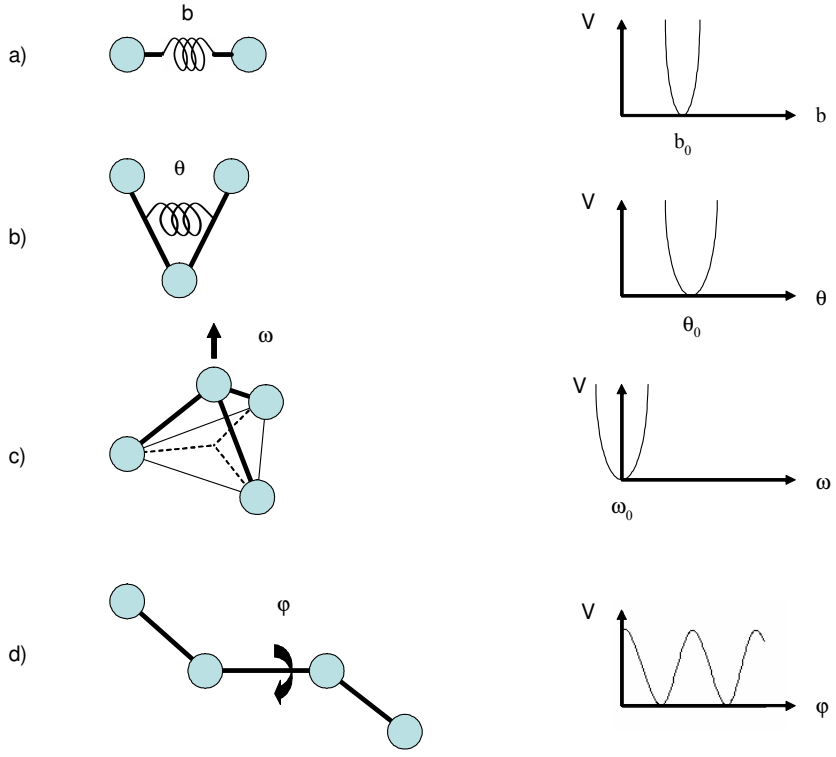
$$V_{\text{non liantes}} = \sum_{i,j} 4\varepsilon_{i,j} \left(\left(\frac{\sigma_{i,j}}{r_{i,j}} \right)^{12} - \left(\frac{\sigma_{i,j}}{r_{i,j}} \right)^{6} \right) + \sum_{i,j>1} \frac{q_i q_j}{4\pi\varepsilon r_{i,j}}$$

La forme complète de la fonction potentielle de CHARMM est donc :

$$V(R) = \sum_{liaisons} k_b (b - b_0)^2 + \sum_{angles} k_\theta (\theta - \theta_0)^2 + \sum_{impropres} k_\omega (\omega - \omega_0)^2 + \sum_{dièdres} k_\phi^n (1 + \cos(n\phi - \delta)) +$$

$$\sum_{i,j} 4\varepsilon_{i,j} \left(\left(\frac{\sigma_{i,j}}{r_{i,j}} \right)^{12} - \left(\frac{\sigma_{i,j}}{r_{i,j}} \right)^{6} \right) + \sum_{i,j>1} \frac{q_i q_j}{4\pi\varepsilon r_{i,j}}$$

Les 2 premiers termes rendent compte de la déformation des longueurs des liaisons covalentes (élongation), et des angles formés par 2 liaisons covalentes (ciseau). Le 3ème terme décrit les variations d'angles impropres (non descriptibles par le terme précédent) (**Figure 30c**). Ces 3 premiers termes sont qualifiés de durs car un faible déplacement par rapport aux valeurs de référence (b_0, θ_0, ω_0) conduit à d'importantes variations d'énergie. Ainsi, leur description se fait à l'aide d'une approximation harmonique.
Le 4ème terme représente la torsion autour d'une simple liaison covalente, fonction périodique développée en série de Fourier, qui se traduit par une variation sinusoïdale de l'énergie potentielle (**Figure 30d**) comme cela a été démontré pour l'éthane (Pitzer 1983).



<u>Figure 30</u> : Termes de la fonction d'énergie potentielle de CHARMM pour les interactions liantes : a) liaisons, b) angles, c) angles impropres et d) torsions dièdres.

Les 2 derniers termes traduisent les interactions entre atomes non liés de manière covalente. Dans le champ de forces de CHARMM, seuls les atomes séparés de plus de 3 liaisons sont considérés comme pouvant avoir des interactions non liées. Ce choix se justifie par le fait que les interactions, entre atomes séparés par une, deux ou trois liaisons, sont suffisamment bien décrites par les termes de déformation des longueurs de liaisons et d'angles formés par 2 liaisons covalentes.

Ainsi, le 5ème terme de la fonction d'énergie est le potentiel de Lennard-Jones qui traduit les interactions de van der Waals. Il est constitué d'un terme répulsif en r^{-12} représentant le principe d'exclusion de Pauli et la répulsion internucléaire à courte distance. Il contient également un terme attractif en r^{-6} représentant les interactions d'origine dipolaire dues aux fluctuations des nuages électroniques de chacun des atomes (forces de London).

Enfin le 6ème et dernier terme est le potentiel coulombien qui traduit les interactions électrostatiques entre les différentes paires d'atomes distantes de $r_{i,,j}$ où ε représente la constante diélectrique du milieu.

Il est également possible d'ajouter un terme supplémentaire représentant explicitement l'énergie des liaisons hydrogène. Il est cependant classiquement admis que les liaisons hydrogène sont majoritairement de nature électrostatique et il a été démontré qu'elles sont correctement représentées par les termes de van der Waals et électrostatiques.

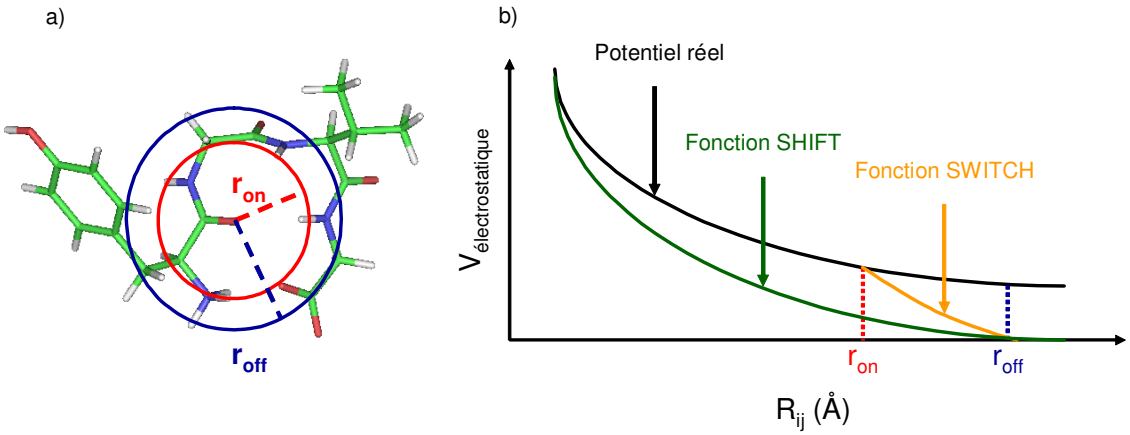
B.1.b. Traitement des interactions non liantes

On a vu que le terme électrostatique était calculé pour chaque paire d'atomes (*i, j*) possible du système étudié, ce qui est extrêmement coûteux en temps de calcul. Il est possible de diminuer le nombre d'interactions non liantes pour chaque atome en ayant recours à une approximation qui implique une limite au delà de laquelle les interactions électrostatiques ne seront plus prises en compte : on définit une sphère de rayon r_{off}, appelé rayon de coupure ou de troncature ou encore "cut-off" (**Figure 31a**).

Cependant, une coupure brutale du traitement des interactions électrostatiques risque d'entraîner des artéfacts de simulation et des instabilités énergétiques. Pour remédier à cela, on peut utiliser la fonction SHIFT qui multiplie les termes d'interaction non liées par une fonction s'annulant en r_{off} ainsi que sa dérivée, ce qui permet une décroissance douce de la fonction au niveau du rayon de coupure r_{off}.

L'utilisation de la fonction SWITCH est également possible; un nouveau rayon r_{on} est défini jusqu'auquel les interactions électrostatiques sont calculées normalement. Au-delà de r_{on}, les

termes des interactions non liées sont modifiés pour permettre également l'annulation de la fonction au niveau de r_{off}.



<u>Figure 31</u> : Traitement des interactions non liantes :
a) définition des rayons de coupure (troncature, cut-off)
b) effets des fonctions SHIFT et SWITCH sur le potentiel électrostatique.

Les deux types de fonctions SHIFT et SWITCH (Brooks et al. 1985) sont :

- Fonction SHIFT :

$$\text{pour} \quad r \le r_{off} \qquad S(r) = \left(1 - \left(\frac{r}{r_{off}}\right)^2\right)^2$$

$$\text{pour} \quad r > r_{off} \qquad S(r) = 0$$

- Fonction SWITCH :

$$\text{pour} \quad r \le r_{on} \qquad S(r) = 1$$

$$\text{pour} \quad r_{on} < r \le r_{off} \qquad S(r) = \frac{\left({r_{off}}^2 - r^2\right)^2 \left({r_{off}}^2 + 2r^2 - 3{r_{on}}^2\right)}{\left({r_{off}}^2 - {r_{on}}^2\right)^3}$$

$$\text{pour} \quad r > r_{off} \qquad S(r) = 0$$

D'autres fonctions dérivant des précédentes ont été proposées permettant un calcul plus précis des interactions non liantes. :

- Fonction FORCESHIFT (Brooks et al. 1985) :

$$\text{pour} \quad r \le r_{off} \qquad S(r) = \left(1 - \frac{r}{r_{off}}\right)^2$$

$$\text{pour} \quad r > r_{off} \qquad S(r) = 0$$

- **Fonction FORCESWITCH** (Steinbach et al. 1994) : où l'énergie potentielle est donnée par:

pour $r \le r_{on}$

$$V(r) = cr_{off}^4\left(r^{-1} - r_{off}^{-1}\right)\frac{\left(r_{off}^2 - 3r_{on}^2\right)}{\left(r_{off}^2 - r_{on}^2\right)^3} + 6cr_{on}^2 r_{off}^2\frac{\left(r_{off} - r\right)}{\left(r_{off}^2 - r_{on}^2\right)^3} - c\left(r_{off}^3 - r^3\right)\frac{\left(r_{on}^2 - r_{off}^2\right)}{\left(r_{off}^2 - r_{on}^2\right)^3} + \frac{2c\left(r_{off}^5 - r^5\right)}{5\left(r_{off}^2 - r_{on}^2\right)^3}$$

$$+ 8cr_{on}^2 r_{off}^2\frac{\left(r_{off} - r_{on}\right)}{\left(r_{off}^2 - r_{on}^2\right)^3} - \frac{8c\left(r_{off}^5 - r_{on}^5\right)}{5\left(r_{off}^2 - r_{on}^2\right)^3}$$

pour $r_{on} < r \le r_{off}$ $\quad V(r) = -\int_{r_{off}}^{r} S(r)F(r)dr$

pour $r > r_{off}$ $\quad V(r) = 0.$

avec S(r) : fonction SWITCH précédemment décrite qui effectue la troncature de la force F(r).

Les fonctions SWITCH sont généralement utilisées pour les interactions de van der Waals car elles ne modifient pas les interactions à courte distance. Il est possible d'utiliser les fonctions SWITCH ou SHIFT pour les interactions électrostatiques, tout en choisissant avec soin les valeurs de r_{off} ou de r_{on}. Il ne faut cependant pas perdre de vue que de tels choix influencent de manière non négligeable les propriétés du système étudié (Loncharich et al. 1989) (Norberg et Nilsson 2000).

D'autres méthodes ont été développées pour le traitement des interactions non liantes. La méthode de sommation d'Ewald (Ewald 1921) est celle qui est la plus couramment utilisée. Elle permet de calculer les interactions électrostatiques d'un atome de charge q_i avec tous les autres atomes du système, ainsi qu'avec toutes ses images périodiques. Il n'y a donc plus de troncature des interactions électrostatiques ce qui permet d'avoir une meilleure modélisation de ce type d'interaction. "Particle Mesh Ewald" est une méthode efficace de calcul de la sommation d'Ewald (Darden et al. 1993) (Essmann et al. 1995) qui a été implémentée dans CHARMM, mais elle demande un effort calculatoire beaucoup plus important et de plus, nécessite d'ajouter des contre-ions afin de neutraliser électriquement le système étudié.

B.1.c. Traitement du solvant

Au niveau expérimental, il est évident que le solvant dans lequel baigne une molécule influence grandement ses propriétés physico-chimiques. La manière dont est traitée

la solvatation influe directement sur les interactions non liantes et est donc importante pour modéliser au mieux le comportement des molécules étudiées.
Il existe 2 manières de simuler la présence de solvant : en mimant sa présence implicitement dans les calculs de la fonction d'énergie (solvant implicite), ou bien en introduisant effectivement les molécules de solvant dans le système étudié (solvant explicite).

α. Solvatation implicite

La manière la plus simple de traiter le solvant consiste à donner à la constante diélectrique ε du potentiel coulombien la valeur correspondant au solvant désiré. Ainsi, pour l'eau on utilisera $\varepsilon = 80$, tandis que pour mimer l'intérieur d'une protéine on fixe ε entre 3 et 4.

D'autres modèles de solvant implicite ont été mis au point, que l'on peut classer selon 2 approches différentes : les **modèles empiriques** et les modèles basés sur un **continuum électrostatique**.
Les modèles **empiriques** partent du principe que l'énergie libre de solvatation du soluté est une somme des contributions de groupes d'atomes. Chaque contribution est évaluée par une fonction linéaire de sa surface accessible au solvant (Eisenberg et McLachlan 1986) (Schiffer et al. 1992) ou par le volume occupé à l'intérieur de la zone de solvatation (Gibson et Scheraga 1967) (Colonna-Cesari et Sander 1990). Ces modèles incorporent les composantes hydrophobes et électrostatiques de la solvatation, mais omettent de prendre en compte les interactions entre les charges au sein du solvant, qui doivent être introduites sous forme d'un terme supplémentaire.

Les modèles de solvant basés sur un **continuum électrostatique** définissent le soluté et le solvant comme étant des régions ayant des constantes diélectriques différentes (**Figure 32**) et l'énergie libre de solvatation est calculée en résolvant l'équation de Poisson-Boltzmann (Kirkwood et Westheimer 1938). Cette méthode traite de manière rigoureuse le continuum électrostatique, et a été appliquée avec succès à des problèmes biologiques (Bashford et Case 2000) (Simonson 2001). Toutefois l'inconvénient principal de cette méthode est qu'elle est trop coûteuse en temps de calcul pour être utilisée de manière courante. Pour pallier ce problème, plusieurs traitements analytiques et semi-analytiques de modèles de continuum électrostatique ont été proposés (Simonson 2001) (Bashford et Case 2000) (Feig et Brooks 2004) (Feig et al. 2004).

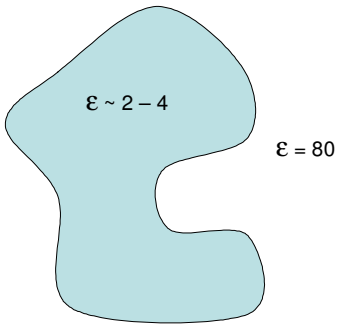


Figure 32 : Distribution des constantes diélectriques dans un modèle de solvant implicite basé sur le continuum électrostatique; exemple d'une protéine (en bleu).

Durant cette étude, nous avons eu recours à un modèle de solvant implicite **empirique**, particulièrement adapté à l'étude des peptides et des petites protéines, mis au point par Ferrara (Ferrara et al. 2002) et implémenté dans le module **SASA** (Solvent Accessible Surface Area) de CHARMM. Il est basé sur une approximation du calcul de la surface accessible au solvant (Wodak et Janin 1980), et l'utilisation de deux paramètres de solvatation. Ces deux paramètres de solvatation prennent en compte seulement les groupements hydrophobes et hydrophiles tandis que la solvatation des atomes d'hydrogène est négligée. Tout comme dans un modèle antérieur de solvant implicite nommé EEF1 (Effective Energy Function 1), les résidus chargés sont neutralisées et un écrantage dépendant des distances des charges du soluté, $\varepsilon(r) = 2r$ est utilisé (Lazaridis et Karplus 1999).

Nous avons également utilisé un modèle de **continuum électrostatique** "Generalized Born" (Qui et al. 1997) implanté dans le module **GBORN** de CHARMM (Generalized BORN) (Dominy et Brooks 1999). Il est basé sur le calcul de l'énergie libre de solvatation qui est calculée comme étant la somme d'un terme de cavité solvant-solvant (G_{cav}), d'un terme de van der Waals soluté-solvant (G_{vdW}) et d'un terme de polarisation électrostatique soluté-solvant (G_{pol}) :

$$G_{sol} = G_{cav} + G_{vdW} + G_{pol}$$

G_{cav} et G_{vdW} sont calculés ensemble en évaluant la surface accessible au solvant de la protéine:

$$G_{cav} + G_{vdW} = \sum_{k=1}^{N} \sigma_k SA_k$$

G_{pol} dépend de la constante diélectrique ε, de la charge q et du rayon de Born α des atomes :

$$G_{pol} = -166{,}0\left(1 - \frac{1}{\varepsilon}\right)\frac{q^2}{\alpha}$$

Nous avons également utilisé le module **GBMV** (Generalized Born using Molecular Volume) basé sur la même approche, mais qui évalue la surface accessible au solvant de la même manière que le module SASA (Lee et al. 2003). Cette méthode est plus précise, mais également plus coûteuse en temps de calcul.

β. Solvatation explicite

Traiter le solvant de manière explicite, c'est-à-dire ajouter les molécules de solvant au système étudié, est très certainement la manière la plus rigoureuse d'appréhender le comportement d'une molécule. Cependant le coût calculatoire est extrêmement important puisqu'il faut tenir compte dans le calcul de la fonction d'énergie potentielle, non seulement des termes qui concernent le soluté, mais aussi de ceux qui concernent le solvant, entraînant ainsi des temps de calcul autrement plus longs que ceux résultant d'un simple traitement implicite du solvant. De plus, il est nécessaire d'avoir les paramètres physico-chimiques de la molécule de solvant calculés spécifiquement pour le champ de forces que l'on utilise.
Nous avons utilisé deux types de solvant spécialement paramétrés pour le champ de forces de CHARMM : l'eau avec le modèle TIP_3P (Jorgensen et al. 1983) et le TFE (Chen et al. 2002).

Pour étudier une molécule en solvant explicite, il est nécessaire de la solvater, c'est-à-dire de la plonger entièrement dans une "boîte de solvant". Cette boîte de solvant, quelle que soit sa forme (cubique, parallélépipédique ou octaédrique tronquée) doit avoir une densité très proche de la valeur expérimentale du solvant. Pour cela, il est nécessaire que le nombre de molécules N de solvant soit calculé précisément suivant le volume désiré de la boîte de solvant :

$$N = \frac{\mathrm{N}_0 \times d \times V}{M}$$

avec N_0 : nombre d'Avogadro, d : densité du solvant, V : volume de la boite de solvant et M : masse molaire du solvant.
Par exemple, dans le cas d'une boîte d'eau cubique de 50 Å d'arête, le nombre de molécules d'eau nécessaires pour avoir $d_{eau} = 1.0$ (conditions standard) est :

$$N = \frac{6{,}02.10^{23} \times 1.0 \times \left(50.10^{-9}\right)^3}{18.10^{-3}} \approx 4181 \text{ molécules d'eau}$$

Pour éviter non seulement que la densité du solvant au voisinage du soluté ne varie, mais aussi que la zone de solvant ne soit au contact direct du vide (effets de bord), des conditions périodiques aux limites (Brooks et al. 1988) sont imposées. Le système soluté-solvant

constitue la boîte primaire qui est reproduite sous forme d'images dans toutes les directions de l'espace pour un total de [(3 x 9) - 1] = 26 images (**Figure 33**).

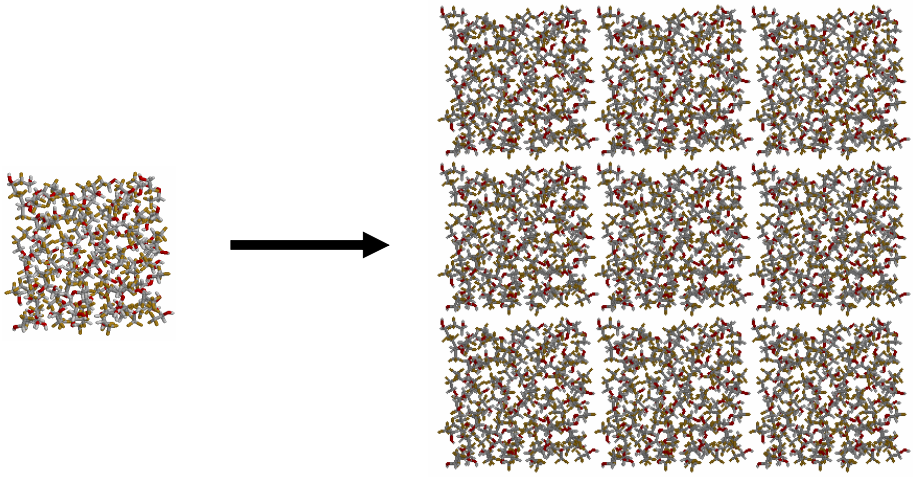


Figure 33 : Création des 26 images autour du système initial pour appliquer les conditions périodiques aux limites.

La fonction d'énergie potentielle n'est calculée que pour le système central et ses interactions éventuelles avec les atomes des boîtes images. Les éventuels mouvements de chacun des atomes sont ensuite répercutés dans toutes les images. Ainsi, si une molécule de solvant sort du système principal, par le jeu des conditions périodiques aux limites, l'image de cette même molécule est réintroduite dans la boîte primaire ce qui permet d'avoir toujours une densité constante du solvant.

Il est à préciser que, puisque les molécules de solvant sont explicitement présentes, elles vont directement participer à l'évaluation de l'énergie du système et donc la constante diélectrique ε aura une valeur fixée à 1.

B.1.d. Méthodes de minimalisation

A l'exception de petits systèmes très simples, la fonction d'énergie potentielle d'un système est compliquée du fait du grand nombre de variables et paramètres dont elle dépend. On parle ainsi d'hypersurface d'énergie potentielle.

Pour un système à N atomes, l'énergie potentielle est une fonction de 3N coordonnées cartésiennes ou de 3N - 6 coordonnées internes (non redondantes). Il est donc impossible de représenter graphiquement l'hypersurface d'énergie potentielle d'une molécule, sauf pour les cas simples où elle n'est fonction que de une ou deux variables.

La modélisation moléculaire s'intéresse particulièrement aux points minima de l'hypersurface d'énergie potentielle qui peuvent être plus ou moins nombreux suivant la complexité du système. En effet, ces points d'énergies minimales correspondent à des arrangements des atomes qui traduisent des états stables du système. Le point de plus basse valeur d'énergie est appelé minimum global d'énergie par opposition aux autres minima qui sont définis comme étant des minima locaux d'énergie. La recherche, pour identifier les géométries du système correspondant à des points minima de l'hypersurface d'énergie potentielle, s'effectue à l'aide d'algorithmes de minimalisation. Ces algorithmes vont graduellement changer les coordonnées des atomes du système pour l'amener de plus en plus proche de la structure correspondant à un point minimum de l'hypersurface d'énergie potentielle.

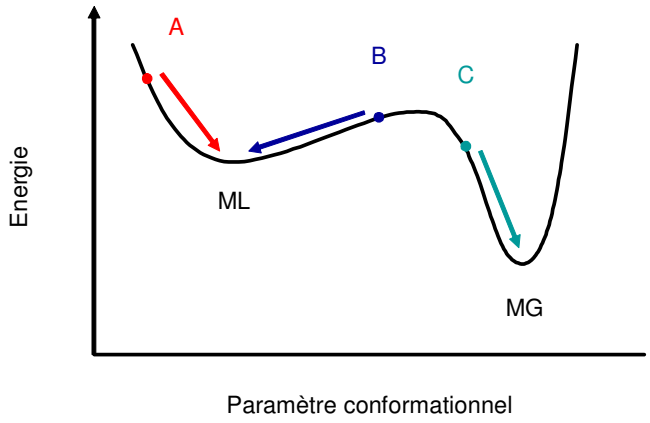


Figure 34 : Représentation d'une surface d'énergie selon un paramètre conformationnel. Les structures A, B "tombent", après minimalisation, dans le minimum local (ML) et la structure C dans le minimum global (MG).

Il est à noter que le minimum global d'énergie est très difficile à trouver car les paramètres conformationnels à partir desquels la minimalisation va être effectuée sont primordiaux : les méthodes de minimisation ne font en général que proposer le minimum d'énergie le plus proche (**Figure 34**).

Les 3 algorithmes de minimalisation auxquels nous avons eu recours durant cette étude sont ceux les plus utilisés dans les travaux de modélisation moléculaire.

α. *Méthode "Steepest Descent" SD*

L'algorithme "Steepest Descent", SD, (plus grande pente) est basé sur une méthode de minimalisation du premier degré. C'est une méthode "grossière" qui converge difficilement vers le minimum d'énergie car elle a tendance à osciller autour de celui-ci. Elle est cependant conseillée pour éliminer les contacts stériques entre des atomes trop proches les

uns des autres. Elle est utilisée en première approximation et généralement en association avec l'une des deux autres méthodes décrites ci-dessous.

β. Méthode du Gradient Conjugué :

Cette méthode tient compte des informations de la fonction d'énergie potentielle accumulées durant les itérations précédentes pour optimiser les itérations suivantes (Fletcher et Reeves 1964). C'est une méthode robuste qui converge plus rapidement que la méthode SD car elle n'a pas le même comportement oscillatoire au voisinage du point minimum d'énergie. Même si elle converge bien lorsque le système est très éloigné du minimum, il conseillé de l'utiliser seulement après avoir déjà eu recours à la méthode SD.

γ. Méthode ABNR :

La méthode ABNR (Adopted Basis Newton-Raphson) est en fait une variante de la méthode Newton-Raphson qui utilise à la fois les dérivées premières et secondes des gradients pour localiser les minima d'énergie. Elle est particulièrement adaptée pour les systèmes de grande taille. C'est une méthode qui converge très bien au voisinage du minimum d'énergie; aussi il est donc préférable d'utiliser au préalable la méthode SD pour s'approcher dans un premier temps d'un minimum d'énergie. L'inconvénient majeur de cette technique est son coût plus élevé en temps de calcul, que celui des autres méthodes précédemment décrites.

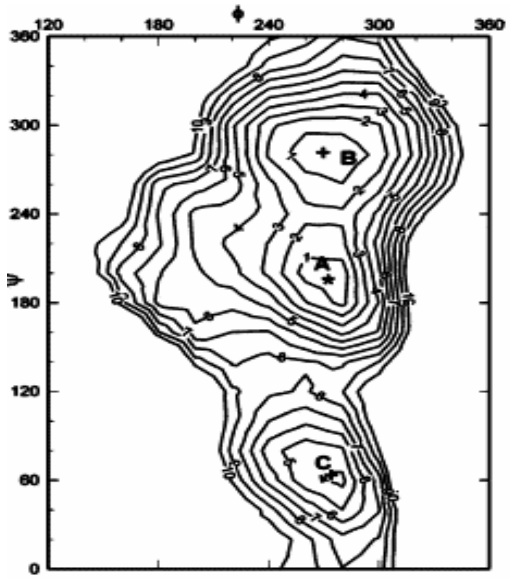
B.2. Exploration de l'hypersurface d'énergie potentielle

Nous avons vu qu'effectuer une minimalisation à partir d'une structure quelconque permettait d'obtenir un minimum local de l'hypersurface d'énergie potentielle et donc la structure associée. Il est cependant beaucoup plus intéressant de connaître la structure associée au minimum global de l'hypersurface d'énergie potentielle. En effet, cette conformation est la plus stable, donc la plus probable et devrait correspondre à la structure native. Il est donc *à priori* possible d'explorer la surface d'énergie potentielle en engendrant un panel de structures qui vont être minimalisées pour essayer d'atteindre le minimum global de l'hypersurface d'énergie potentielle. Pour cela, on peut générer de manière systématique toutes les conformations possibles de la molécule, c'est ce que l'on appelle la méthode de la carte adiabatique. On peut également générer un ensemble de conformations du système par

des changements, effectués au hasard, de la position des atomes, c'est ce type d'approche qui est à la base des méthodes de Monte-Carlo.

B.2.a. Carte adiabatique

Le terme adiabatique correspond à l'évaluation de l'énergie du système sans échange de chaleur avec l'extérieur.



<u>Figure 35</u> : Carte adiabatique d'un disaccharide : A indique le minimum global, B et C indiquent des minima locaux (Stortz 2004).

La **Figure 35** présente l'exemple d'un système ne dépendant que de deux angles ϕ et ψ.
Il est possible de décrire la structure d'une protéine ou d'un peptide à partir des angles dièdres ϕ et ψ du squelette protéique. La valeur de l'angle dièdre ω influence peu la structure car il est généralement contraint en position *trans* autour de 180°, la position *cis* n'étant pas énergétiquement favorable, elle n'est retrouvée que rarement dans la structure des protéines.
Il est bien évidemment impossible, pour une protéine, de tracer une véritable carte bidimensionnelle à cause du nombre trop important de résidus, donc de paramètres ϕ et ψ.
Cependant, par extension, cette méthode peut encore être appelée carte adiabatique car nous déterminons une valeur d'énergie associée à un ensemble de couples d'angles dièdres.
Générer l'ensemble des conformations d'une molécule protéique est donc théoriquement possible en échantillonnant les valeurs des angles dièdres ϕ et ψ du squelette protéique. Cependant si on prend le cas d'un seul acide aminé et si l'on génère toutes les valeurs d'angles dièdres entre - 180° et + 180° par pas de 1°, on obtient déjà 129.600 conformations.

Un petit peptide de 4 résidus conduirait ainsi à $2{,}8.10^{20}$ conformations. Si chaque minimalisation nécessite 1 ps (1.000 fs !) de calcul et d'écriture, on obtient un temps de calcul de l'ordre de 10 ans ! Il est évidemment possible d'augmenter le pas d'intégration pour diminuer le temps de calcul, mais même ainsi les échelles de temps de calculs sont trop importantes pour être envisageables dans le cas de notre étude. De plus, le risque principal de prendre un pas trop important est d'avoir une mauvaise description des conformations possibles.

Un bon compromis entre le temps de calcul et une description précise des conformations possibles que peuvent prendre des peptides a été trouvé en se basant sur l'analyse statistique des motifs structuraux observés au sein des protéines. Ainsi, les cartes de ϕ et ψ, ou cartes de Ramachandran (Ramachandran et al. 1963) mettent en évidence dix zones particulières (**Tableau 03**) où les résidus ont de grandes probabilités de se localiser (Derreumaux 1997).

Régions	ϕ	ψ	Structure secondaires
1	**- 60°**	**- 40°**	**hélice α**
2	- 120°	+ 120°	brin β
3	- 162°	- 162°	étendue
4	**- 60°**	**+ 150°**	**polyproline II**
5	- 90°	0°	coude β_1
6	- 110°	+ 70°	intermédiaire
7	+ 60°	+ 60°	hélice α_L
8	+ 90°	0°	coude β_2
9	+ 60°	- 120°	II'
10	+ 60°	- 60°	coude γ

Tableau 03 : Angles dièdres des 10 régions définies pour la carte adiabatique, en gras : régions accessibles à la proline.

Les structures secondaires régulières sont définies dans la région 1 pour l'hélice α, 2 pour le brin β et 4 pour la polyproline II. Il n'y a pas de région correspondant à l'hélice 3_{10} car les valeurs d'angles dièdres (ϕ = - 49° ; ψ = - 26°) sont trop proches de celles de l'hélice α.

Pour les coudes, il est nécessaire de définir plusieurs régions suivant leurs angles dièdres caractéristiques (Hutchinson et al. 1994).

Les seconds résidus des coudes β de type I, VIa, VIb et II' sont décrits par la région 5 ou β_1, tandis que les seconds résidus des coudes β de type II et I' sont décrits par la région 8 appelée β_2. La région 9 correspond aux angles dièdres du premier résidu du coude β de type II'. Enfin

la région 10 est la région de la carte de Ramachandran où l'on trouve le coude γ (Rose et al. 1985).
Pour une bonne description des régions visitées par les résidus, il a été également nécessaire d'inclure deux autres régions : la région 6, appelée intermédiaire (Abagyan et Totrov 1994) et la région 3 pour décrire les structures étendues.
Ce découpage de la carte de Ramachandran en ces 10 régions a déjà prouvé sa validité lors de différentes études (Derreumaux 2001; Floquet et al. 2004; Floquet et al. 2005).

La construction des peptides et l'ensemble des calculs de mécanique moléculaire ont été effectués grâce au champ de forces CHARMM.
Les peptides étudiés ont été construits en imposant à chacun de leurs résidus les valeurs d'angles dièdres des 10 régions, tout en laissant aux résidus N- et C-terminaux leurs valeurs par défaut. La proline fait exception car étant caractérisée par une conformation cyclique, elle ne peut adopter que les angles dièdres correspondant aux régions 1, 4 et 6 (hélice α, polyproline II et intermédiaire). Pour exemple : un peptide de k = 7 résidus, sans proline, présentera donc $10^{k-2} = 10^5$ conformations possibles avec cette méthode.

Trois étapes consécutives de minimalisations (100 pas de SD et 100 pas de GC) sont appliquées au système, durant lesquelles le solvant est traité avec le modèle de solvant implicite empirique, implémenté dans le module **SASA** (Solvent Accessible Surface Area) du programme CHARMM.
On impose aux angles dièdres ϕ et ψ du squelette protéique, dans la première étape, une contrainte de 50 kcal.mol^{-1} ; puis, dans la deuxième étape, une contrainte de 10 kcal.mol^{-1} et finalement pour la dernière étape plus aucune contrainte n'est appliquée. Ce protocole est utilisé de manière à avoir une minimalisation "douce" du système, les premières étapes ayant plutôt un but de "relaxation" des chaînes latérales des résidus constituant le peptide.
L'ensemble des conformations obtenues permet d'avoir une bonne description de l'hypersurface d'énergie potentielle puisque les valeurs retenues des angles dièdres des différents résidus correspondent aux régions les plus visitées de la carte de Ramachandran.
Malheureusement, si ce type de méthode est très performant, il est très dur de l'appliquer à des peptides ayant une taille supérieure à 10 résidus. En effet, pour un nombre k de résidus, le nombre de conformations à calculer est de 10^{k-2}, (un milliard pour k = 11 !) et le temps de calcul devient alors vite très important.

B.2.b. Simulation de Monte Carlo (MC)

Une simulation de Monte Carlo permet d'engendrer aléatoirement un ensemble de configurations d'un système en générant des changements de positions, d'orientations et de conformations. On évalue les énergies des différentes configurations du système pour obtenir celle de plus basse énergie (minimum global d'énergie de l'hypersurface d'énergie potentielle).

Le critère de comparaison des configurations du système le plus souvent utilisé est celui de Metropolis (Metropolis et al. 1953).

En pratique :

- une première configuration du système est générée ; son énergie E_1 calculée est l'énergie associée à l'état de référence
- une deuxième configuration du système est générée et son énergie E_2 est calculée.

Les valeurs d'énergie E_1 et E_2 sont comparées et deux cas de figure sont possibles :

- $\mathbf{E_1 \leq E_2}$: la nouvelle conformation est conservée et devient le nouvel état de référence,
- $\mathbf{E_1 > E_2}$: le critère de Metropolis entre en jeu. Un nombre aléatoire **R,** compris entre 0 et 1, est généré et :

la conformation est conservée si : $e^{-\frac{E_1-E_2}{k_bT}} > \mathrm{R}$

la conformation est rejetée si : $e^{-\frac{E_1-E_2}{k_bT}} < \mathrm{R}$

avec k_b : constante de Boltzmann et T : température absolue en K.

Nous avons utilisé pour les simulations de Monte Carlo, la version 4.0 du programme FANTOM (**Fa**st **N**ewton–**R**aphson **To**rsion **M**inimizer) (Schaumann et al. 1990) qui utilise l'algorithme de minimalisation de Newton-Raphson dans l'espace des angles de torsion (Abe et al. 1984).

Il est nécessaire de définir les angles de torsion qui vont être aléatoirement modifiés (von Freyberg et Braun 1993), puis minimaliser la conformation obtenue selon l'algorithme de Newton-Raphson.

La fonction énergie potentielle est décrite sous la forme :

$$E_{total} = E_{ECEPP/2} + E_{solv}$$

Le terme $E_{ECEPP/2}$ inclut les contributions des interactions de van der Waals, électrostatiques, des liaisons hydrogène, ainsi que celle des angles de torsion (Némethy et al. 1983).
Le terme de solvatation E_{solv} est calculé de la manière suivante (Eisenberg et al. 1986) :

$$E_{solv} = \sum_{i=1}^{N} \Delta\sigma_i ASA_i$$

avec ASA_i : surface accessible au solvant de l'atome i, $\Delta\sigma_i$: paramètres de solvatation atomique et N : nombre d'atomes contribuant au calcul de l'énergie de solvatation.
Nous avons utilisé les paramètres de solvatation des solvants eau, octanol, cyclohexane, calculés par Efremov et collaborateurs (Efremov et al. 1999)

B.3. Dynamique Moléculaire

La dynamique moléculaire permet d'avoir une description temporelle du comportement d'un système en s'appuyant sur la connaissance de la fonction énergie potentielle et la résolution de l'équation du mouvement de Newton pour chacun des atomes constituant le système étudié. On obtient ainsi un ensemble de "trajectoires atomiques" qui va spécifier comment les positions et les vitesses des atomes du système varient au cours du temps auquel correspond un ensemble de conformations successives du système.

B.3.a. Principe

L'équation du mouvement, pour la particule *i*, donnée par la seconde loi de Newton ($F = m\,a$), s'écrit :

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

avec $\vec{F}_i$: la force s'exerçant sur la particule *i*, m_i : sa masse et $\vec{a}_i$: son accélération.

A partir de la connaissance de la position de la particule *i* à l'instant t; $\vec{r}_i(t)$, et de la force $\vec{F}_i(t)$ s'exerçant sur cette même particule, on peut déterminer la position et la force à l'instant $t + \Delta t$ en effectuant un développement en série de Taylor grâce à l'algorithme de Verlet (Verlet 1967) :

$$\vec{r}_i(t+\Delta t) = \vec{r}_i(t) + \frac{d\vec{r}_i(t)}{dt}\Delta t + \frac{1}{2}\frac{d^2\vec{r}_i(t)}{dt^2}\Delta t^2 + \frac{1}{6}\frac{d^3\vec{r}_i(t)}{dt^3}\Delta t^3 + o(\Delta t^4)$$

$$= \vec{r}_i(t) + \vec{v}_i(t)\Delta t + \frac{1}{2}\vec{a}_i(t)\Delta t^2 + \frac{1}{6}\frac{d^3\vec{r}_i(t)}{dt^3}\Delta t^3 + o(\Delta t^4) \qquad \text{(Eq.1)}$$

De même :

$$\vec{r}_i(t-\Delta t) = \vec{r}_i(t) - \frac{d\vec{r}_i(t)}{dt}\Delta t + \frac{1}{2}\frac{d^2\vec{r}_i(t)}{dt^2}\Delta t^2 - \frac{1}{6}\frac{d^3\vec{r}_i(t)}{dt^3}\Delta t^3 + o(\Delta t^4)$$

$$= \vec{r}_i(t) - \vec{v}_i(t)\Delta t + \frac{1}{2}\vec{a}_i(t)\Delta t^2 - \frac{1}{6}\frac{d^3\vec{r}_i(t)}{dt^3}\Delta t^3 + o(\Delta t^4) \qquad \text{(Eq.2)}$$

avec respectivement $\vec{v}_i(t)$: vitesse et $\vec{a}_i(t)$: accélération, de la particule *i* au temps *t*.

En combinant Eq.1 et Eq.2, on obtient :

$$\vec{r}_i(t+\Delta t) = 2\vec{r}_i(t) - \vec{r}_i(t-\Delta t) + \vec{a}_i(t)\Delta t^2 + o(\Delta t^4)$$

et

$$\vec{v}_i(t) = \frac{\vec{r}_i(t+\Delta t) - \vec{r}_i(t-\Delta t)}{2\Delta t} + o(\Delta t^3)$$

Pour utiliser l'algorithme de Verlet, il est nécessaire de spécifier en premier lieu les positions initiales $\vec{r}_i(0)$ et les vitesses initiales $\vec{v}_i(0)$ des différents atomes du système.

- Les coordonnées initiales sont obtenues à partir de données expérimentales de biologie structurale (cristallographie des rayons X, RMN) ou bien à partir de données théoriques (simulation de Monte-Carlo, modèle 3D construit par homologie).

- Les vitesses initiales sont attribuées de façon aléatoire selon une distribution de Maxwell.

La position suivante du système peut alors être calculée en appliquant la formule :

$$\vec{r}_i(\Delta t) = \vec{r}_i(0) + \vec{v}_i(0)\Delta t \ .$$

L'algorithme de Verlet peut ensuite être appliqué itérativement en déterminant les accélérations $\vec{a}_i(t)$ à partir des forces :

$$\vec{F}_i = \frac{-dV}{d\vec{r}_i} = m_i\vec{a}_i = m_i\frac{d^2\vec{r}_i}{dt_2}$$

V étant la fonction énergie potentielle décrite précédemment.

Le pas d'intégration Δt doit être choisi avec soin car il doit être suffisamment petit pour que l'énergie potentielle ne varie pas trop rapidement entre 2 pas d'intégration. Le pas d'intégration doit donc être pour cela bien inférieur à la période des mouvements d'oscillation les plus rapides, correspondant à la vibration d'élongation ν(**X-H**) (**X** = **C**, **N**, **O**, **S**, ...) qui sont de l'ordre de 10 femtosecondes (10 fs = 1 / ν <=> ν ~ 3.300 cm^{-1}).

Par contre, plus on prend un pas d'intégration grand, plus le temps de calcul diminue. Ainsi la procédure SHAKE (Ryckaert et al. 1977) permet, en contraignant les longueurs de certaines

liaisons, d'augmenter le pas d'intégration pour gagner du temps de calcul. A l'heure actuelle, il est admis qu'imposer une contrainte avec l'algorithme SHAKE sur les liaisons mettant en jeu un atome d'hydrogène, permet l'utilisation d'un pas d'intégration de 2 fs.

B.3.b. Description d'un protocole

Nous avons travaillé dans l'ensemble microcanonique NVE où N : nombre d'atomes N, V : volume et E : énergie du système, doivent rester constants au cours de la simulation de dynamique moléculaire.

Les 5 phases principales nécessaires pour réaliser ce type d'étude sont :

(i) Préparation du système

La première étape, primordiale, consiste à obtenir un fichier renfermant les coordonnées de l'ensemble des atomes du système. Dans le cas d'une étude en solvant explicite, il est nécessaire de plonger la molécule dans le solvant puis de retirer les molécules de solvant dans l'environnement proche du soluté. Dans le cas de l'eau, toutes les molécules aqueuses dont l'atome d'oxygène est situé à moins de 2,8 Å d'un atome lourd du soluté sont supprimées.

(ii) Minimalisation

Pour une solvatation implicite, le système est simplement minimalisé par l'utilisation successive des deux méthodes SD et GC.

Dans le cas d'une solvatation explicite, le protocole est plus élaboré. Les minimalisations s'effectuent successivement : sur le soluté seul (le solvant étant fixé), puis sur le solvant seul (le soluté étant fixé). Ces 2 étapes sont répétées une duexième fois puis, le système soluté-solvant est minimalisé dans son ensemble, sans aucune contrainte.

La structure finale obtenue correspond à un point minimum de l'hypersurface d'énergie potentielle pour une température théorique de 0 K **(Figure 36)**.

(iii) Chauffage

Cette étape permet l'apport progressif et uniforme d'énergie cinétique (E_{cin}) au système pour atteindre la température de simulation souhaitée qui est le plus souvent 300 K (**Figure 36**). A intervalles de temps réguliers, on attribue aléatoirement selon une distribution de Boltzmann, une vitesse à chacun des *N* atomes du système.

La température est directement reliée aux vitesses par la relation :

$$E_{cin} = \sum_{i=i}^{N} \frac{1}{2} m_i v_i^2 = \frac{3N}{2} k_B T$$

avec v_i : la vitesse de l'atome *i*, m_i : sa masse, k_B : constante de Boltzmann et T : température absolue.

(iv) Equilibrage

Une fois la température désirée atteinte, il est nécessaire d'équilibrer le système, c'est-à-dire stabiliser l'énergie totale et par conséquent la température (**Figure 36**).

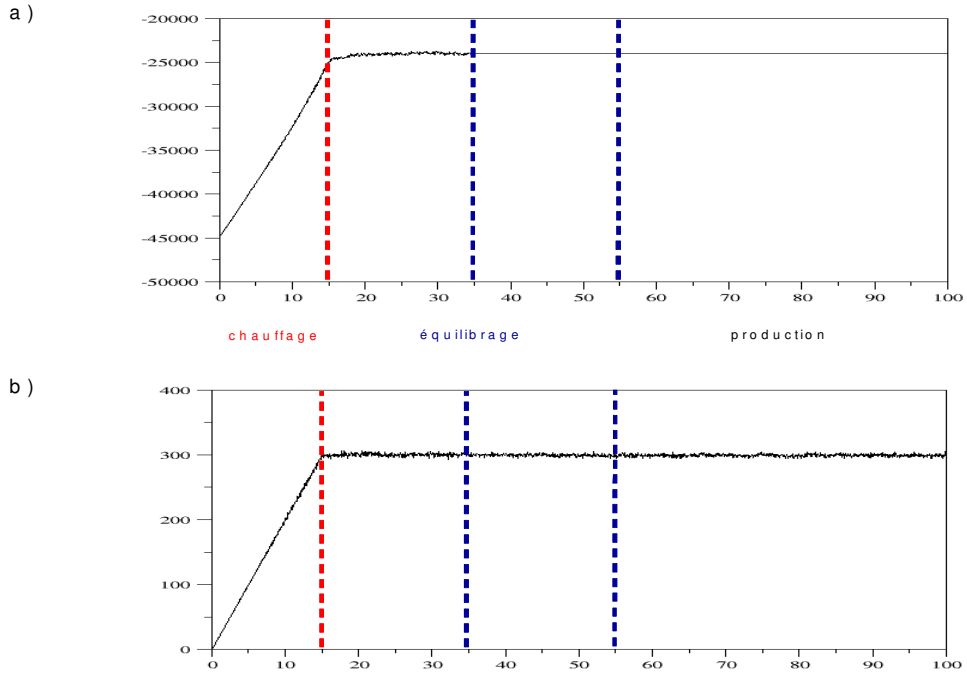
Dans nos simulations, nous avons eu recours à 2 étapes d'équilibrage :

- dans la première, on laisse évoluer le système librement tout en surveillant l'énergie totale que l'on stabilise en jouant sur la température par l'intermédiaire des vitesses comme durant la phase de chauffage;

- dans la seconde étape d'équilibrage, le contrôle de l'énergie totale est également effectué *via* la température, mais avec un contrôle par "réajustement", la température ne devant pas s'écarter de ± 10 K par rapport à la température définie.

(v) Production

Durant la phase de production, le système évolue librement, sans contrainte. C'est cette phase qui donne la trajectoire de simulation de dynamique moléculaire proprement dite et qui va servir à l'analyse du comportement du système. Pour vérifier qu'elle est correcte, il faut au minimum surveiller la stabilité du système par l'intermédiaire de l'énergie totale et de la température (**Figure 36**).



<u>Figure 36</u> : Suivi (a) de l'énergie totale et (b) de la température du système au cours des phases de chauffage, équilibrage, et production d'une dynamique moléculaire.

Si l'une de ces grandeurs ne reste pas constante, c'est-à-dire que sa valeur dérive, cela signifie qu'il y a eu une erreur ou une incompatibilité entre les paramètres de simulation utilisés, par exemple : la phase d'équilibrage a été trop courte ou bien le pas d'intégration n'est pas adapté.

B.3.c. Analyse

A la suite d'une simulation de dynamique moléculaire on analyse la trajectoire obtenue à l'aide de grandeurs et/ou paramètres judicieusement choisis.

α. Séries temporelles

Nous avons utilisé le module CORREL de CHARMM afin de suivre certaines propriétés particulières des molécules en fonction du temps, comme par exemple une distance entre 2 atomes, la valeur d'un angle dièdre…

β. RMSD et carte des RMSD croisés

Pouvoir déterminer quantitativement une différence structurale entre 2 conformations est important pour pouvoir rendre compte de la déformation spatiale qui existe entre elles. Pour cela, on utilise une grandeur scalaire : l'écart quadratique moyen (RMSD : Root Mean Square Deviation) qui représente la moyenne des différences de positions des atomes appartenant à deux conformations différentes *i* et *j*. Un RMSD ayant une valeur de 1, signifie qu'en moyenne, les positions des atomes entre 2 conformations diffèrent de 1 Å.
Cette grandeur, généralement calculée après superposition optimale des deux conformations, est donnée par :

$$RMSD(i,j) = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left\|\vec{r}_{i,n} - \vec{r}_{j,n}\right\|^2}$$

avec N : nombre d'atomes du système et $\vec{r}_{i,n}$: position de l'atome n dans la conformation i.

En général, cette grandeur est calculée tout au long de la simulation en prenant comme structure de référence la conformation initiale, mais on peut tout aussi bien choisir n'importe quelle autre structure, complète ou partielle avec seulement un sous-ensemble de résidus d'intérêt.
L'intérêt du calcul du RMSD prend toute son importance lorsque l'on calcule sa valeur pour chacune des conformations (i fixé) de la simulation de dynamique moléculaire par rapport à toutes les conformations obtenues durant cette même simulation (j variable). (**Figure 37)**.

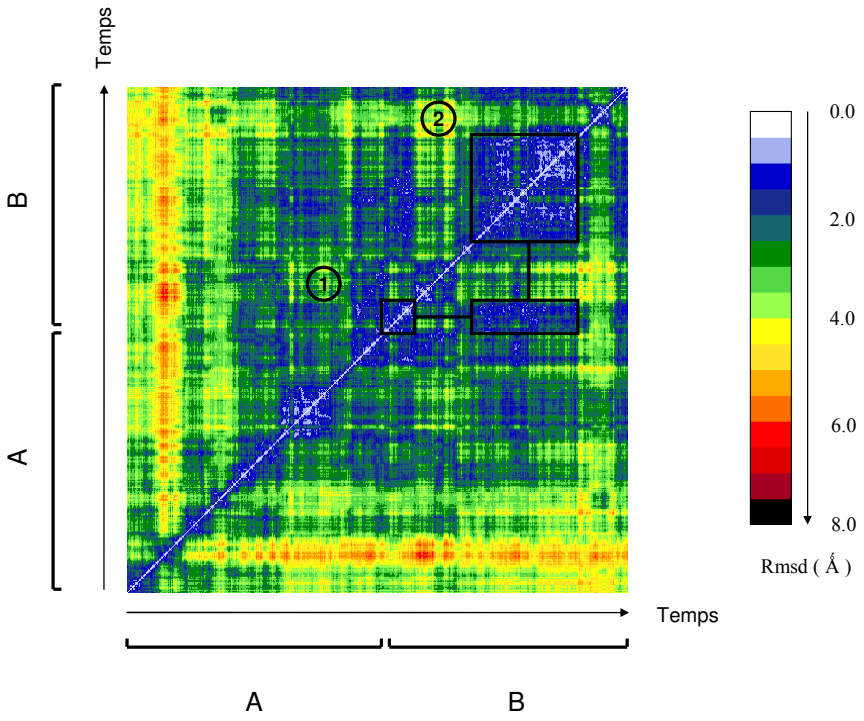


Figure 37 : Carte des RMSD croisés, calculés sur le squelette du peptide Sgal. Deux ensembles de conformations A et B sont discernables comme l'indiquent les zones de couleurs "froides" (bleues-vertes). Les amas de conformations 1 et 2 sont proches comme l'indiquent les "zones bleues" hors diagonale mises en évidence dans les cadres noirs.

On construit alors une carte des RMSD croisés qui est coloriée suivant les valeurs de RMSD obtenues pour chaque couple (i, j) de conformations : une couleur froide signifiant un RMSD faible, tandis qu'une couleur chaude indique un RMSD élevé. La diagonale de la carte correspond aux couples $(i = j)$ et donne donc des RMSD nuls.

L'intérêt de ce type de carte est de bien mettre en évidence la persistance et la stabilité de certains états structuraux de la molécule étudiée, qui ne sont pas nécessairement consécutifs dans le temps. Ainsi les zones de couleurs froides correspondent à des amas de conformations très proches (RMSD faible) et réciproquement, les zones de couleurs chaudes signifient que les conformations sont structurellement éloignées.

Le calcul des RMSD a été effectué avec le module `CORREL` de CHARMM ou bien avec le programme de graphisme moléculaire MOLMOL (Koradi et al. 1996).

L'exemple donné **Figure 37** met en évidence une zone hors diagonale de couleur froide ce qui traduit la proximité des conformations 1 et 2, qui pourtant ne sont pas consécutives dans le temps.

γ. Analyse des structures secondaires

Nous avons utilisé le programme STRIDE (Frishman et Argos 1995) mis au point dans le même but que DSSP (Kabsch et Sander 1983), pour analyser à partir des coordonnées atomiques les structures secondaires des molécules étudiées. En utilisant à la fois la présence de liaisons hydrogène et les angles dièdres du squelette peptidique, le programme détecte les hélices α, 3_{10} et π, les conformations étendues, les ponts isolés, les coudes et enfin tout ce qui ne peut entrer dans l'une des catégories précédentes (boucles ou "coil").

L'utilisation de STRIDE dans le cas de petits peptides n'est pas toujours pertinente si les structures secondaires ordonnées sont rares, et si les seules structures observées sont les coudes β.

δ. Analyse des coudes β

Comme indiqué plus haut en introduction (§ **IB2c**), le coude de type IV est en fait un coude qui ne peut être classé dans aucun des autres types. Il est cependant tout de même intéressant de connaître le type dont il se rapproche le plus. Pour ce faire, une distance euclidienne D entre le coude calculé et un coude canonique est définie dans l'espace des angles dièdres des résidus centraux de la conformation d'intérêt :

$$D = \sqrt{\Delta^2_{\phi(i+1)} + \Delta^2_{\psi(i+1)} + \Delta^2_{\phi(i+2)} + \Delta^2_{\psi(i+2)}}$$

avec

$$\Delta_{\varphi(i+1)} = \phi(i+1)_{observé} - \phi(i+1)_{canonique}$$

Plus D est faible, plus le coude β de type IV est proche du type canonique correspondant.

Cette approche est parfaitement adaptée à notre étude car dans le cas des peptides étudiés, il y a peu, ou même pas du tout de structures secondaires ordonnées et les coudes observables sont souvent peu stables. Les peptides, du fait de leur petite taille, possède une grande flexibilité intrinsèque qui entraîne une réelle difficulté à observer un coude d'un type bien défini. Quand un coude apparaît, il existe bien souvent un équilibre entre un type défini (I, I', II, II', VIa1, VIa2, VIb et VIII) et un type IV proche de ce dernier.

ε. Analyse par zones de la carte de Ramachandran

Même en absence de structures secondaires ordonnées, il semble tout de même intéressant de connaître quelles sont, au cours de la simulation de dynamique moléculaire, les zones de la carte de Ramachandran les plus visitées pour chacun des résidus dont on a pu extraire les valeurs des angles dièdres ϕ et ψ. Nous avons ainsi vérifié si les résidus se

localisaient dans les 10 zones correspondant aux régions définies pour la construction des cartes adiabatiques (Derreumeaux 2001).

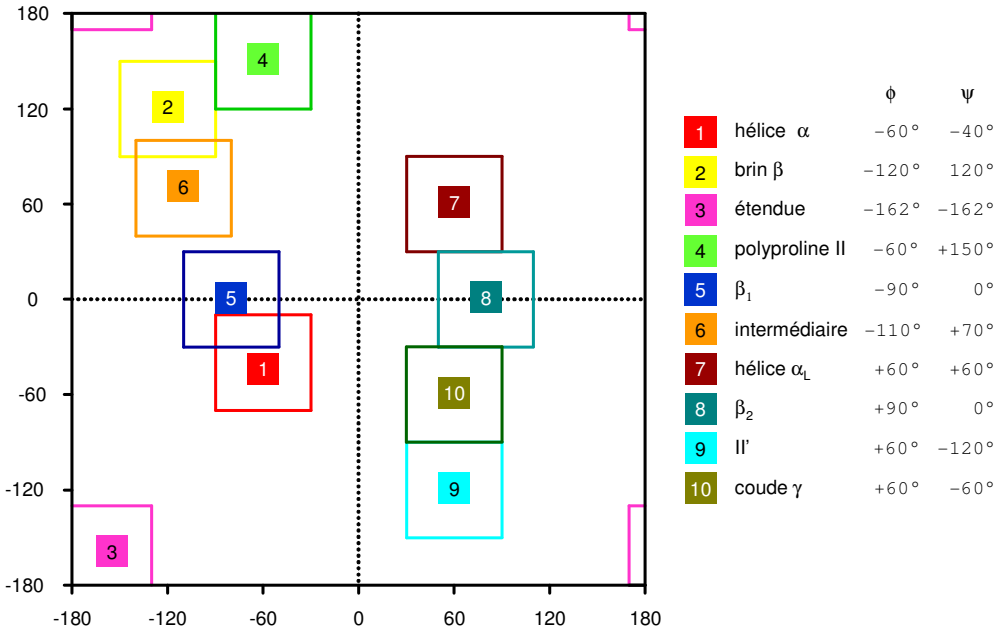


Figure 38 : Représentation des 10 zones, avec leurs valeurs d'angles dièdres ϕ et ψ, de la carte de Ramachandran utilisées pour le calcul d'une carte adiabatique .

La localisation de toutes les zones est donnée ci-dessus **Figure 38**, ainsi que leurs valeurs d'angles dièdres canoniques. Les zones sont délimitées par une tolérance de ± 30° autour des angles ϕ et ψ. Lorsque deux régions se chevauchent comme c'est le cas pour les zones 1 et 5 ou 2 et 6, nous avons défini une distance euclidienne, de la même manière que celle décrite pour l'analyse des coudes β, afin d'affecter le résidu dans l'une ou l'autre des deux zones

ϕ. Analyse par amas de conformations (cluster)

L'analyse de la simulation de dynamique moléculaire peut se faire également par regroupement d'amas de conformations proches. Nous avons, pour cela, utilisé l'algorithme de "clustérisation" ART-2 implémenté dans le programme CHARMM (Karpen et al. 1993). Cet algorithme est basé sur une approche de type réseau de neurones (Carpenter et Grossberg 1987). A partir d'un vecteur de *N* paramètres qui sont, dans notre cas, les angles dièdres ϕ et ψ des résidus il réalise une partition optimalisée du vecteur suivant un rayon de troncature. Le rayon de troncature est choisi de manière à ce que les amas de conformations proposés correspondent au mieux aux transitions structurales de la molécule.

χ. Calcul de variation d'énergie libre ΔG

Au cours de nos trajectoires de dynamique moléculaire, il est possible qu'une molécule oscille entre deux types de conformations. Pour vérifier si nous sommes bien en présence d'un équilibre thermodynamique, il est possible à partir de la formule de Gibbs de calculer la variation d'énergie libre ΔG entre deux types de conformations :

$$\Delta G_{1 \to 2} = -k_b T \ln\left(\frac{P_2}{P_1}\right)$$

avec K_b, la constante de Boltzmann ; T, la température ; P_1 ou P_2 la probabilité d'avoir la molécule dans l'état 1 ou 2.

Dans le cas d'un peptide se repliant en type VIII, la variation d'énergie libre se calculera de la manière suivante :

$$\Delta G_{non\ type\ VIII \to type\ VIII} = -k_b T \ln\left(\frac{P_{type\ VIII}}{P_{non\ type\ VIII}}\right)$$

B.4. Amarrage moléculaire

L'interaction entre une protéine et son substrat est la première étape de la plupart des réactions biologiques. Comprendre son mode de fonctionnement et définir quels sont les résidus mis en jeu, est donc primordial pour pouvoir expliquer les mécanismes qui influent sur l'affinité entre deux molécules. De même, la découverte de nouvelles drogues activant ou inhibant l'activité biologique d'une protéine ne peut se faire qu'en prédisant leur affinité respective. C'est dans ce but que des techniques de modélisations moléculaires, regroupées sous le nom "amarrage moléculaire" ou "docking", ont été développées. Il est à noter que les substrats et les inhibiteurs sont définis par le terme générique de "ligand", tandis que le terme "récepteur" est utilisé pour la cible sur laquelle le ligand va se fixer, que ce soit une enzyme ou un récepteur protéique.

Afin d'obtenir le minimum global d'énergie d'interaction entre le ligand et le récepteur, la plupart des méthodes d'amarrage moléculaire repose donc sur des minimalisations d'énergie, des recuits simulés, des calculs de Monte-Carlo, de complémentarité de surface, d'énergie libre ou encore sur des algorithmes génétiques. Parmi tous les programmes d'amarrage moléculaire qui ont déjà prouvé leur efficacité, comme GOLD (Jones et al. 1995), FlexX

(Rarey et al. 1996) ou bien DOCK (Kuntz et al. 1982), nous avons choisi d'utiliser AutoDock (Automated Docking of flexible ligands to receptors) (Goodsell et Olson 1990).
AutoDock permet d'imposer au système l'exploration de l'espace conformationnel en autorisant certains degrés de liberté au ligand, tout en maintenant fixe le récepteur.
Pour chaque variation de conformation, l'énergie d'interaction résultante est évaluée. Le contrôle des paramètres définissant les propriétés du ligand est réalisé avec un programme indépendant appelé AutoTors; le programme AutoGrid effectue le calcul des différentes grilles et AutoDock réalise les calculs d'amarrage moléculaire proprement dits. Nous avons également utilisé l'interface graphique "ADT" (AutoDock Tools) qui a été développé par l'équipe qui a créé le programme AutoDock. ADT permet, notamment, de visualiser la localisation précise de l'espace à l'intérieur duquel vont se faire les calculs et de créer les scripts nécessaires pour AutoTors, AutoGrid et AutoDock.

La première étape consiste à préparer le ligand et le récepteur en définissant les charges partielles de leurs atomes. Il est possible de le faire directement par l'intermédiaire d'AutoTors, en leur attribuant des charges de type "Kollman" calculées avec le champ de forces empirique Amber (Weiner et Kolmann. 1981). Cependant, pour le ligand, bien qu'il soit possible de calculer les charges partielles des atomes par l'intermédiaire d'AutoTors, en utilisant la méthode de Gasteiger-Marsili (Gasteiger et Marsali 1980), nous avons préféré calculer les charges partielles des atomes grâce au programme MOPAC / AMPAC implémenté dans INSIGHT II, grâce à la méthode semi-empirique AM1 (Dewar et al. 1985). AutoTors permet ensuite de contrôler les mouvements que peut effectuer le ligand. Il permet de définir un "centre de masse" pour les mouvements rigides externes que sont les rotations et les translations et les angles torsions permettant les mouvements de déformations internes du ligand.

Pour augmenter la vitesse d'évaluation de l'énergie du système, le récepteur est plongé dans une grille tridimensionnelle dont le volume est défini par l'intermédiaire du nombre de noeuds ou points n_x, n_y et n_z constituant cette grille (**Figure 39**). Pour chaque type d'atomes (généralement **C**, **H**, **O**, **N**), un atome "sonde" est placé en chacun des nœuds de la grille et l'affinité entre le type d'atome et la molécule cible est calculée par AutoGrid. Une grille d'affinité vis-à-vis de la molécule cible est ainsi obtenue pour chaque type d'atome.

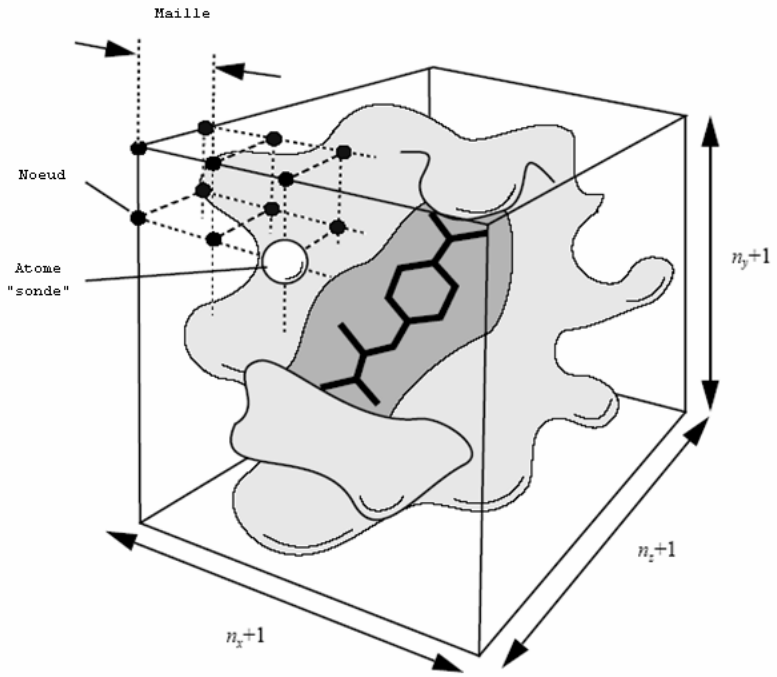


Figure 39 : Schéma d'une grille calculée par AutoGrid.

Parallèlement, une autre grille de potentiel électrostatique est également calculée en résolvant l'équation de Poisson-Boltzmann. Cette procédure permet ainsi d'avoir des temps de calculs 100 fois plus courts que s'il était nécessaire de lister l'ensemble des interactions non liantes entre le ligand et le récepteur pour chacune des modifications de la position ou de la structure du ligand. La maille de la grille est fixée par défaut à 0,375 Å qui correspond au quart de la longueur d'une simple liaison **C-C**, mais elle peut varier de 0,2 à 1,0 Å suivant la précision désirée de la grille.

Suivant la structure du ligand et son emplacement vis-à-vis du récepteur, AutoDock évalue leur affinité en utilisant une fonction de score basée sur une forme empirique de l'énergie libre d'interaction entre le ligand et le récepteur :

$$\Delta G_{\text{interaction}} = \Delta G_{\text{van der Waals}} + \Delta G_{\text{électrostatique}} + \Delta G_{\text{liaison hydrogène}} + \Delta G_{\text{desolvatation}} + \Delta G_{\text{torsion}}$$

Les trois premiers termes traduisent respectivement les termes classiques de mécanique moléculaire d'attraction et de répulsion entre les atomes, d'interaction électrostatique et de liaison hydrogène. Le quatrième terme modélise les effets hydrophobes et la désolvatation après la liaison du ligand au récepteur. Le dernier terme permet de prendre en compte les mouvements internes et les rotations / translations du ligand.

Pour réaliser automatiquement l'échantillonnage des structures, des positions et des orientations du ligand par rapport au récepteur, nous avons utilisé l'algorithme génétique de

Lamarck implanté dans la version 3.0 d'AutoDock (Morris et al. 1998). C'est une méthode hybride basée sur un algorithme génétique couplé à une méthode de recherche locale reposant sur un algorithme optimalisé par Solis et Wets (Solis et Wets 1981). L'avantage de cette méthode est qu'elle permet d'ajouter à l'algorithme génétique, la possibilité de modifier le phénotype suivant la fonction de score qui à son tour modifie le génotype (**Figure 40**).

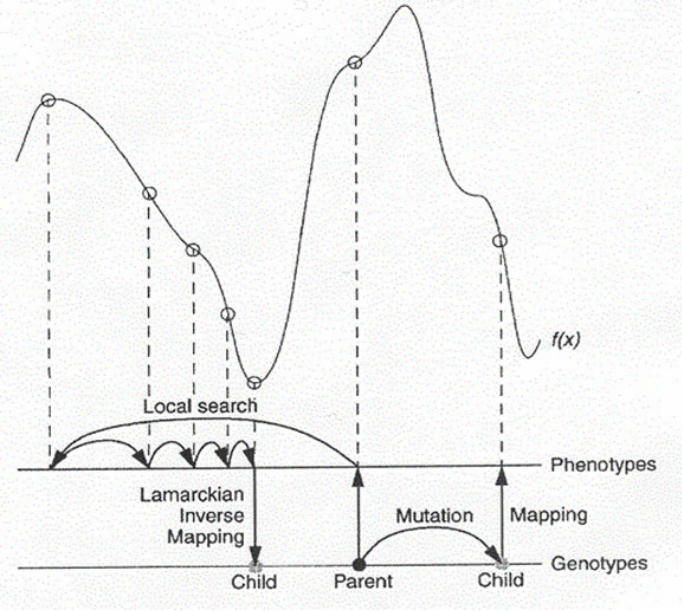


Figure 40 : Principe de l'algorithme génétique de Lamarck.

B.5. Moyens informatiques

B.5.a. Logiciels utilisés

Le programme principalement utilisé durant ce travail a été le logiciel de modélisation moléculaire CHARMM version c27b2 (Brooks et al. 1983; MacKerrel et al. 1998) qui a permis les calculs de modélisation et de simulation de dynamique moléculaire. Nous avons également utilisé le logiciel commercial INSIGHT II (Accelrys Inc), un autre programme de modélisation moléculaire, pour les calculs des charges partielles des ligands pour les calculs d'amarrage moléculaire.

Bien qu'il ait été possible de n'avoir recours qu'au seul logiciel CHARMM, nous avons préféré utiliser des programmes spécialement développés pour certaines techniques de modélisations moléculaires particulières :

- pour le Monte Carlo, le programme FANTOM (Schaumann et al. 1990) et
- pour l'amarrage moléculaire, AutoDock (Goodsell et Olson 1990).

Enfin, pour la visualisation moléculaire, plusieurs logiciels de graphisme moléculaire ont été utilisés suivant leurs spécificités : RasMol (Sayle et Milner-White 1995), MOLMOL (Koradi et al. 1996), VMD (Humphrey et al. 1996) et PyMOL.

B.5.b. Moyens de calcul

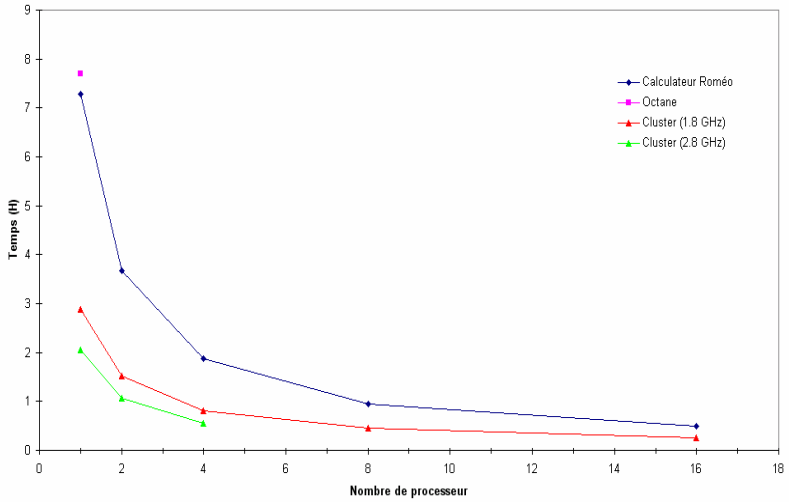
Au début de cette thèse, les moyens de calculs du laboratoire comportaient, au delà des PCs standards et d'un ordinateur haut de gamme DEC α 3000, des stations Silicon Graphics Indigo (monoprocesseur) et Octane DUO (biprocesseurs et/ou 2 postes de travail) pour la visualisation et les calculs de modélisation moléculaire pour lesquels ce type de station n'est pas vraiment destiné, engendrant ainsi des temps de calculs extrêmement longs.

Depuis juin 2002, nous avons pu avoir accès au calculateur "Romeo" de l'Université de Reims Champagne-Ardenne financé par le Conseil Régional. Il s'agit d'un Sunfire 6800 (SMP 24 x UltraSparc III @900MHz, 24 Go de mémoire) permettant d'effectuer des calculs en parallèle allant jusqu'à 4 processeurs.

Finalement, le laboratoire s'est doté en janvier 2003, grâce à un financement par le Contrat de Plan Etat-Région CPER 2002-2007, d'un cluster de 8 PCs haut de gamme biprocesseurs (2 x 1,8 GHz, 1 Go de mémoire) sous Linux, complété depuis par 2 nouveaux PCs biprocesseurs (2 x 2,8 GHz, 1 Go de mémoire).

Un comparatif des performances (benchmark) des divers moyens de calculs de ces machines est donné plus bas **Figure 41.**

L'exemple choisi est le calcul de simulation d'une dynamique moléculaire d'une durée de 10 ns (par pas de 1 fs soit 10 millions de calculs) sur un petit peptide en solvant explicite comprenant 10.304 atomes dont 10.107 atomes de molécules d'eau.



<u>Figure 41</u> : Comparatif de la puissance de calcul des différentes machines utilisées.

Partie III

SITE de LIAISON du RECEPTEUR

A. ETUDE DU RECEPTEUR

Le récepteur de l'élastine est constitué de trois sous-unités, une neuraminidase, une protéine protectrice et l'EBP. C'est cette dernière qui est responsable des propriétés réceptrices vis-à-vis des ligands du récepteur de l'élastine. Pour comprendre la structure responsable de l'interaction entre les ligands et l'EBP, nous avons essayé dans un premier temps de construire un modèle par homologie de l'EBP.

Pour construire un modèle par homologie, il est nécessaire de trouver une protéine de structure connue avec un taux d'homologie de séquence le plus élevé possible avec l'EBP. La recherche d'homologies de séquence parmi les banques de données des protéines dont la structure a été résolue a été effectuée grâce au programme BLASTP (Altschul et al. 1997). Malheureusement, aucune des séquences proposées ne possédait un taux d'homologie suffisant pour permettre de construire un tel modèle. La recherche de séquences homologues au sein de protéines ayant la même fonction, c'est-à-dire la capacité de se fixer à l'élastine, comme la Galectine-3 humaine ou certaines MAGP (Microfibrille Associated GlycoProtein) ou bien de se fixer au β-galactose (β-galactosidase d'E.coli ou β-glucuronidase humaine) n'ont pas montré de meilleurs taux d'homologies de séquence.
Il n'est donc pas actuellement raisonnable, en l'absence de donnée structurale complémentaire, de construire un modèle par homologies de l'EBP humaine

Ne pouvant obtenir de modèle de la structure tridimensionnelle de l'EBP, nous avons, tout d'abord, regardé quelles informations pouvaient être apportées par sa séquence protéique.
Nous avons restreint ensuite notre étude à la partie réceptrice de 15 résidus de l'EBP que l'on sait capable de se lier encore à l'élastine sous forme de peptide, nommé peptide Sgal.
Des informations expérimentales sur la structure du peptide Sgal ont été obtenues par spectroscopies RMN et de dichroïsme circulaire (DC).
Nous avons également construit par homologie le peptide Sgal et étudié *in silico* par Monte-Carlo et dynamique moléculaire les conformations qu'il peut adopter.

Enfin, la séquence de 32 résidus, spécifique de l'EBP par rapport à son variant d'épissage la β-galactosidase, a également été étudiée expérimentalement sous forme de peptide du point de vue de sa structure.

A.1. Prédictions de structures secondaires

Plusieurs programmes de prédiction des structures secondaires ont été utilisés. Un estimateur de leur qualité de prédiction, en terme de pourcentage de résidus bien prédits appelé Q_3, (**Tableau 04**), a été évalué sur le serveur EVA :

http://www.cubic.bioc.columbia.edu/eva/.

Programmes	Q_3 (%)	Références
PsiPred	76,4	**Jones 1999**
Sable2	77,1	**Wagner et al. 2005**
SSpro4	78,1	**Pollastri et al. 2002**
Porter	80,4	**Pollastri et al. 2005**

Tableau 04 : Qualité de prédiction des structures secondaires.

Le programme Porter est celui qui donne les meilleurs résultats avec 80,4 % de bonnes prédictions. Il aurait été également possible d'utiliser le programme Rosetta (Somons et al. 1999) qui est connu comme étant basé sur l'une des meilleures méthodes de prédiction de structures secondaires.

Les 4 programmes de prédiction de structures secondaires donnent des résultats assez homogènes, avec en moyenne :

17,7 % d'hélices α, 24,6 % de brins β et 57,7 % de structures non ordonnées (random coil).

La partie spécifique de l'EBP, contenant donc le site d'interaction de 32 résidus avec l'élastine ou les peptides dérivant de l'élastine, est prédite en structure non ordonnée. Ces résultats de prédictions de structures secondaires vont également permettre la prédiction des épitopes linéaires de l'EBP.

L'ensemble des prédictions de structures secondaires a été reporté **Figure 42**.

```
1- 50   : MPGFLVRILL LLLVLLLLGP TRGLRNATQR MFEIDYSRDS FLKDGQPFRY
PsiPred : CCHHHHHHHH HHHHHHHHHH HHHHCCCCEE EEEEEECCCE EEECCCEEEE
Sable2  : CCHHHHHHHH HHHHHHHCC HHHHHHCCCC EEEEECCCCE EECCCCEEEE
SSpro4  : CCHHHHHHHH HHHHHHHCCC HHHHHCCCCC EEEECHHHHH HHHCCCCEEE
Porter  : CCCHHHHEEE HHCEEEECCC CCEEECCCCC EEEEEECCCE EEECCCEEEE

51-100 :  ISGSIHYSRV PRFYWKDRLL KMKMAGLNAI QTLPGSCGQV VGSPSAQDEA
PsiPred : EEEEEEECCC CHHHHHHHHH HHHHCCCEEE EECCCCCCCC CCCCCCCCCC
Sable2  : EECCCCCCCC CHHHHHHHHH HHHHHHHHHH HCCCCCCCCC CCCCCCCCCC
SSpro4  : EEEECCCCCC CHHHHHHHHH HHHHHHCCCE EECCCCCCCC CCCCCCCCCC
Porter  : EECEECCCCC CHHHHHHHHH HHHHCCCCEE EECCCCCCCC CCCCCCCCCC

101-150 : SPLSEWRASY NSAGSNITDA FLSQRKCEPK GPLINSEFYT GWLDHWGQPH
PsiPred : CCCCCCCEEE CCCCCCCHHH HHHHHHCCCC CCEEEEECCC CCCCCCCCCC
Sable2  : CCCCCCCEEC CCCCCHHHHH HHHHHCCCCC CCEEEEECC CCCCCCCCCC
SSpro4  : CCCCCCCCEE ECCCCCHHHH HHHHCCCCCC CCEEEEECCC CCCCCCCCCC
Porter  : CCCCCEEEEE CCCCCCHHHH HHHHHCCCCC CCEEEEECC CCCCCCCCCC

151-200 : STIKTEAVAS SLYDILARGA SVNLYMFIGG TNFAYWNGAN SPYAAQPTSY
PsiPred : CCCCHHHHHH HHHHHHCCCC EEEEEEECCC CCCCCCCCCC CCCCCEEEEE
Sable2  : CCCCHHHHHH HHHHHHHCC EEEEEEEECC CCEEECCCCC CCCCCCCECC
SSpro4  : CCCCHHHHHH HHHHHHCCCC CEEEEEEECC CCCCCCCCCC CCCCCCCCEC
Porter  : CCCCHHHHHH HHHHHHCCCC CEEEEEEECC CCCCCCCCCC CCCCCCEEEC

201-250 : DYDAPLSEAG DLTEKYFALR NIIQKFEKVP EGPIPPSTPK FAYGKVTLEK
PsiPred : CCCCCCCCCC CCCCCHHHHH HHHHHHCCCC CCCCCCCCCC CCCCCEEEEE
Sable2  : CCCCCCCCCC CCCHHHHHHH HHHHHHCCC CCCCCCCCCC CCCCCCEEEE
SSpro4  : CCCCCCCCCC CCCHHHHHHH HHHHHHCCCC CCCCCCCCCC CCCCCCCCCH
Porter  : CCCCCCCCCC CCCHHHHHHH HHHHHHCCCC CCCCCCCCCC CCCCCCCHHH

251-300 : LKTVGAALDI LCPSGPIKSL YPLTFIQVKQ HYGFVLYRTT LPQDCSNPAP
PsiPred : EECCHHHHHH HCCCCCCCCC CEEEHHHCCC CCCCEEEEEE ECCCCCCCCE
Sable2  : ECCHHHHHHH HCCCCCCCCC CCEEEEECCC CCEEEEEEEE ECCCCCCCCC
SSpro4  : CHHHHHHHHH CCCCCCCCCC CCCCHHHCCC CCEEEEEEEE ECCCCCCCCC
Porter  : HHHHHHHHHH CCCCCCCCCC CCCCHHHHCC CCEEEEEEEE CCCCCCCCCC

301-350 : LSSPLNGVHD RAYVAVDGIP QGVLERNNVI TLNITGKAGA TLDLLVENMG
PsiPred : EEEEECCCCE EEEEEECCEE EEEEEEEEEE EEECCCCCCC EEEEEEEECC
Sable2  : CCCCCCCCCC EEEEEECCCC EEEEECCCEE EEEECCCCCC EEEEEEEEEC
SSpro4  : CCECCCCCCC CEEEEECCEE EEEEECCEEE EEEECCCCCC EEEEEEEECC
Porter  : CCCECCCCCC EEEEEECCCE EEEEECCEEE EEECCCCCCC EEEEEEEECC

351-400 : RVNYGAYIND FKGLVSNLTL SSNILTDWTI FPLDTEDAVR SHLGGWGHRD
PsiPred : CCCCCCCCCC CCCCCCCEEE CCEECCCCEE EECCCCCHHH HHCCCCCCCC
Sable2  : CCCCCCCCCC CCCCCCCEEE CCCCCCCCEE EECCCHHHHH HHCCCCCCCC
SSpro4  : CCCCCCCHCC CCCCCCCEEE CCEECCCCEE EECCCCCCEE EECCCCCCCC
Porter  : CCCCCCCCCC CCCECCCEEE CCCEECCEEE EECCCCCCHH CCCCCCCCCC
```

```
401-450 : SGHHDEAWAH NSSNYTLPAF YMGNFSIPSG IPDLPQDTFI QFPGWTKGQV
PsiPred : CCCCCCCCCC CCCCCCCCEE EEEEEECCCC CCCCCCCEEE EECCCCEEEE
Sable2  : CCCCCCCCCC CCCCCCCCEE EEEEEECCCC CCCCCCCEEE ECCCCCEEEE
SSpro4  : CCCCCCCCCC CCCCCCCCEE EEEEECCCCC CCCCCCCCEE ECCCCCCCEE
Porter  : CCCCCCHHCC CCCCCCCCEE EEEEEECCCC CCCCCCCEEE ECCCCCCEEE

451-500 : WINGFNLGRY WPARGPQLTL FVPQHILMTS APNTITVLEL EWAPCSSDDP
PsiPred : EECCCCCCEE ECCCCCEEEE ECCCHHCCCC CCEEEEEEEE CCCCCCCCCC
Sable2  : EECCCCCCCE CCCCCCCCEE ECCCCCECCC CCCEEEEEEE CCCCCCCCCC
SSpro4  : EEECCCCCCE CCCCCCCEEE EECCHHCCCC CCCEEEEEEE CCCCCCCCCC
Porter  : EEECCCCCEE ECCCCCCEEE EECCCHCCCC CCCEEEEEEE CCCCCCCCCC

501-546 : ELCAVTFVDR PVIGSSVTYD HPSKPVEKRL MPPPPQKNKD SWLDHV
PsiPred : CEEEEEECCC CCCCCCCCEE CCCCCCHHCC CCCCCCCCHH HHHHCC
Sable2  : CCEEEEEECC CEECCCECCC CCCCCCCCCC CCCCCCCCCH HHHHHH
SSpro4  : CCCEEEEEEC CEEECCCCCC CCCCCCCCCC CCCCCCCCCC CHHCCC
Porter  : CCCEEEEECC CECCCCCCCC CCCCCCHHCC CCCCCCCCCC CCCCCC
```

Figure 42 : Prédictions des structures secondaires de l'EBP la séquence spécifique de 32 résidus est représentée en bleu; H : hélice α, E : brin β et C : random coil.

A.2. Recherche d'épitopes linéaires

Un épitope, qui est également appelé antigène, est un site reconnu de manière spécifique par un anticorps. On le dit linéaire (coudé) lorsqu'il est constitué exclusivement d'une séquence de résidus consécutifs, appartenant au même segment protéique. Il existe également des épitopes non linéaires, mais dont la prédiction est extrêmement difficile puisque dépendant du repliement de la protéine. Du point de vue structural, les résidus formant un épitope linéaire (coudé) doivent être exposés à la surface de la protéine, hydrophiles, flexibles et ne pas être en structure hélice α ou brin β, mais en coude.

La méthode PEOPLE (Predictive Estimation Of Protein Linear Epitope), développée au laboratoire (Alix 1999), permet par la combinaison de la prédiction de ces 4 états, de calculer un indice d'antigénicité reflétant les chances que plusieurs résidus consécutifs constituent un épitope linéaire coudé :

- **l'hydrophilie** est calculée à l'aide de 4 méthodes différentes (Hopp et Woods 1981), (Kyte et Doolittle 1982), (Parker et al. 1986) et (Efremov et Alix 1993).

- **l'accessibilité au solvant** est calculée selon l'échelle de Janin (Janin 1979) et d'Emini (Emini et al. 1985) corrigée par Alix (Alix, 1999).

- **la flexibilité des résidus** est calculée selon l'échelle de Karplus et Schulz (Karplus et Schulz 1985) et de Vihinen (Vihinen et al. 1994).

- la prédiction des états structuraux a été initialement effectuée par PEOPLE avec les méthodes de type Chou-Fasman (Chou et Fasman 1978) et GOR (Garnier et al. 1978), mais nous avons également utilisé les résultats des méthodes de prédiction des structures secondaires citées plus haut.

Lorsqu'au moins quatre résidus consécutifs sont considérés comme hydrophiles, accessibles au solvant, flexibles et non prédits en hélice α ou brin β (prédit coudé), alors cette séquence est considérée comme étant un épitope linéaire (coudé).

Les épitopes de l'EBP prédits par la méthode PEOPLE sont reportés **Tableau 05**.
On peut constater, tout d'abord, que la séquence **PSAQDEASPL**, qui appartient à la partie spécifique de l'EBP humaine par rapport à la β-galactosidase humaine, est prédite comme étant un épitope linéaire. Or, cette séquence est bien reconnue de manière spécifique par un anticorps.
La séquence **LMPPPPQKNDS** est reconnue également par un autre anticorps servant pour la mise en évidence de la présence de l'EBP ou de la β-galactosidase humaine (Hinek et Rabinovich 1994).

Séquences	Positions
PSAQDEASPL	**94-102**
EPKJP	128-132
GQPHS	147-151
NGANS	187-191
EKVPEGPIPPSTP	227-239
PQDCSN	292-297
GHRDSGHH	397-404
AHNSS	409-413
GIPDLPQ	430-436
PARGP	462-466
PCSSDDPE	494-501
LMPPPPQKNDS	**530-541**

Tableau 05 : Epitopes de l'EBP prédits par PEOPLE (Alix 1999).

B. ETUDE DU PEPTIDE Sgal

Comme il est impossible d'étudier l'EBP humaine dans son ensemble, nous avons donc restreint notre étude à la séquence responsable de son interaction avec ses ligands qui sous forme peptidique, appelé peptide Sgal, conserve ses propriétés réceptrices vis-à-vis de ses ligands. Pour obtenir des informations structurales sur le peptide Sgal, nous l'avons étudié par spectroscopies RMN et de dichroïsme circulaire. Par modélisation moléculaire, sa structure et sa dynamique ont été étudiées par Monte-Carlo et simulations de dynamique moléculaire.

B.1. Résultats expérimentaux

B.1.a. Spectroscopie de dichroïsme circulaire électronique

Les spectres de dichroïsme circulaire électronique du peptide Sgal ont été enregistrés à différentes températures en solution aqueuse (**Figure 43**) et dans le TFE (**Figure 44**).

α. DCE de Sgal dans l'eau

L'interprétation des spectres de DCE du peptide Sgal dans l'eau est assez complexe, mais plusieurs caractéristiques remarquables peuvent y aider (**Figure 43**).

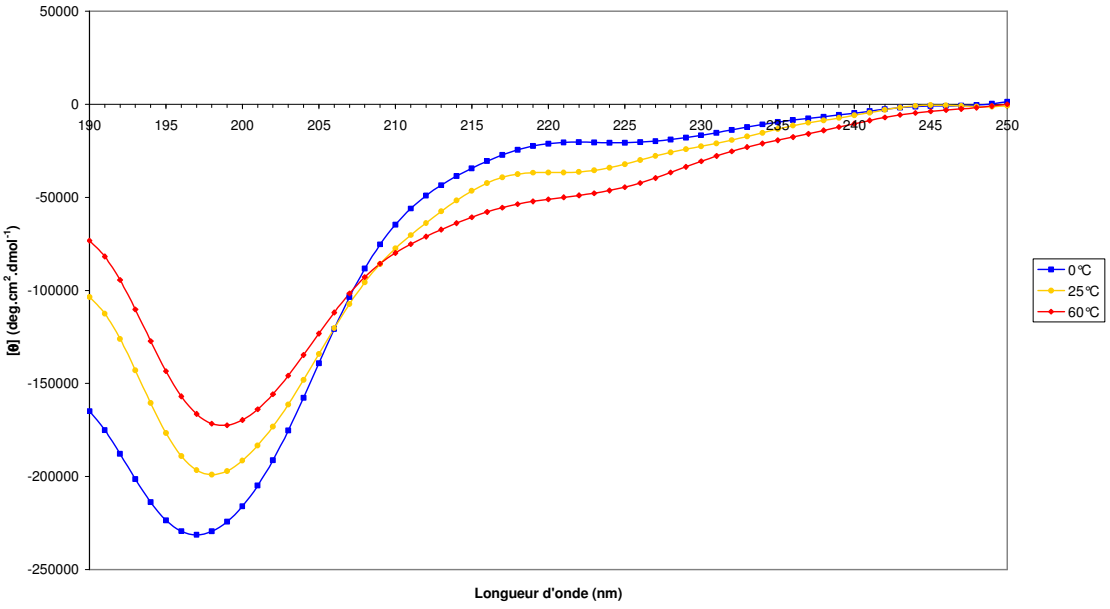


Figure 43 : Spectres de DCE de Sgal dans l'eau.

Tout d'abord, on peut voir que les trois spectres pris à différentes températures ont un point commun à 208 nm, c'est ce que l'on appelle un point isoelliptique. La présence d'un point isoelliptique indique que la molécule est en équilibre entre deux types de conformation.

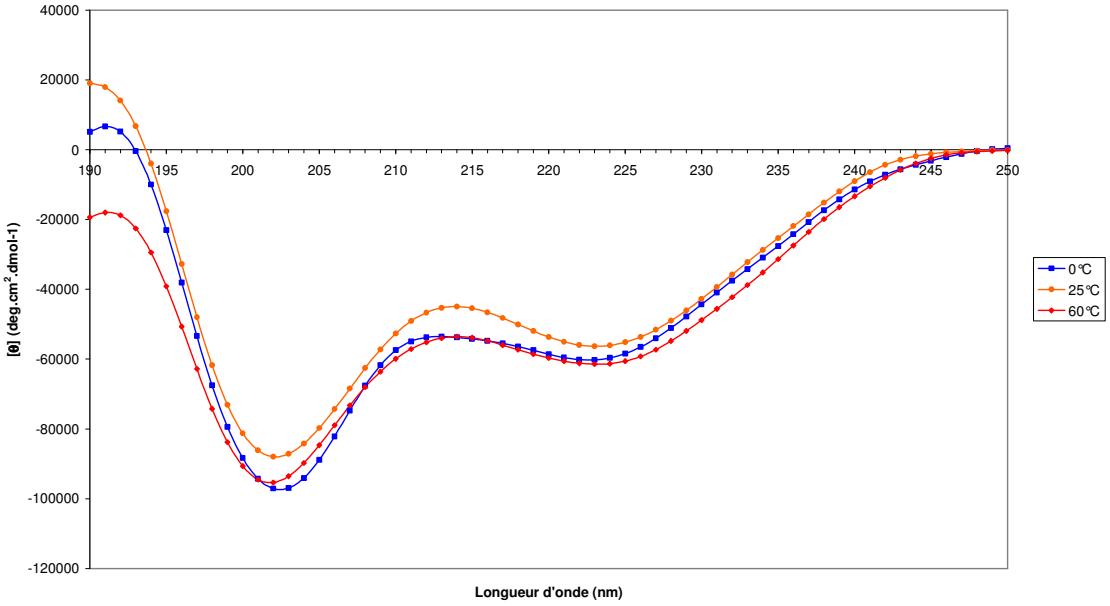
Aucune bande caractérisant la présence d'hélices α, de brins β, de coudes β de type I ou II n'est visible. Mais, une bande négative intense est visible, centrée autour de 198 nm, et qui, avec une augmentation de la température, voit son intensité diminuer. La présence de cette bande est habituellement associée à la présence de structure non régulière (random coil), mais cela ne semble pas être le cas ici, puisque l'augmentation de la température devrait renforcer son intensité.

De plus, on peut voir que la diminution de la température provoque l'apparition d'une bande qui tend à devenir positive autour de 219 nm. Cette modification du spectre en fonction de la température permet de rejeter la possibilité de la présence de coude β de type VIII de manière stable au sein du peptide Sgal.

Le comportement conjoint de ces deux bandes (198 nm et 219 nm) suivant la température est caractéristique de la présence d'une structure **polyproline II**. Mais le fait que la bande autour de 219 nm ne devienne pas positive avec la diminution de la température indique que le peptide Sgal n'adopte pas qu'une structure en polyrpoline II, mais que cette dernière est en équilibre avec une structure non ordonnée, très probablement en **random coil**.

β. DCE de Sgal dans le TFE

A l'inverse, les spectres de dichroïsme circulaire du peptide Sgal dans le TFE (**Figure 44**) sont plus faciles à interpréter que ceux obtenus dans l'eau.



<u>Figure 44</u> : Spectres de DCE de Sgal dans le TFE.

Il est, dans un premier temps, intéressant de noter que la structure du peptide Sgal est remarquablement stable en fonction de la température puisque les spectres pris à différentes températures sont quasiment identiques. En effet, quelle que soit la température, deux bandes négatives sont observables, une à 202 nm et une autre à 223 nm, ainsi qu'une bande positive à 190 nm. La présence de ces trois bandes est caractéristique d'une structure en hélice α ou d'un coude β de type I. La spectroscopie de DCE ne permet pas de conclure en faveur de l'une ou l'autre de ces structures. Cependant, le fait que la première bande négative soit à 202 nm et non pas à 208 nm, laisse penser que nous sommes plutôt en présence du spectre de DCE d'un coude β de type I plutôt qu'en présence de celui d'une hélice α. De plus, des résultats préliminaires de spectroscopie RMN ont clairement montré l'absence d'hélice α.

B.1.b. Spectroscopies de résonance magnétique nucléaire

Les spectres TOCSY, NOESY et ROESY du peptide Sgal ont été enregistrés dans un mélange 90 % d'H_2O / 10 % de D_2O à une température de 298 K.

Les valeurs des déplacements chimiques des protons sont reportées **Tableau 06**.

résidus	déplacements chimiques (ppm)				$^3J_{NH-H\alpha}$ (Hz)
	NH	Hα	Hβ	autres	
V_1		3,86	2,21	1,01	
V_2	8,57	4,15	2,05	0,98	7,0
G_3	8,59	3,98			11,6
S_4	8,20	4,80	3,85		
P_5	-	4,49	2,33	2,02 3,74 / 3,86	
S_6	8,32	4,42	3,89		7,0
A_7	8,31	4,32	1,42		
Q_8	8,28	4,29	1,98 / 2,11	2,36	
D_9	8,35	4,69	2,94 / 2,85		7,0
E_{10}	8,23	4,37	1,98 / 2,13	2,47	7,3
A_{11}	8,25	4,35	1,39		
S_{12}	8,26	4,78	3,85		6,6
P_{13}	-	4,49	2,33	2,02 3,74 / 3,86	
L_{14}	8,30	4,38	1,66	1,66 0,95 / 0,89	
S_{15}	8,09	4,42	3,94 / 3,89		7,6

Tableau 06 : Déplacements chimiques des protons de Sgal dans l'eau (90 % d'H_2O / 10 % de D_2O) à 298 K.

Seuls des effets NOE intra-résidus ou séquentiels des protons du squelette peptidique sont observables, ce qui laisse penser que nous sommes en présence d'une structure non ordonnée. Cependant, le fait que l'intensité des effets δα NOE séquentiels soit plus forte que celles des effets δα NOE intra-résidus, laisse supposer que le peptide Sgal présente un certain taux d'hélice polyproline II.

De plus, les coefficients de température du proton des groupements amides du squelette protéique sont supérieurs à 5,3 ppm.K^{-1}, indiquant ainsi que ces protons ne sont pas impliqués dans une liaison hydrogène, mais qu'ils sont exposés au solvant. Nous sommes donc en présence d'une structure plutôt étendue, compatible avec une hélice de type polyproline II.

Enfin, l'analyse des constantes de couplage $^{3}J_{NH-H\alpha}$ permet d'obtenir des informations concernant l'angle de torsion ϕ grâce à l'équation de Karplus (Karplus 1959). Pour une même constante de couplage, plusieurs solutions sont possibles. Ainsi les résidus V_2 et D_9 ont une valeur de 7.3 Hz ce qui équivaut à un angle ϕ de - 82,2° ou bien à un angle ϕ de - 157,8 °. La première valeur correspond à celle de l'angle dièdre d'un résidu en hélice polyproline II et la deuxième à un résidu en structure étendue.

Seule la glycine **G_3** du fait de la nature particulière de sa chaîne latérale, constituée d'un seul atome d'hydrogène, donne une valeur de constante de couplage ne permettant pas d'obtenir, selon l'équation de Karplus, une solution de valeur d'angle dièdre ϕ.

Les résidus (**V_2, S_6, D_9, E_{10}, S_{12}** et **S_{15}**), dont les constantes de couplage ont pu être attribuées, ont soit une valeur d'angle dièdre ϕ correspondant à celle d'une structure en polyproline II, soit une valeur ϕ correspondant à celle d'une structure étendue.

En conclusion, les données RMN sont en accord avec l'analyse des spectres de DCE dans l'eau : le peptide Sgal semble pouvoir adopter une structure en hélice polyproline II et / ou une structure non ordonnée qui pourrait être une structure plutôt de type étendue. La modélisation moléculaire trouve dans l'étude de ce peptide tout son intérêt puisqu'elle va permettre d'apporter d'autres informations structurales que celles, dans notre cas, limitées, provenant des résultats expérimentaux.

B.2. Résultats théoriques

B.2.a. Construction d'un modèle par homologie du peptide Sgal

Il a été démontré précédemment que le peptide Sgal correspondait à la séquence de l'EBP interagissant avec l'élastine (Hinek et al. 1993; Hinek et Rabinovich 1994). Un anticorps dirigé contre peptide Sgal peut également se fixer sur l'élastase pancréatique porcine

(Pancreatic Porcine Elastase : PPE) de manière à inhiber son activité élastolytique. De plus, une homologie de séquence a été soulignée entre le peptide Sgal et la partie N-terminale de certaines élastases, laissant à penser qu'il s'agirait d'une séquence caractéristique d'un domaine de liaison à l'élastine.

Nous avons reporté **Figure 45** les homologies de séquences proposées initialement dans les articles précédemment cités, complétées par les résultats obtenus après une recherche BLAST dans la banque de données protéiques Swiss-Prot.

Au vu des homologies de séquences et de structures, puisqu'un anticorps reconnaît à la fois le peptide Sgal et la PPE, il semble donc que le site de fixation de l'élastine chez ces protéines soit un épitope linéaire de structure similaire. Une construction par homologie de la structure du peptide Sgal est donc tout à fait réalisable.

```
EBP Humaine, peptide SGAL :                    VVGSP SAQDE ASPLS
EBP de Mouton :                                VVGGT EAQRN SWPLQ
Elastase Pancréatique Porcine (1B0E) :         VVGGT EAQRN SWPSQ
Elastase Leucocytaire Humaine (1HNE) :         IVGGR RARPH ANPFM
Elastase Pancréatique de Saumon (1ELT) :       VVGGR VAQPN SWPWQ
Elastase Pancréatique de Rat :                 VVGGA EARRN SWPSQ
Elastase Pancréatique de Souris :              VVGGQ EATPN TWPWQ
Elastase Pancréatique Humaine :                VVGGE EARPN SWPWQ
Elastase Dermique Humaine :                    VVGGT EAGRN SWPSQ
```

Figure 45 : Homologies de séquences entre les EBP humaine et de mouton et des élastases (les codes PDB correspondent aux structures résolues).

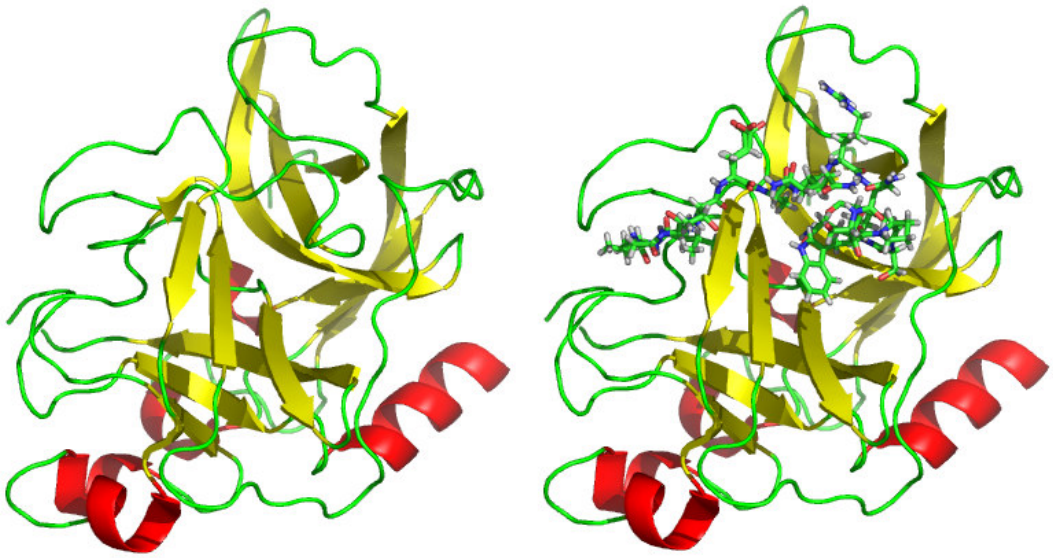
Nous avons choisi d'utiliser comme structure modèle la PPE car, du point de vue séquence, elle est extrêmement proche de celle de l'EBP de mouton avec un taux d'identité de 93,33 %, mais également très proche de l'EBP humaine avec un taux d'identité de 46,15 % (**Figure 46**).

```
EBP Humaine :                          VVGSP SAQDE ASPLS
                                       :::.   ::  .  .  :
Elastase Pancréatique Porcine :        VVGGT EAQRN SWPSQ
```

Figure 46 : Homologie entre l'EBP humaine et l'Elastase Pancréatique de Porc (PPE).

Du point de vue de la structure, on la sait très proche de celle de la PPE puisque un anticorps dirigé contre le peptide Sgal peut également se fixer à la PPE. La structure de

l'Elastase Pancréatique de Porc résolue à 1,80 Å que nous avons utilisée (**Figure 47**), est référencée sous le code PDB : 1B0E (Cregge et al. 1998).



(a) ***(b)***

Figure 47 : Structure 3D de l'Elastase Pancréatique de Porc (1BOE) en représentation *ruban (a)* et partie servant de modèle en représentation *bâtons (b)*.

De manière intéressante, en étudiant sa structure, on peut s'apercevoir que les résidus du site ayant de fortes homologies de séquence avec le peptide Sgal, sont bien exposés en surface de la protéine, renforçant l'hypothèse qu'il s'agit d'un site d'interaction. De plus, ce site est relativement éloigné du site responsable de l'hydrolyse de l'élastine, ce qui permet d'exclure la possibilité que la disparition de l'activité élastolytique de l'élastase PPE en présence d'anticorps anti-Sgal, puisse être due à l'encombrement du site d'élastolyse.

Enfin, les structures, correspondant au site probable de fixation de l'élastine, des différentes élastases sont très proches les unes des autres avec un RMSD inférieur à 0,3 Å, laissant penser à l'existence d'une structure commune responsable de la fixation de l'élastine.

Le modèle par homologie du peptide Sgal a été construit à partir de la structure correspondant à la séquence homologue de l'Elastase Pancréatique de Porc grâce au module `Homology` du programme INSIGHT98. Les atomes d'hydrogène ont été rajoutés avec le module `HBUILD` (Brünger et Karplus 1988) du programme CHARMM. Ce modèle par homologie du peptide Sgal est ainsi formé (**Figure 49**) d'une hélice de polyproline II (**SPS**) et de 4 coudes (**VVGS, QDEA, EASP, SPLS**).

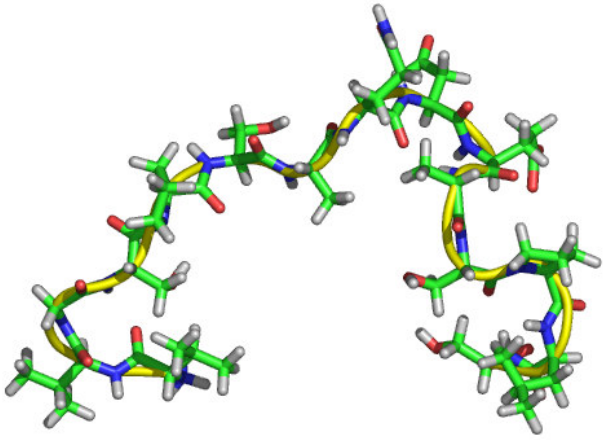


Figure 48 : Structure du peptide Sgal construit par homologie.

Séquence :	**V**	**V**	**G**	**S**	**P**	**S**	**A**	**Q**	**D**	**E**	**A**	**S**	**P**	**L**	**S**
coude β de type II	V	V	G	S	–	–	–	–	–	–	–	–	–	–	–
polyproline II	–	–	–	S	P	S	–	–	–	–	–	–	–	–	–
coude β de type II	–	–	–	–	–	–	–	Q	D	E	A	–	–	–	–
coude β de type IV (VIII)	–	–	–	–	–	–	–	–	–	E	A	S	P	–	–
coude β de type I	–	–	–	–	–	–	–	–	–	–	–	S	P	L	S

Figure 49 : Structures secondaires du modèle par homologie du peptide Sgal.

Une autre séquence présentant une homologie de séquence de 77,8 % sur 9 acides aminés a été trouvée dans les banques de données de structure de protéines (**Figure 50**).

```
Peptide Sgal :        VVGSPSAQDEASPLS
                          :: ::: ::
Intégrine αVβ3 :      EDYRPSQQDECSPRE
```

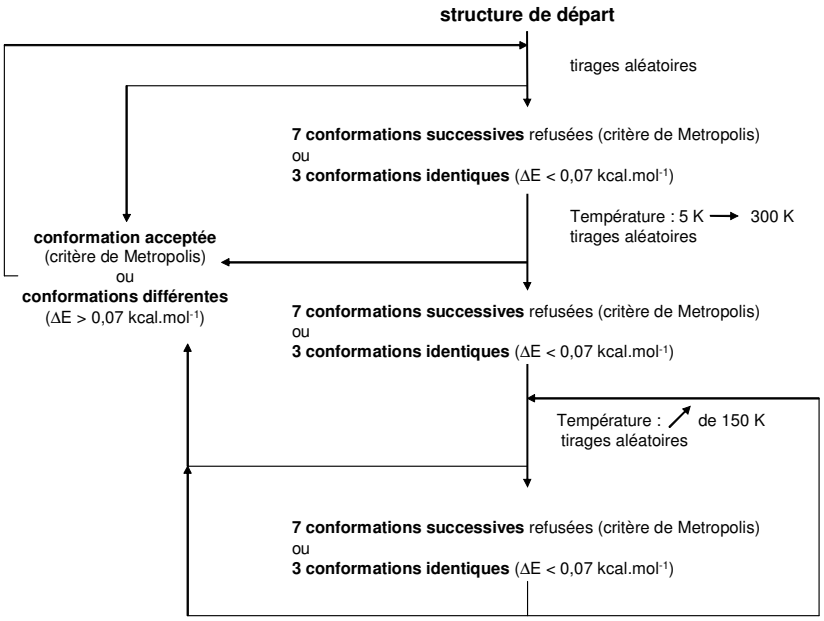
Figure 50 : Homologies de séquences entre le peptide Sgal et l'intégrine $\alpha_V\beta_3$.

Cette séquence homologue à celle du peptide Sgal se situe dans la sous-unité β_3 de l'intégrine $\alpha_V\beta_3$ dans le sous-domaine particulier appelé fibronectine II. L'intégrine est également un récepteur membranaire impliqué dans de nombreux processus biologiques au sein de la MEC. Il a été démontré que cette intégrine était capable de se lier à la tropoélastine mais non pas à travers un site **GXXPG** comme l'EBP, mais par l'extrémité C-terminale de la tropoélastine (Rodgers et Weiss 2004). De plus, la séquence homologue du peptide Sgal est située dans une partie non résolue de la protéine.

B.2.b. Simulations de Monte-Carlo du peptide Sgal

La structure obtenue par homologie a servi de modèle de départ pour les simulations de Monte-Carlo, ses extrémités étant chargées (NH_3^+, COO^-) et l'aspartate (**D_9**) ainsi que le glutamate (**E_{10}**) sont sous forme carboxylate (COO^-). Les calculs ont été effectués en utilisant la version 4.0 du programme FANTOM (Schaumann et al. 1990).

Tous les cycles de Monte-Carlo ont été réalisés avec un cut-off de 30 Å, et suivant le même protocole d'adaptation de la température (**Figure 51**) : lorsque 7 conformations successives n'ont pas été acceptées suivant le critère de Metropolis ou lorsque 3 conformations successives sont considérées comme identiques, c'est-à-dire que leur énergie diffèrent de moins de 0,07 kcal.mol^{-1}, alors la température des simulations de Monte-Carlo suivantes passe de 5 K à 300 K, puis augmente par pas de 150 K jusqu'à l'acceptation de la conformation par le critère de Metropolis ou l'obtention de structures différentes. Un tel traitement de la température permet une large exploration de l'espace conformationnel du peptide Sgal et évite d'être piégé dans un minimum local d'énergie.



<u>Figure 51</u> : Protocole d'adaptation de la température en simulation Monte-Carlo.

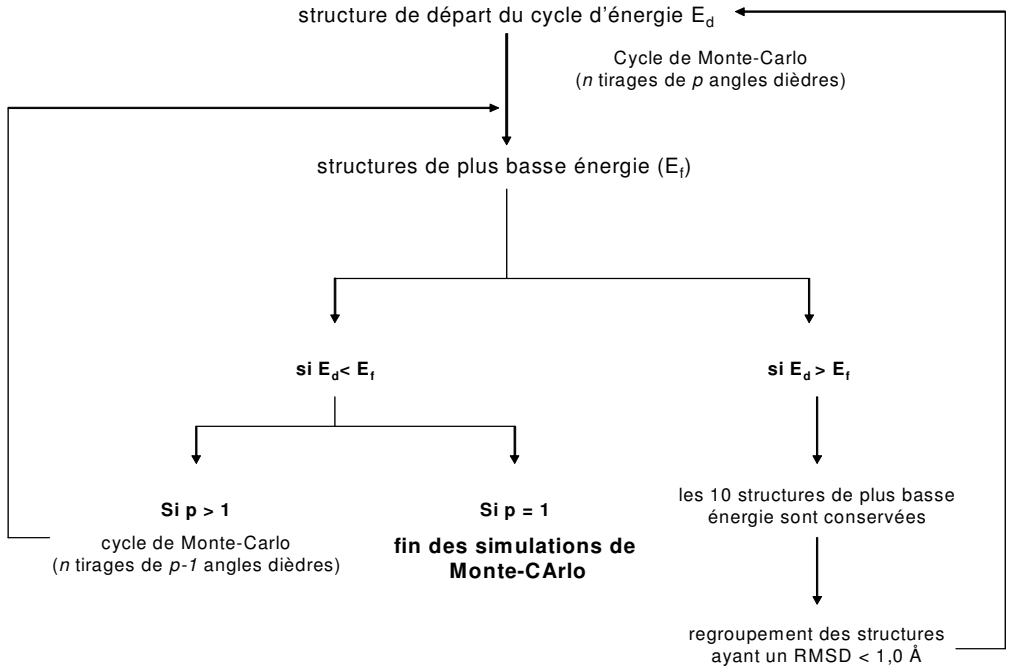
Chaque cycle de Monte-Carlo effectue *n* tirages aléatoires de modifications de *p* angles de torsion suivis de 90 pas de minimalisation par la méthode ABNR (**Figure 52**). Le nombre *p* d'angles torsions peut prendre les valeurs successives 7, 3 et 1.

Les 10 conformations de plus basse énergie obtenues après chaque cycle E_f sont comparées à l'énergie de la structure de départ E_d du cycle correspondant :

- si : $\mathbf{E_f > E_d}$, la structure obtenue est énergétiquement moins favorable que la structure de départ, on reprend alors la structure de départ qui va servir à un nouveau cycle, mais pour lequel le nombre de modifications d'angle de torsion *p* diminue à chaque tirage. Si *p* était déjà égal à 1, alors l'exploration conformationnelle s'arrête.

- si : $\mathbf{E_f < E_d}$, on compare le RMSD des structures : chaque conformation dont le RMSD est supérieur à 1,0 Å par rapport aux autres conformations va servir de structure de départ pour un autre cycle de Monte-Carlo.

Les solvants eau et cyclohexane ont été utilisés (Efremov et al. 1999), pour comparer le comportement du peptide en milieu polaire et apolaire. L'utilisation d'un solvant apolaire permet de mimer un environnement protéique comme par exemple au voisinage de l'élastine.

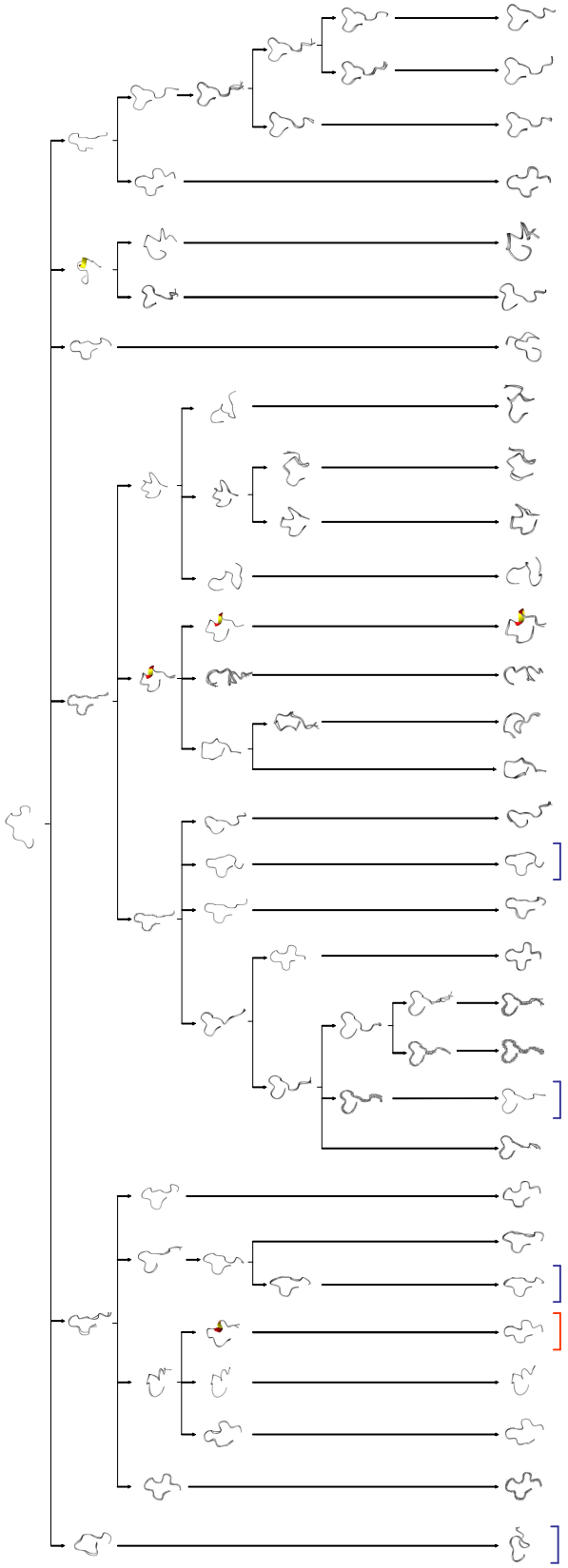


<u>Figure 52</u> : Protocole de choix des structures de départ des différents cycles Monte Carlo

α. Simulation MC dans l'eau

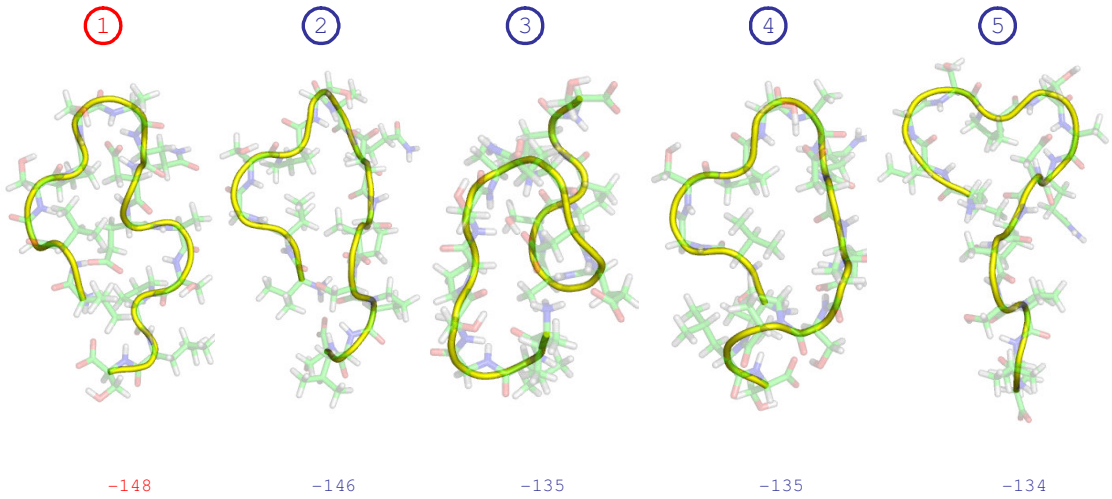
Les calculs de Monte-Carlo dans l'eau ont nécessité 145 cycles de 5.000 tirages pour un total d'environ 73.000 structures retenues selon les critères précédemment décrits. La **Figure 53** présente toutes les structures d'énergie minimum obtenues.

Par souci de clarté, les nombreuses structures intermédiaires obtenues après chacun des cycles de Monte-Carlo n'ont pas été reportées dans la **Figure 53**.



<u>Figure 53</u> : Structures résultant des calculs de Monte Carlo dans l'eau,
en rouge : structure de plus basse énergie, en bleu les 4 suivantes de plus basse énergie.

Nous avons, en fait, obtenu 31 minima d'énergie allant de - 121 à - 148 kcal.mol^{-1} pour le minimum global d'énergie. Certains de ces minima d'énergie correspondant à des structures très proches les unes des autres, nous avons donc regroupé les structures obtenues possédant un RMSD inférieur à 1 Å, la conformation représentative du groupe étant celle ayant la plus basse énergie. Pour l'analyse de ces résultats de calculs de Monte-Carlo, nous avons retenu les cinq conformations représentatives de plus basse énergie (Cf **Figure 54)** : le minimum global d'énergie est donné en rouge et les 4 autres de plus basse énergie en bleu.



<u>Figure 54</u> : Les 5 structures Monte-Carlo de plus basses énergies (en kcal.mol^{-1}).

Aucune de ces structures n'est repliée en structure secondaire de type hélice ou brin.

En revanche, elles ont toutes comme point commun d'avoir une conformation assez compacte stabilisée par jusqu'à 4 liaisons hydrogène, avec les 4 mêmes coudes β :

- **V_2GSP_5** de type VIII ou IV proche du type VIII
- **P_5SAQ_8** de type I ou IV proche du type I
- **S_6AQD_9** de type IV sans être proche d'un type bien déterminé
- **$E_{10}ASP_{13}$** de type VIII ou IV proche du type VIII.

De ce fait, leur partie N-terminale présente un enchaînement de 3 coudes β imbriqués les uns dans les autres, sur le motif **$V_2GSPSAQD_9$** (**Figure 55**), associé à la présence du coude β sur **$E_{10}ASP_{13}$**. Il est à noter que ce type de structure permet notamment au résidu **Q_8** d'être très exposé au solvant.

La conformation 3 est la seule à posséder un coude β supplémentaire de type IV sur **V_1VGS_4** présentant ainsi une partie N-terminale plus compacte que les autres conformations, ainsi qu'un coude β de type I sur **D_9EAS_{12}**.

Enfin, de manière intéressante, la conformation 1, qui est la structure de plus basse énergie, se distingue des autres par la présence d'un coude β de type IV, proche du type VIII, sur la séquence **Q_8DEA_{11}**.

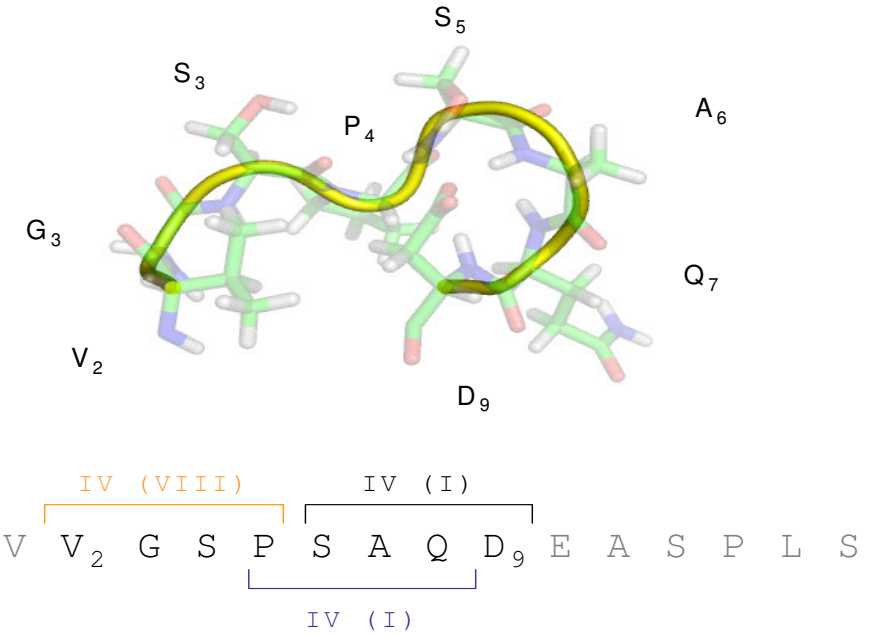


Figure 55 : Détails de la partie N-terminale des conformations 1, 2, 4 et 5.

Conformation	Types de coudes β						
	V_1VGS_4	V_2GSP_5	P_5SAQ_8	S_6AQD_9	Q_8DEA_{11}	D_9EAS_{12}	$E_{10}ASP_{13}$
1		IV (VIII)	I	IV	IV (I)		**IV (VIII)**
2		VIII	IV (I)	IV			**IV (VIII)**
3	IV	IV (VIII)	IV (I)	IV		I	**IV (VIII)**
4		VIII	IV (I)	IV			**IV (VIII)**
5		IV (VIII)	IV (I)	IV			**VIII**

Tableau 07 : Présence des coudes β dans les différentes conformations obtenues après simulations de Monte-Carlo dans l'eau.

β. Simulation MC dans le cyclohexane

53 cycles de 5000 tirages ont été nécessaires pour les calculs de Monte-Carlo dans le cyclohexane et 27.000 conformations ont été retenues selon le critère de Métropolis. Les structures d'énergie minimum, sans les conformations intermédiaires obtenues après chaque cycle, ont été reportées en **Figure 56**.

Nous avons obtenu 8 minima d'énergie, la plus élevée étant - 105 kcal.mol^{-1}, tandis que le minimum global se situe à - 118 kcal.mol^{-1} (**Figure 57**). Les 8 structures correspondantes ont toutes entre elles un RMSD supérieur à 1 Å.

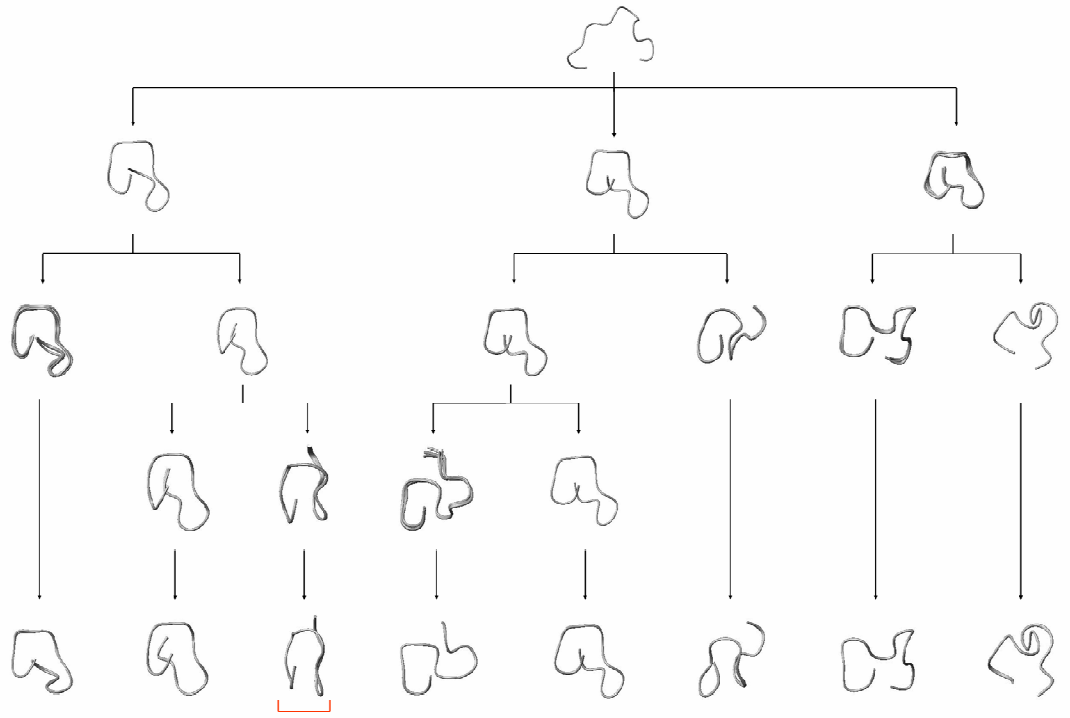


Figure 56 : Structures obtenues par Monte Carlo dans le cyclohexane.
La structure de la plus basse énergie est soulignée en rouge.

Les 8 conformations de plus basses énergies ne présentent aucune structure secondaire régulière, hormis un nombre variable de coudes β (**Figure 57**). Ces structures sont stabilisées par jusqu'à 4 liaisons hydrogène et ont toutes en commun 2 coudes β :

- **V_1VGS_4** : de type IV proche du I' pour les conformations 1, 5 et 8 ou de type II, ou IV proche du type II, pour les conformations 2, 3, 4, 6 et 7

- **D_9EAS_{12}** : de type I ou IV proche du type I.

D'autres coudes β sont retrouvés dans la majorité des conformations :

- de type VIII, ou IV proche du type VIII, sur la séquence **$E_{10}ASP_{13}$** dans toutes les conformations exceptées dans la 8 : celle de moins basse énergie. Ce même type de coude β a été retrouvé dans les conformations de plus basse énergie, obtenues par les calculs de Monte-Carlo dans l'eau et est également présent dans le modèle construit par homologie.

- sur **S_6AQD_9** un coude β, de type différent, est présent sur 5 des 8 conformations.

La présence de certains coudes β sont ainsi caractéristiques de certaines conformations comme ceux de type IV et I sur, respectivement, les séquences **P_5SAQ_8** et **Q_8DEA_{11}** de la

conformation 3. Le coude β de type IV sur **S_4PSA_7** est, de même, retrouvé exclusivement dans la conformation 6.

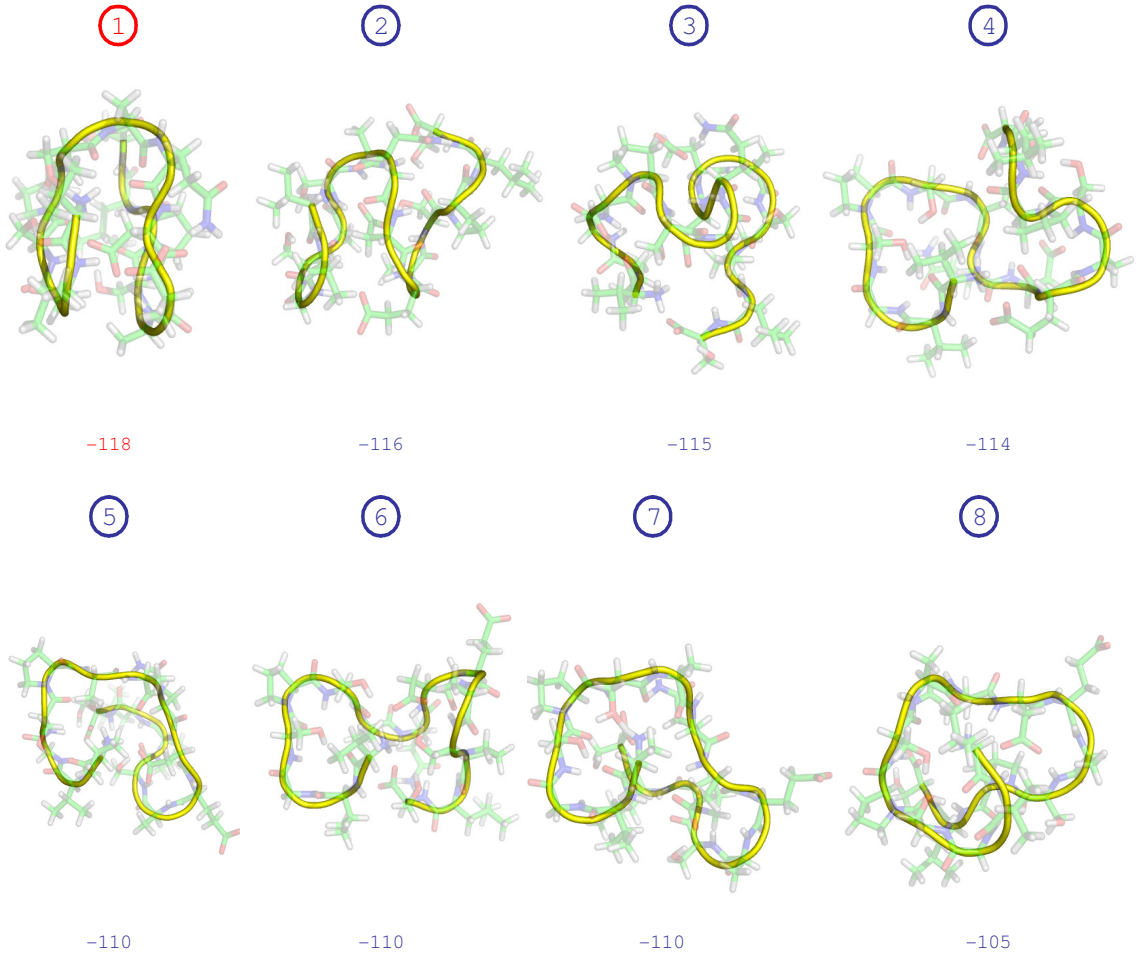


Figure 57 : Structures obtenues par Monte Carlo dans l'eau,
en rouge la structure de plus basse énergie et en bleu celles ayant les plus basses énergies.

Conformation	Types de coudes β						
	V_1VGS_4	S_4PSA_7	P_5SAQ_8	S_6AQD_9	Q_8DEA_{11}	D_9EAS_{12}	$E_{10}ASP_{13}$
1	IV (I')			IV (II')		IV (I)	IV (VIII)
2	II			IV		I	IV (VIII)
3	II		IV	IV	I	I	IV (VIII)
4	IV (II)					IV (I)	IV (VIII)
5	IV (I')					I	VIII
6	II	IV		II		I	IV (VIII)
7	II			IV		I	IV (VIII)
8	IV (I')					IV (I)	

Tableau 08 : Présence de coudes β dans les différentes conformations obtenues après simulations de Monte-Carlo dans le cyclohexane.

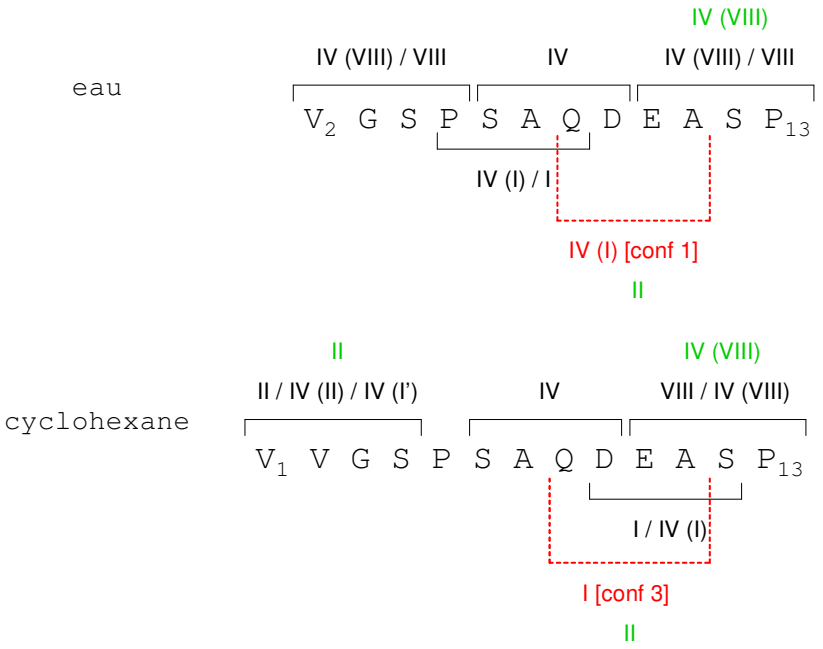
χ Conclusions

Les calculs de Monte-Carlo réalisés ont pour résultats de proposer des structures de plus basses énergies, plutôt compactes, caractérisées par la présence de plusieurs coudes β.

Dans l'eau, ces structures de plus basses énergies ont notamment en commun 4 coudes β et forment ainsi un motif structural particulier englobant 12 résidus (**V_2GSPSAQDEASP$_{13}$**) sur 15. De plus, la structure représentative du minimum global, présente également un coude β sur **Q_8DEA$_{11}$**, renforçant la stabilité de ce motif.

Dans le cyclohexane, les structures de plus basses énergies sont caractérisées par un coude β sur **V_1VGS$_4$**, mais aussi par 3 coudes imbriqués sur la séquence **S_6AQDEASP$_{13}$**.

L'intérêt de cette étude dans l'eau et dans le cyclohexane était de proposer une structure pouvant être responsable de la fixation du peptide Sgal à la tropoélastine.



<u>Figure 58</u> : Types de coudes β dans les structures de plus basse énergie obtenues par Monte-Carlo dans l'eau et le cyclohexane. En vert : coudes β présents dans le modèle construit par homologie et en rouge : ceux présents dans une seule conformation (numéro de celle-ci précisé entre crochets).

Les structures obtenues par les calculs de Monte-Carlo dans ces deux solvants sont schématisées **Figure 58**. Il semble donc que quelle que soit la nature du solvant, le domaine **S_6AQDEASP$_{13}$** est structuré par plusieurs coudes β qui permettent de bien exposer les résidus **A_7**, **Q_8** et **D_9** au solvant.

Toutes ces conformations, possédant les plus basses énergies d'après les calculs de Monte-Carlo, présentent des structures compactes qui semblent être peu compatibles avec les résultats expérimentaux proposant une conformation en polyproline II et / ou en random coil.
La comparaison entre nos résultats de Monte-Carlo dans le cyclohexane et les résultats expérimentaux est difficile. En effet, expérimentalement, les peptides précipitent dans une solution de cyclohexane (on utilise même cette propriété durant les étapes de purification de la synthèse peptidique chimique). Cependant, puisque le peptide Sgal peut se fixer à l'élastine qui est une protéine extrêmement apolaire, l'intérêt d'une étude de Monte-Carlo dans le cyclohexane est de connaître le comportement structural théorique au sein d'un environnement très apolaire. C'est pour cela qu'expérimentalement, le TFE est utilisé pour mimer un environnement plus apolaire que l'eau, mais comparer des résultats obtenus dans le TFE et le cyclohexane ne serait pas correct à cause de leur trop grande différence de propriétés physico-chimiques.

L'absence de structure hélice polyproline II peut être due à la méthode utilisée, les calculs de Monte-Carlo ayant été réalisés en conditions de solvant implicite. Même si ce type de modèle de solvant, dont les termes de solvatation sont calculés en tenant compte de la surface accessible au solvant, est précis et bien adapté aux protéines et aux peptides, il n'en demeure pas moins que la présence de molécules de solvant est importante pour la stabilité, et influence grandement l'évaluation de l'énergie conformationnelle. Un peptide en structure l'hélice polyproline II, en particulier, est stabilisé par les molécules d'eau en formant des liaisons hydrogène avec certains atomes du squelette peptidique (Sreerama et Woody 1999). Il est donc envisageable, pour pallier ce problème de solvatation, d'utiliser le module `MC` du programme CHARMM qui permettrait des calculs de Monte-Carlo dans des conditions de solvant explicite.

De plus, des structures obtenues par Monte-Carlo devraient posséder, notamment au niveau des coudes β, de forts effets NOE visibles en RMN. Cependant, les informations structurales obtenues en RMN ou en DC sont une moyenne de toutes les conformations que peut prendre, suivant sa flexibilité, une structure peptidique. La taille encore relativement petite du peptide Sgal laisse à penser qu'il possède certainement une grande flexibilité.

Ainsi, même si le minimum d'énergie global du peptide Sgal, obtenu par cette méthode de Monte-Carlo, donne une structure repliée, elle est très certainement, expérimentalement, noyée dynamiquement dans la somme des conformations du peptide Sgal.

C'est pourquoi une étude en simulation de dynamique moléculaire de ce peptide a été effectuée avec comme modèle de départ, la structure de plus basse énergie obtenue suite aux calculs de Monte-Carlo, c'est-à-dire la conformation 1 obtenue dans l'eau.

B.2.c. Simulations de dynamique moléculaire du peptide Sgal

Les simulations du peptide Sgal ont été effectuées avec la version c27b3 du programme CHARMM (Brooks et al. 1983) avec le jeu de paramètres "tout atome" PARAM22 (MacKerell et al. 1998). Les interactions non liantes ont été calculées dans un rayon de 14 Å avec la fonction FORCE-SWITCH pour les interactions électrostatiques et la fonction VSHIFT pour les interactions de van der Waals.

Le peptide Sgal a été plongé dans des boîtes de solvant cubiques dont les dimensions ont été choisies de manière à avoir, dans toutes les directions de l'espace, au minimum 10 Å entre le peptide et les limites de la boîte de solvant. Des conditions périodiques aux limites ont été ensuite appliquées. Deux solvants ont été utilisés : l'eau, modèle TIP3P (Jorgensen et al. 1983) et le TFE. Le peptide a été plongé dans la boîte de solvant, précédemment équilibrée et les molécules de solvant situées à une distance inférieure à 2,8 Å des atomes lourds ont été éliminées.

Après la solvatation, le peptide a été minimalisé seul avec 500 pas de SD suivi de 500 pas de GC; le solvant a été ensuite minimalisé seul, puis à nouveau le peptide et le solvant seul encore à son tour. Enfin l'ensemble du système (solvant + peptide Sgal) a été minimalisé de la même façon.

La température du système a été augmentée jusque 300 K durant 15 ps. La phase d'équilibrage du système s'est déroulée en deux étapes consécutives de 20 ps. Dans la première étape, les vitesses des atomes ont été attribuées tous les 50 pas suivant une fonction de distribution gaussienne. Dans la deuxième étape, les vitesses des atomes ont été modifiées de manière à ce que la température ne sorte pas d'une fenêtre de ± 10 K autour de 300 K. Pour la phase de production dans le TFE, la température de 300 K a été régulée par un couplage avec un bain externe, la constante de couplage étant de 5 ps (Berendsen et al. 1984). Dans l'eau, en revanche, aucune contrainte n'a été imposée pour maintenir la température à 300 K. Ce protocole de simulation a déjà prouvé son efficacité (Huynh et al. 2002).

Les modèles de départ des simulations ont été soit le modèle construit par homologie soit le modèle de plus basse énergie obtenue par les calculs de Monte-Carlo dans l'eau (§ **B2bα**). Le **Tableau 09** résume les principales caractéristiques des différentes simulations.

Modèles de départ	Solvants		Durées (ns)	contraintes
	Eau	TFE		
Homologie	X		25	aucune
Homologie	X		30	aucune
Monte Carlo	X		15	aucune
Homologie	X		10	de distance
Homologie		X	40	aucune

Tableau 09 : Conditions de simulation des dynamiques moléculaires de Sgal.

Durant l'analyse, les amas de conformations proches ont été calculés avec l'algorithme de clustérisation ART-2 (Carpenter et Grossberg 1987), implémenté dans le programme CHARMM en tenant compte des angles dièdres des résidus composant le peptide Sgal.

α. Simulations de dynamique moléculaire du peptide Sgal dans l'eau

Les simulations de DM du peptide Sgal ont été effectuées à l'aide d'une boîte cubique de 47 Å d'arête, comprenant 3.472 molécules d'eau. Le système après solvatation comprenait 10.304 atomes dont 10.107 atomes de solvant (3.369 molécules d'eau).

Les deux premières simulations de DM dans l'eau de respectivement 25 et 30 ns ont été effectuées en solvant explicite, en partant du même modèle construit par homologie du peptide Sgal. Ces deux simulations ont été effectuées sur deux plates-formes de calcul différentes, leur donnant ainsi des vitesses initiales différentes.

(i) DM de 25 ns dans l'eau du peptide Sgal (modèle par homologie)

L'analyse de la phase de production de 25 ns n'a révélé la présence d'aucune structure de type hélice ou brin selon le programme d'analyse de structure secondaire STRIDE (Frishman et Argos 1995). Seuls des coudes β ont put être mis en évidence au cours de la trajectoire.

Les pourcentages des différents types de coudes β ont été reportés dans le **Tableau 10**. Le peptide Sgal (**VVGSPSAQDEASPLS**) étant constitué de 15 résidus, 12 coudes β sont, au maximum, possiblement observables. Le tour d'hélice polyproline II présent sur **SPS** dans le modèle de départ, disparaît rapidement en moins de 1 ns.

Les coudes β majoritaires observés sont ceux de type I avec jusqu'à 51,8 % des conformations pour la séquence **PSAQ**.

Il est important de connaître l'évolution au cours de la trajectoire de dynamique moléculaire, des différents coudes β présents (**Figure 59**).

	Types de coudes β (% de conformations)					
	canoniques et déformés					Total
	I	I'	II	II'	VIII	
VVGS	19,0	–	13,1	–	8,6	40,7
VGSP	–	–	–	1,3	0,7	2,0
GSPS	–	–	–	–	–	–
SPSA	–	–	–	–	–	–
PSAQ	51,8	–	–	–	10,6	62,3
SAQD	43,8	–	–	–	7,0	50,8
AQDE	1,8	–	–	–	12,0	13,7
QDEA	27,6	–	1,7	–	–	29,3
DEAS	29,4	–	–	–	12,2	41,6
EASP	–	–	–	–	7,9	7,9
ASPL	–	–	–	–	–	–
SPLS	30,6	–	–	–	8,9	39,5

Tableau 10 : Pourcentages de coudes β observés dans la structure de Sgal lors de la simulation de 25 ns dans l'eau.

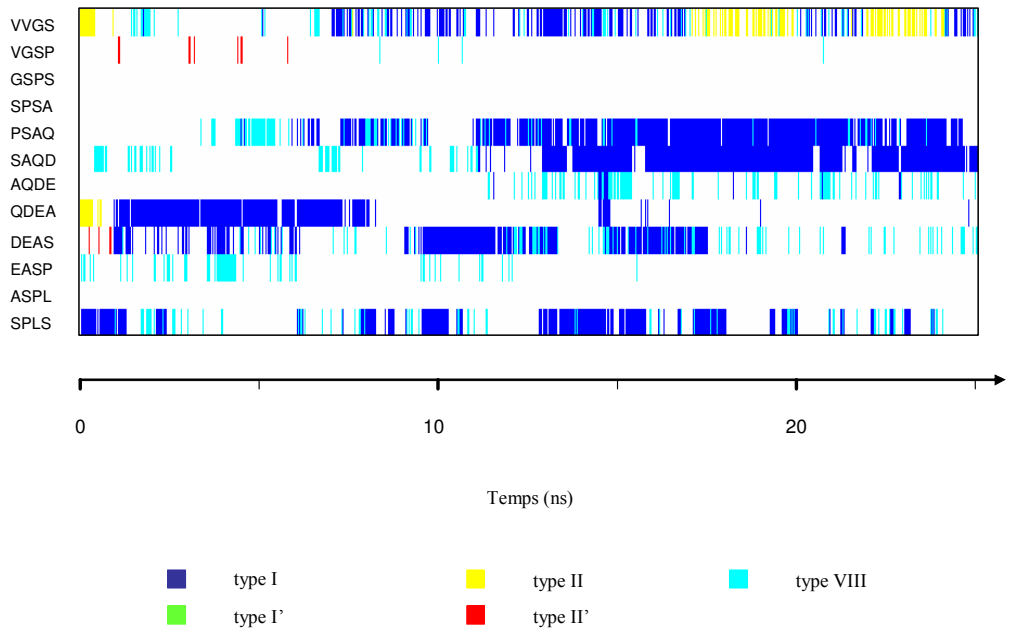


Figure 59 : Présence des différents types de coudes β au cours du temps pour la simulation de 25 ns de Sgal dans l'eau.

La séquence **VVGS** voit son coude β de type II présent initialement dans le modèle par homologie disparaître après 1 ns, pour revenir en coude β de type I ou VIII jusqu'à 16 ns pour revenir en coude β de type II et ensuite passer alternativement du type I ou VIII au type II. Les motifs **PSAQ** et **SAQD** se replient tous deux de manière très stable en coude β de type I après 12 ns de dynamique moléculaire.

La séquence **QDEA** passe à partir d'1 ns d'un coude β de type II comme dans la structure de départ à un coude β de type I jusqu'à 8 ns de trajectoire pour ne plus être en coude β, sauf de manière transitoire jusqu'à la fin de la simulation de dynamique moléculaire.

Le motif **EASP** oscille entre une structure en coude β de type I ou VIII, dont la présence peut être très stable pendant 4 à 5 ns, et une structure non coudée.

On peut enfin voir que la séquence **SPLS** alterne tout au long de la trajectoire entre une structure non coudée et une structure en coude β de type I ou VIII.

Nous avons tracé au cours du temps les zones de la carte de Ramachandran occupées par chacun des résidus (**Figure 60**). Les pourcentages globaux de chacune des zones visitées au cours du temps sont reportés **Tableau 11**.

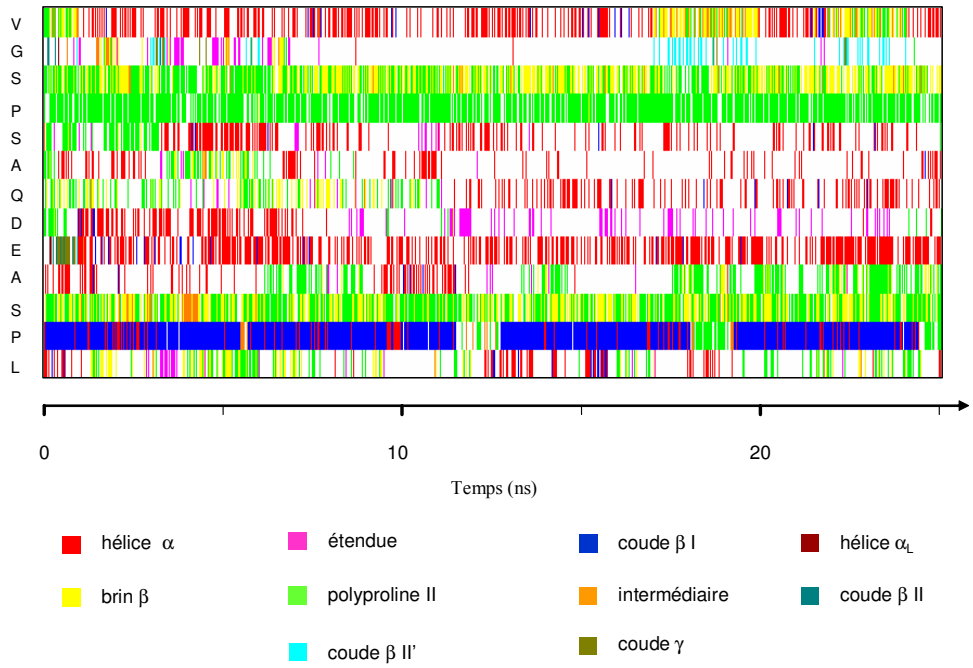


Figure 60 : Zones de la carte de Ramachandran occupées par les résidus de Sgal pour la simulation de 25 ns dans l'eau.

Zones	Pourcentages	Zones	Pourcentages
hélice α	13,0	**intermédiaire**	1,2
brin β	9,9	**hélice α_G**	0,1
étendue	1,5	**coude β II**	0,2
polyproline II	17,7	**coude β II'**	0,5
coude β I	6,5	**coude γ**	0,2
	hors zones	49,3	

Tableau 11 : Pourcentages des zones de la carte de Ramachandran occupées par les résidus du peptide Sgal pour la simulation de 25 ns dans l'eau.

On peut remarquer que les résidus précédant une proline restent dans les zones de brin β ou de la polyproline II (partie nord-ouest de la carte de Ramachandran) comme cela a déjà été observé dans les protéines de structures connues (MacArthur et Thornton 1991). Les autres résidus restent globalement dans les mêmes zones de la carte de Ramachandran, hormis $\mathbf{V_2}$ qui à la fin de la trajectoire alterne entre la zone des hélices α et la zone des brins ou de la polyproline II, expliquant ainsi l'alternance de coudes β de type II avec ceux de type I ou VIII sur le motif **VVGS**. La zone la plus visitée est celle de la polyproline II (17,7 % des résidus), suivie par la zone de l'hélice α puis du brin β et du coude β de type I.

Pour chacune des conformations obtenues au cours de la trajectoire, nous avons calculé sa valeur de RMSD sur les atomes du squelette peptidique, par rapport à toutes les autres conformations de la dynamique. Nous avons tracé la carte des RMSD croisés qui permet de connaître les conformations voisines. On met ainsi en évidence la présence stable d'amas de conformations proches (clusters). Ce type d'analyse est particulièrement utile dans le cas d'une molécule flexible comme le du peptide Sgal (**Figure 61**).

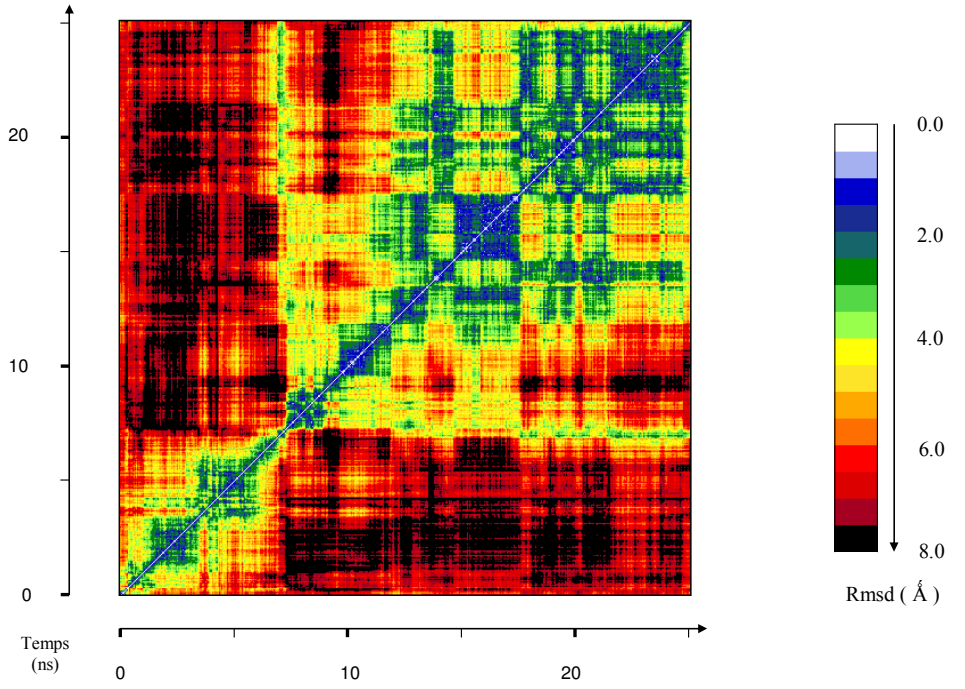


Figure 61 : Carte des RMSD croisés de Sgal pour la simulation de 25 ns dans l'eau.

La carte permet de dégager au moins 3 amas de conformations proches : le premier de 0 à 7 ns, le deuxième de 7 à 12 ns et le dernier de 12 ns jusqu'à la fin de la trajectoire.

L'algorithme de clustérisation ART-2 utilisé avec un rayon de troncature de 69,4° propose 4 amas de conformations proches (**Figure 62**).

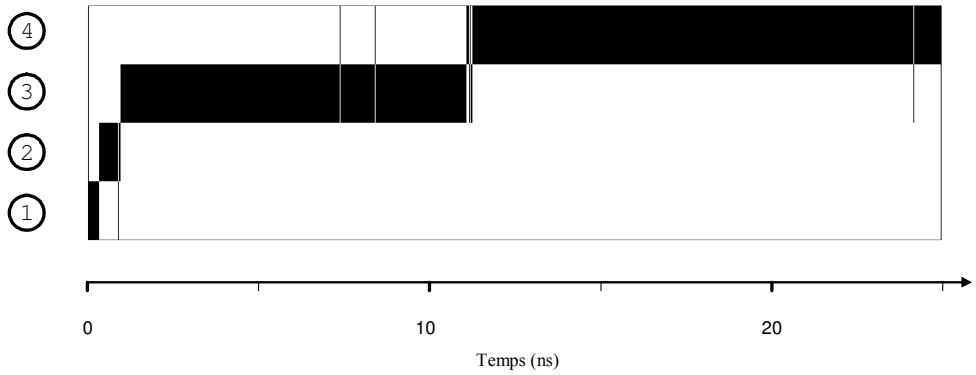


Figure 62 : Amas de conformations de Sgal durant la simulation de 25 ns dans l'eau, calculés par l'algorithme ART-2.

Les deux premiers amas se succèdent rapidement en moins de 1 ns, puis le troisième persiste entre 2 et 12 ns et enfin le quatrième se maintient jusqu'à la fin de la trajectoire. Cette répartition des amas est différente de celle déduite de la carte de RMSD. En fait ART-2 fait rapidement la différence entre 2 amas au début de la trajectoire dans une échelle de temps où il est difficile visuellement de les distinguer. Le troisième amas proposé par ART-2 comprend deux amas de la carte de RMSD croisés, cette différence est probablement due au fait que les deux méthodes ne prennent pas en compte les mêmes paramètres : la carte de RMSD croisés provient du RMSD calculé sur les atomes du squelette protéique, alors que ART-2 tient compte des angles dièdres du squelette protéique. Cette différence fait que ces deux méthodes de clustérisation sont, non pas redondantes, mais complémentaires. Le dernier amas proposé par ces deux méthodes est d'ailleurs le même : il débute à 12 ns de la trajectoire et se maintient jusqu'à sa fin.

amas	Types de coudes β						Occurrences (% des conformations)
	V_1VGS_4	P_5SAQ_8	S_6AQD_9	A_7QDE_{10}	Q_8DEA_{11}	$E_{10}ASP_{13}$	
1	IV (II)	-	-	-	IV (II)	IV (VIII)	1,3
2	-	IV (VIII)	-	-	IV (II)	IV (VIII)	2,4
3	-	IV (I)	-	-	-	-	40,8
4		IV (I)	IV (I)	IV (VIII)		-	55,5

Tableau 12 : Coudes β des amas de conformations proches calculés par ART-2, associés à leur occurrence au cours de la trajectoire de 25 ns de Sgal dans l'eau.

Les structures moyennes de ces amas sont représentées ci-dessous (**Figure 63**) et l'ensemble des coudes β les composants sont décrit ci-dessus **Tableau 12**. On peut y voir, notamment, que le dernier amas comprend 55,5 % de l'ensemble des conformations de la trajectoire.

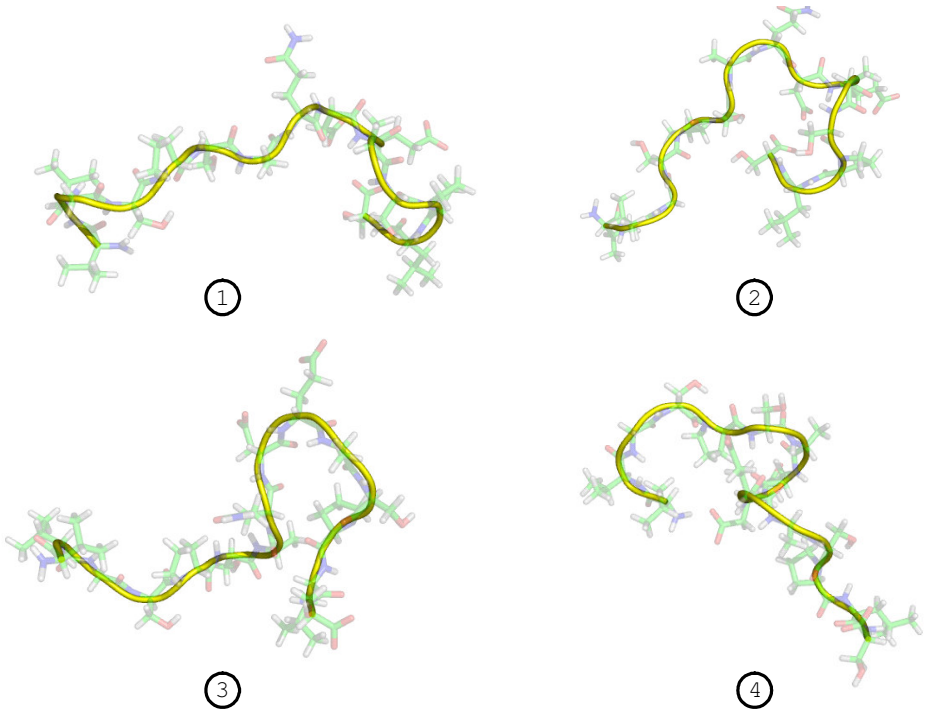


Figure 63 : Structure moyenne de chaque amas de conformations calculées par l'algorithme ART-2 au cours de la trajectoire de 25 ns de Sgal dans l'eau.

Le suivi des liaisons hydrogène au cours de la trajectoire nous montre que seules deux liaisons hydrogène impliquant des atomes du squelette peptidique sont visibles (**Figure 64**).

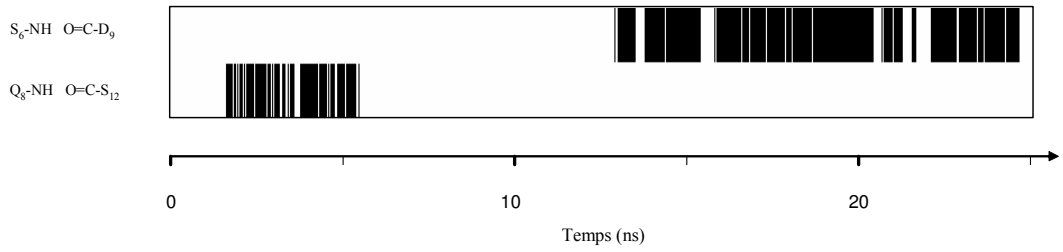


Figure 64 : Présence de liaisons H durant la simulation de 25 ns de Sgal dans l'eau.

De 1,3 à 5,4 ns, la première liaison entre le **N-H** de **Q_8** et le **C=O** de **S_{12}**, stabilise les coudes β de type I sur les séquences **QDEA** et **DEAS**. La seconde liaison H est présente à partir de 12 ns jusqu'à la fin de la trajectoire entre le **NH** de **S_6** et le **C=O** de **D_9**, c'est-à-dire en même

temps de que le dernier amas de conformations, et qui permet de stabiliser le coude β de type I du motif **SAQD**.

(ii) DM de 30 ns dans l'eau du peptide Sgal (modèle par homologie)

La seule structure secondaire détectée est le coude β dont les pourcentages de présence des différents types dans les conformations analysées ont été reportés dans le **Tableau 13**.

	Types de coudes β (% des conformations)					
	canoniques et déformés					**Total**
	I	**I'**	**II**	**II'**	**VIII**	
VVGS	0,8	–	14,5	–	6,1	21,4
VGSP	–	–	–	3,0	3,8	6,8
GSPS	–	–	–	–	–	–
SPSA	23,8	–	–	–	11,0	34,8
PSAQ	2,0	–	–	–	8,8	10,8
SAQD	5,2	–	–	–	5,9	11,1
AQDE	8,9	–	–	–	2,1	11,0
QDEA	45,2	–	2,2	–	4,6	52,0
DEAS	8,7	–	–	–	2,8	11,6
EASP	–	–	–	–	0,6	0,6
ASPL	–	–	–	–	–	–
SPLS	0,7	–	–	–	0,8	1,6

Tableau 13 : Pourcentages de coudes β observés dans la structure de Sgal lors de la simulation de 30 ns dans l'eau.

Ainsi, tout comme dans la première DM de 25 ns, l'analyse de cette phase de production de 30 ns de cette seconde DM n'a pu mettre en évidence la présence de structures secondaires de type hélice ou brin.

Le tour d'hélice polyproline II sur **SPS**, présent dans le modèle de départ, se maintient uniquement pendant les 3 premières nanosecondes pour disparaître ensuite.

La séquence la plus souvent en coude β est **QDEA** avec 45,2 % des conformations en coude β de type I, pour un total de 52,0 % en un quelconque type de coude β sur cette séquence. Les autres séquences qui présentent souvent un coude β sont les motifs **SPSA** et **VVGS** avec respectivement 34,8 % et 21,4 % des conformations échantillonnées au cours de la trajectoire.

La présence, en fonction du temps, des différents types de coudes β sur chacun des motifs est donnée plus bas (**Figure 65**).

Le coude β de type II sur la séquence **VVGS**, initialement présent dans le modèle de départ, se maintient pendant 3 ns puis réapparaît transitoirement et en alternance avec un

coude β de type VIII. La séquence **QDEA** passe, après 2 ns de simulation d'un coude β de type II, initialement présent dans le modèle de départ, à un coude β de type I qui reste observable pratiquement tout au long de la trajectoire. Le motif **SPSA**, qui a de forts pourcentages de coudes β, oscille entre une structure non coudée et une structure en coude β de type I ou VIII.

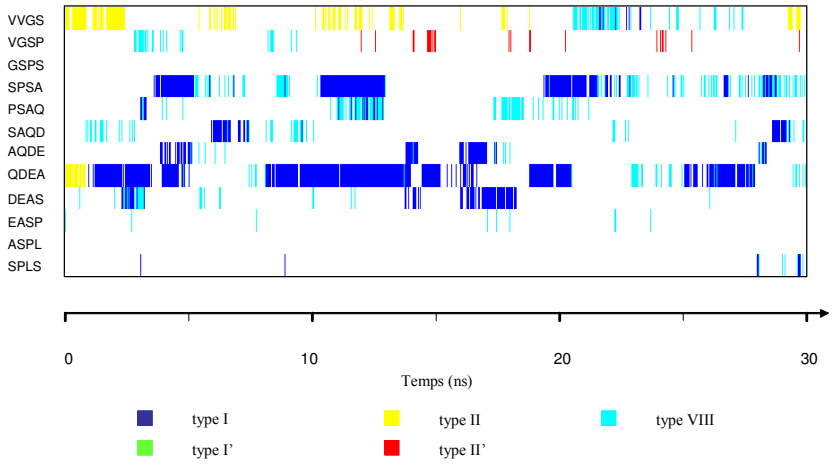


Figure 65 : Présence des différents types de coudes β au cours du temps pour la simulation de 30 ns de Sgal dans l'eau.

Les zones de la carte de Ramachandran explorées par chacun des résidus au cours de la DM sont reportées **Figure 66**, ainsi que les pourcentages d'occupation de ces zones en **Tableau 14**.

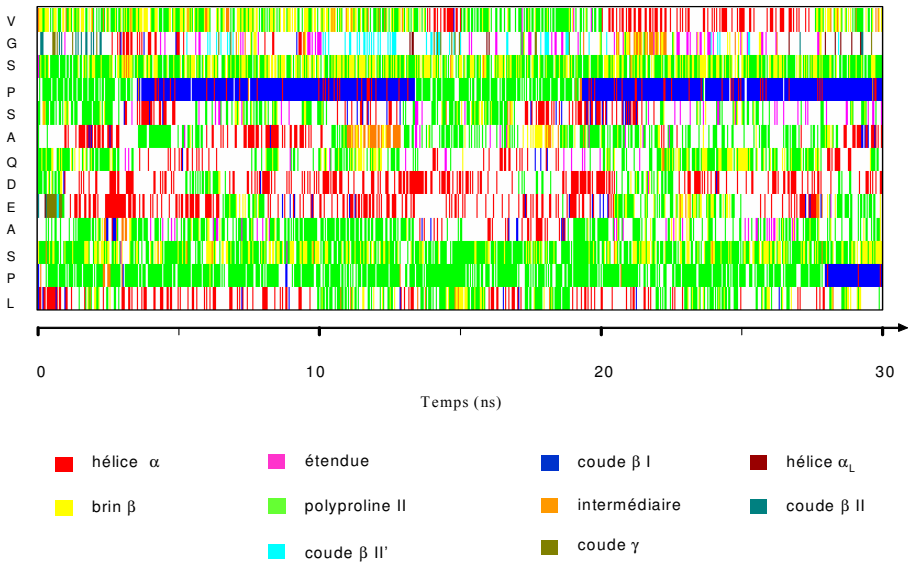


Figure 66 : Zones de la carte de Ramachandran occupées par les résidus du peptide Sgal pour la simulation de 30 ns dans l'eau.

Zones	Pourcentages	Zones	Pourcentages
hélice α	9,7	**intermédiaire**	1,4
brin β	9,8	**hélice α_G**	0,1
étendue	1,4	**coude β II**	0,2
polyproline II	28,6	**coude β II'**	0,5
coude β I	5,9	**coude γ**	0,3
	hors zones	42,1	

Tableau 14 : Pourcentages des zones de la carte de Ramachandran occupées par les résidus de Sgal pour la simulation de 30 ns dans l'eau.

On s'aperçoit que la zone la plus visitée au cours de la DM est celle de la polyproline II (28,6 % des résidus), suivie par celle de l'hélice α, et celle du brin β.

La **Figure 67** représente la carte des RMSD croisés obtenue pour cette simulation, en ne prenant en compte que les atomes du squelette peptidique.

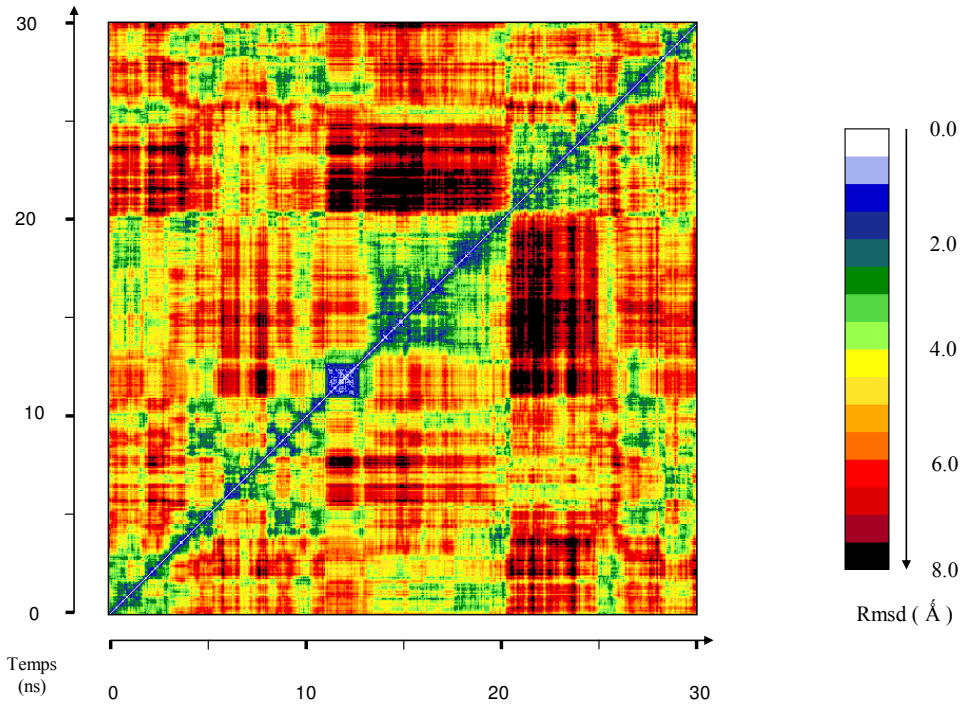


Figure 67 : Carte des RMSD croisés de Sgal pour la simulation de 30 ns dans l'eau.

Il est difficile d'après cette carte de RMSD de dégager des amas de conformations proches. Le peptide Sgal démontre un comportement très entropique et très flexible puisqu'il ne semble pas, dans cette DM, passer par une succession de différents amas de conformations. Cependant, un amas est visible de 14 à 20 ns et un autre de 20 à 25 ns de trajectoire.

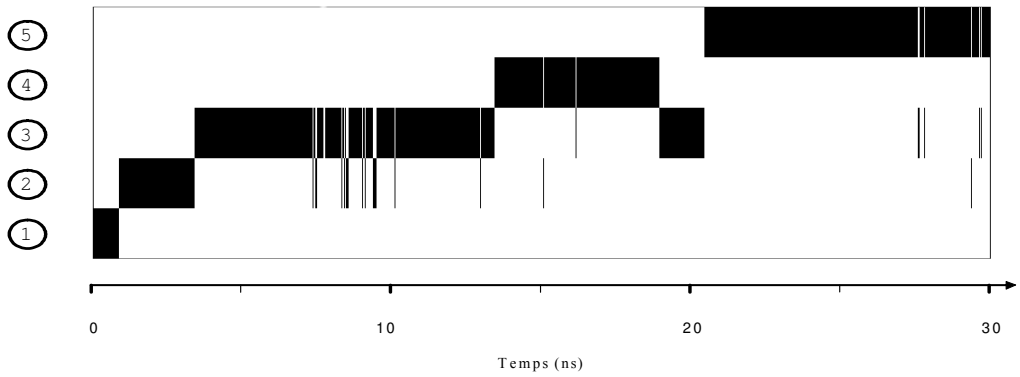


Figure 68 : Amas de conformations de Sgal durant la simulation de 30 ns dans l'eau..

L'algorithme de clustérisation ART-2 prend ici toute son importance (**Figure 68**). Avec un rayon de troncature de 64°, ART-2 nous propose 5 amas de conformations proches, dont les structures moyennes sont représentées ci dessous **Figure 69** et le détail des structures secondaires qui les composent dans le **Tableau 15**.

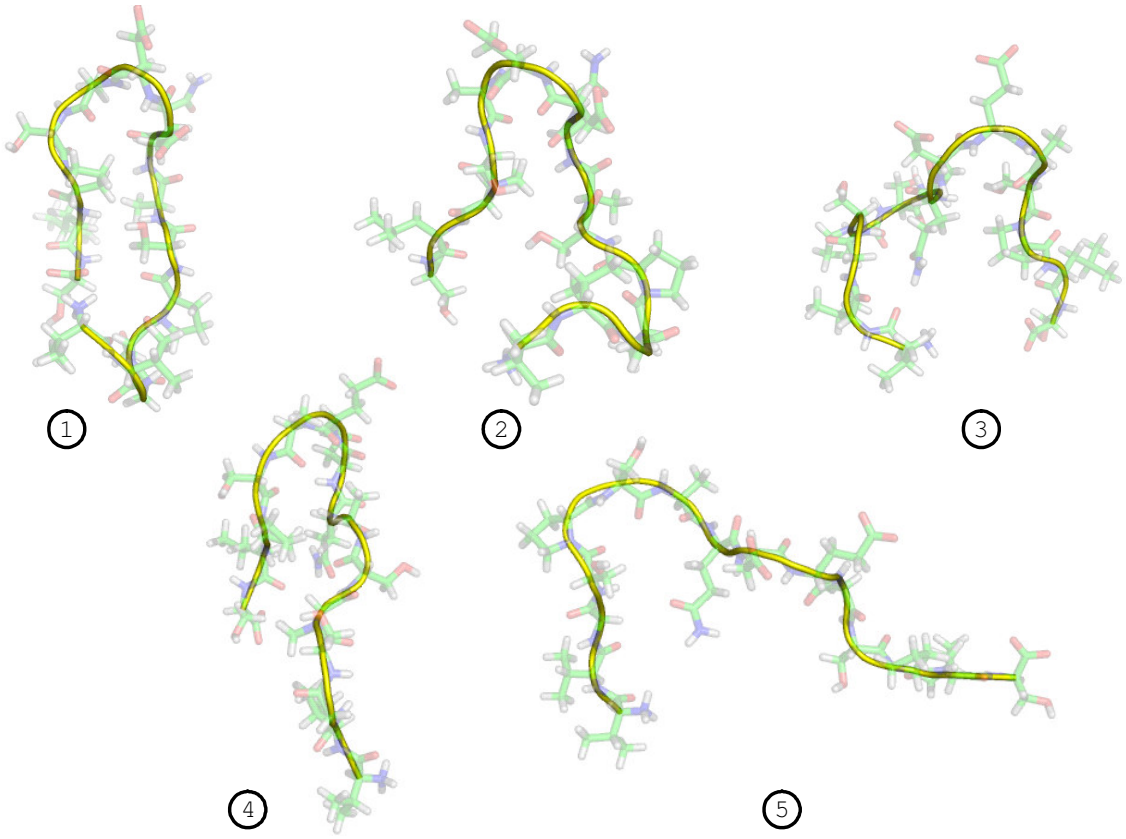


Figure 69 : Structure moyenne de chaque amas de conformations calculés par ART-2 au cours de la trajectoire de 30 ns de Sgal dans l'eau.

amas	Types de coudes β					Occurrences (% des conformations)
	V_1VGS_4	S_4PSA_7	A_7QDE_{10}	Q_8DEA_{11}	D_9EAS_{12}	
1	IV (II)	-	-	IV (II)	IV (VIII)	2,8
2	-		-	IV (I)	IV (VIII)	8,9
3	-	IV (I)	-	-	-	37,6
4	-	-	IV (I)	-	-	18,4
5	-	IV (VIII)		-	-	32,2

Tableau 15 : Coudes β des amas de conformations proches calculés par ART-2, associés à leur occurrence au cours de la trajectoire de 30 ns de Sgal dans l'eau.

(iii) DM de 15 ns dans l'eau du peptide Sgal (modèle MC de plus basse énergie)

Le modèle de départ choisi pour cette simulation dans l'eau est la structure MC de plus basse énergie obtenue avec le programme FANTOM (§ **B.2.b.α.**). L'analyse de la trajectoire d'une durée de 15 ns permet de mettre en évidence la présence de différents coudes β (**Tableau 16**), mais pas celles d'autres structures secondaires.

	Types de coudes β (% des conformations)					
	canoniques et déformés					Total
	I	I'	II	II'	VIII	
VVGS	–	–	2,4	–	–	2,4
VGSP	–	–	–	5,0	34,2	39,2
GSPS	–	–	–	–	–	–
SPSA	33,6	–	–	–	–	33,6
PSAQ	36,1	–	–	–	9,0	45,1
SAQD	20,4	–	0,8	–	11,1	32,3
AQDE	4,4	–	–	–	0,5	4,9
QDEA	21,5	–	–	–	8,4	29,4
DEAS	–	–	–	–	5,4	5,4
EASP	–	–	–	–	6,0	6,0
ASPL	–	–	–	–	–	–
SPLS	44,2	–	–	–	6,8	51,0

Tableau 16 : Pourcentages de coudes β observés dans la structure de Sgal lors de la simulation de 15 ns dans l'eau.

La distribution au cours du temps des différents types de coude β est donnée plus bas **Figure 70**. La séquence **SPLS** présente tout au long de la trajectoire un coude β (51,0 % des conformations analysées) de type I ou VIII. A l'inverse, le motif **PSAQ** se replie en coude β de type I ou VIII de manière ponctuelle : au début de la simulation jusque 2,0 ns puis entre 4,0 et 5,5 ns et enfin de 10,5 ns jusqu'à la fin de la trajectoire. Le coude β de type I sur **SPSA**

n'apparaît qu'à la fin de la trajectoire à partir de 9,5 ns. De même, le coude β de type I sur **QDEA** n'est présent qu'entre 5,0 et 9,0 ns.

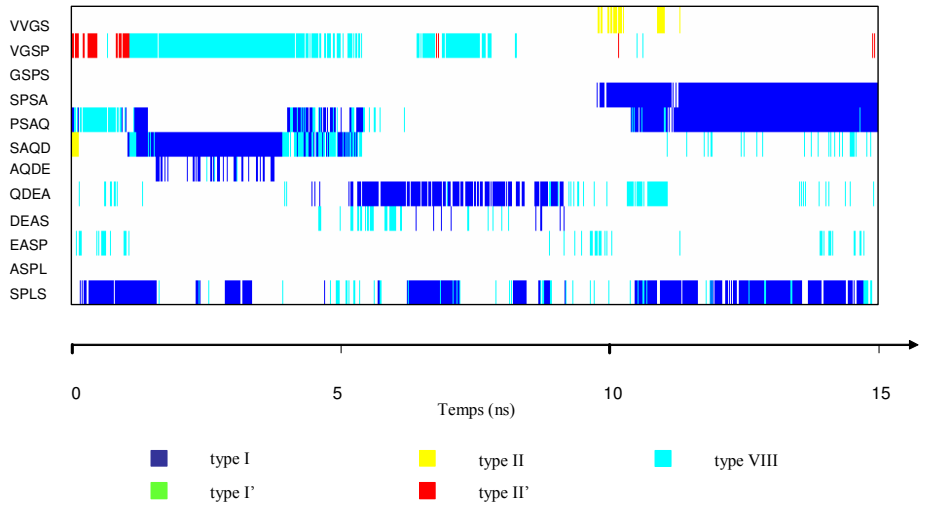


Figure 70 : Présence des différents types de coudes β au cours du temps pour la simulation de 15 ns de Sgal dans l'eau.

Les zones de la carte de Ramachandran explorées au cours du temps par chacun des résidus et les pourcentages des résidus explorant ces zones sont reportés respectivement en **Figure 71** et dans le **Tableau 17**.

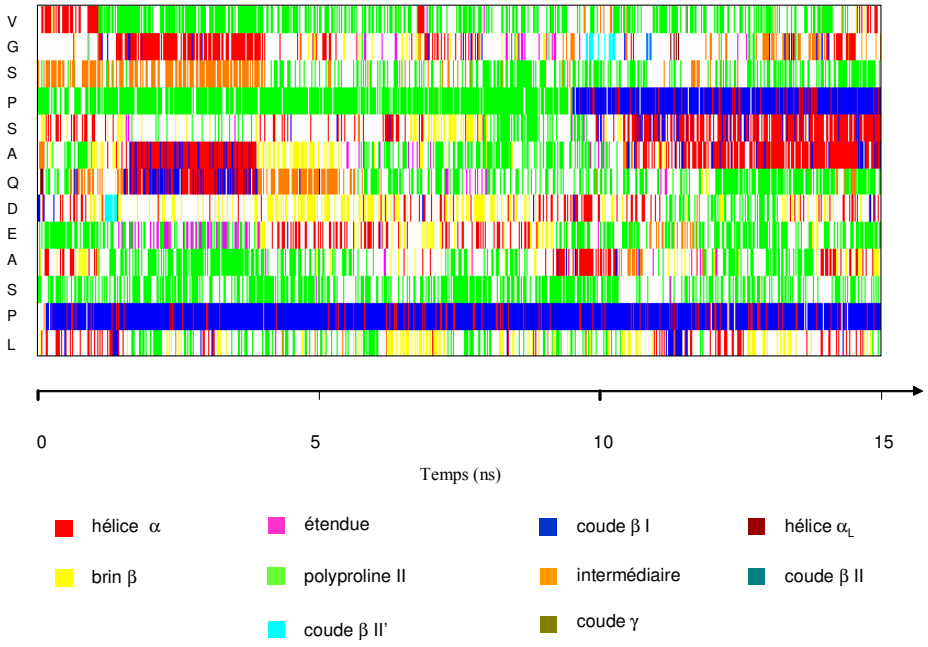


Figure 71 : Zones de la carte de Ramachandran occupées par les résidus de Sgal pour la simulation de 15 ns dans l'eau.

Zones	Pourcentages	Zones	Pourcentages
hélice α	11,7	**intermédiaire**	3,7
brin β	4,8	**hélice α_G**	0,1
étendue	1,2	**coude β II**	0,1
polyproline II	23,6	**coude β II'**	0,2
coude β I	11,1	**coude γ**	0,1
	hors zones	43,5	

Tableau 17 : Pourcentages des zones de la carte de Ramachandran occupées par les résidus de Sgal pour la simulation de 15 ns dans l'eau.

Comme pour les DM précédentes, la zone la plus la visitée est celle de la polyproline II (23,6 % des résidus), puis celle de l'hélice α et celle du coude β de type I.

La carte des RMSD croisés (**Figure 72**) montre une succession d'amas de conformations proches sur de courtes durées, hormis celui qui se maintient de 2 à environ 7 ns.

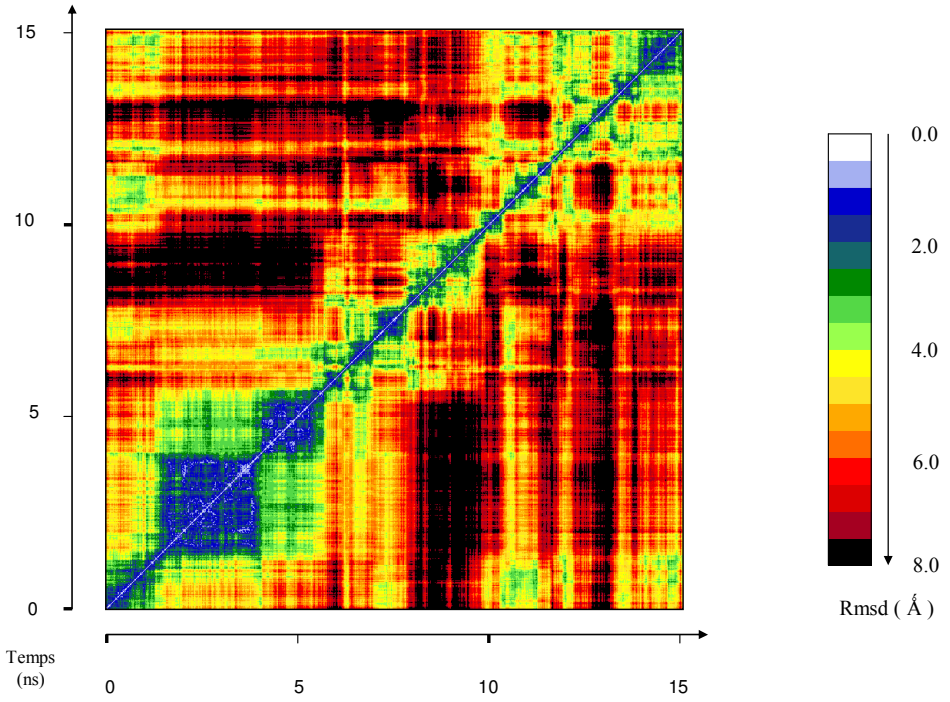


Figure 72 : Carte des RMSD croisés de Sgal pour la simulation de 15 ns dans l'eau.

6 amas de conformations proches, proposés par ART-2 (**Figure 73**) avec un rayon de troncature de 58,3°, se succèdent rapidement au cours de la trajectoire de DM comme le montre la carte des RMSD croisés.

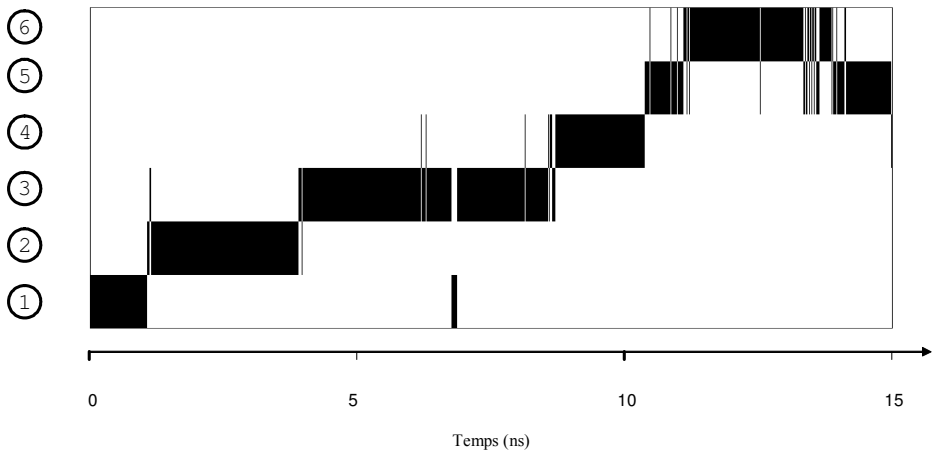


Figure 73 : Amas de conformations de Sgal calculés par ART-2 durant la simulation de 15 ns dans l'eau.

Les structures moyennes des amas de conformations proches sont données en **Figure 74** et leur composition en coudes β dans le **Tableau 18**.

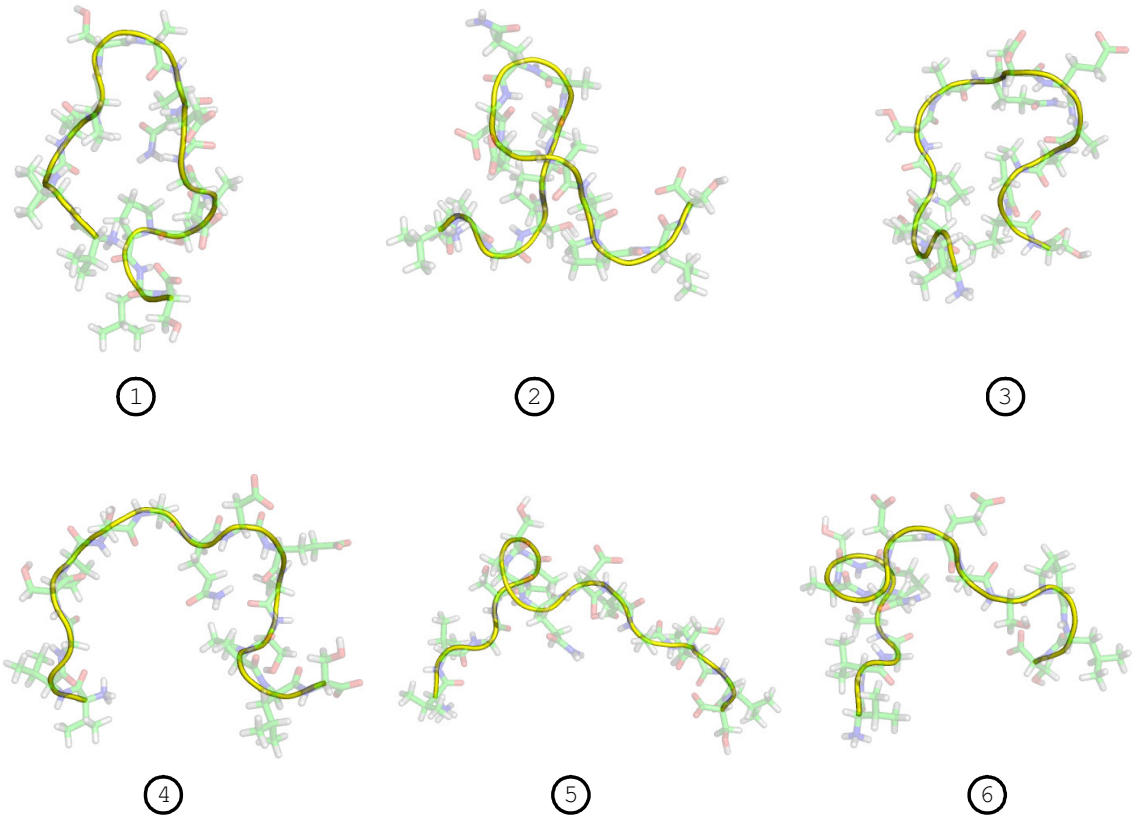


Figure74 : Structure moyenne de chaque amas de conformations calculés par ART-2 au cours de la trajectoire de 15 ns de Sgal dans l'eau.

amas	Types de coudes β					Occurrences (% des conformations)
	V_2GSP_5	S_4PSA_7	P_5SAQ_8	S_6AQD_9	Q_8DEA_{11}	
1	-	-	IV (VIII)	-	-	7,9
2	VIII	-	-	IV (VIII)	-	18,8
3	-	-	-	-	-	30,6
4	-	-	-	-	-	11,9
5	-	IV (I)	IV (I)	IV (VIII)	-	12,7
6	-	IV (I)	IV (I)	IV (VIII)	IV (VIII)	18,1

Tableau 18 : Coudes β des amas de conformations proches calculés par ART-2, associés à leur occurrence au cours de la trajectoire de 15 ns de Sgal dans l'eau.

(iv) Comparaison entre les 3 DM sans contrainte du peptide Sgal dans l'eau

Ces 3 trajectoires de DM effectuées dans l'eau et sans aucune contrainte ont en commun d'avoir la zone de la carte de Ramachandran de la polyproline II comme étant la plus visitée par les résidus avec respectivement 17,7 %, 28,6 % et 23,6 % des résidus dans cette zone.

Maintenant, quand on considère l'addition des 3 simulations de DM effectuées comme une seule trajectoire de DM, il est possible de tracer sa carte des RMSD croisés pour voir quels sont les amas de conformations proches qui se retrouvent au cours des différentes simulations de DM du peptide Sgal dans l'eau (**Figure 75**).

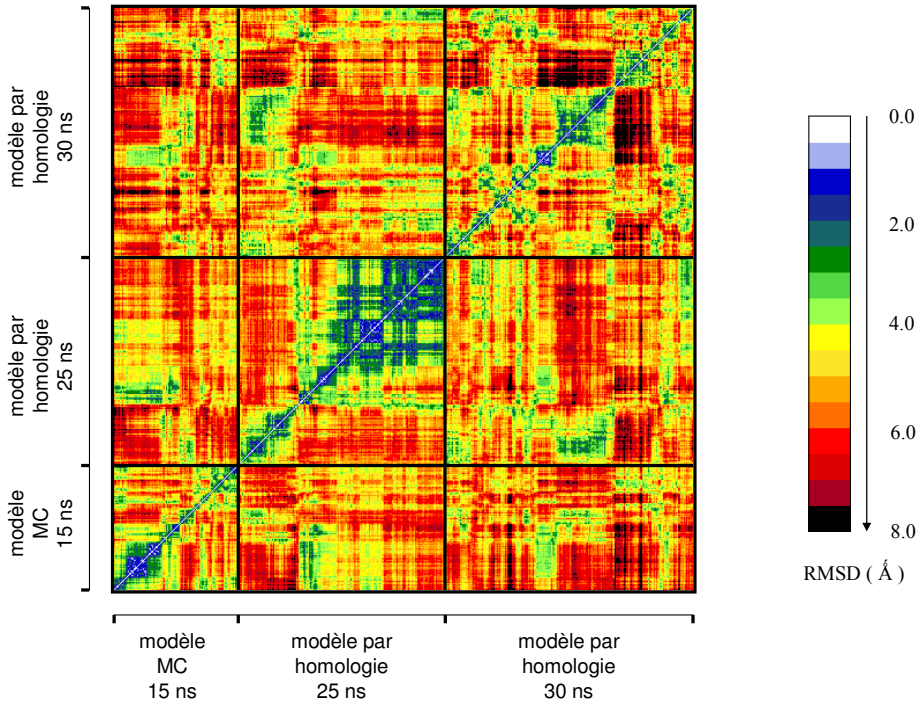
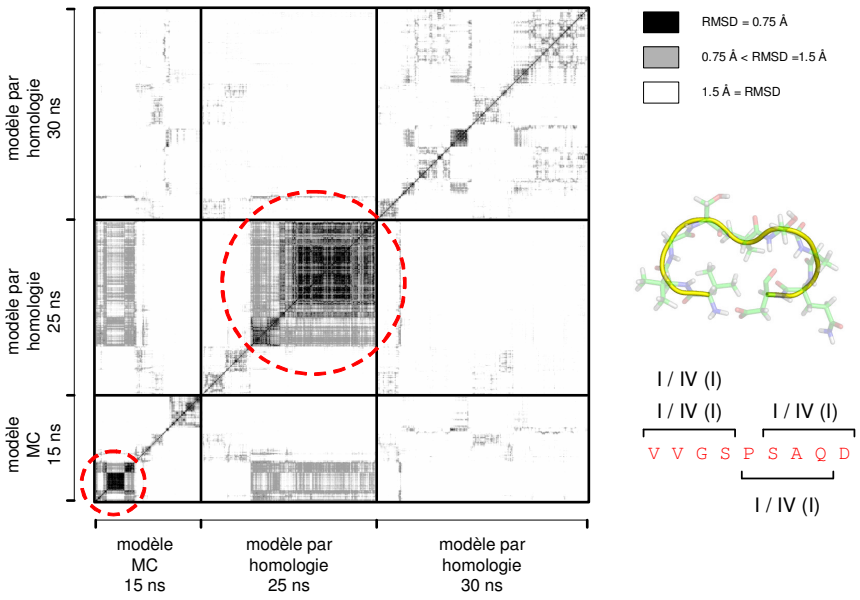


Figure 75 : Carte des RMSD croisés de Sgal pour les 3 simulations dans l'eau.

D'après les calculs de Monte-Carlo dans l'eau, les modèles de plus basses énergies possèdent 3 coudes β successifs : sur **VGSP** de type VIII ou IV proche du type VIII, **PSAQ** et **SAQD**, tous les deux de type VIII ou IV proche du type VIII. De plus, les coudes β sur **PSAQ** et **SAQD** sont les plus fréquents après ceux sur le motif **QDEA**, avec respectivement 42,6 % et 34,8 % d'apparition, au cours des 3 DM.

Nous avons donc tracé la carte des RMSD croisés des 9 premiers résidus (**V_1VGSPSAQD$_9$**) du peptide Sgal (**Figure 76**).



<u>Figure 76</u> : Carte des RMSD croisés des résidus VVGSPSAQD de Sgal pour les 3 DM dans l'eau et structure moyenne des amas de conformations cerclés en rouge.

Deux amas de conformations proches sont visibles durant la majorité de la DM de 25 ns, soit de 7 ns à 25 ns, et durant les 5 premières ns de DM de 15 ns dont le modèle départ est la conformation MC de plus basse énergie. Cependant, même si dans ces amas de conformations proches, les coudes de type I ou IV proche du type I sont bien présents sur les séquences **PSAQ** et **SAQD**, aucun coude β ne peut être mis en évidence sur **VGSP** ; mais un coude β est en revanche observable sur **VVGS**. Ce dernier coude β est d'ailleurs le coude β le plus présent après ceux observés sur **PSAQ** et **SAQD** avec 28,2 % de présence au cours des 3 DM. Ce motif structural est en fait plus proche de celui observé sur la conformation MC de plus basse énergie n°5 comme le montrent les graphiques représentant les RMSD des résidus **VVGSPSAQD** de cette structure par rapport aux conformations obtenues lors des DM de 15 et 25 ns du peptide Sgal (**Figure 77**).

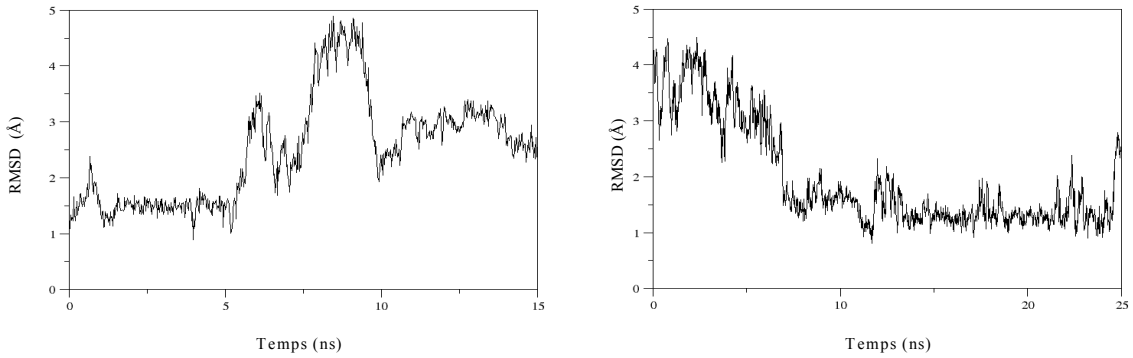


Figure 77 : RMSD des conformations de Sgal par rapport à la structure MC 5 obtenue lors des DM de 15 ns (à gauche) et de 25 ns (à droite).

La présence des amas de conformations proches durant les 5 premières ns de la DM de 15 ns et entre 7 et 25 ns de la DM de 25 ns correspond bien aux périodes de temps où le RMSD est le plus faible.

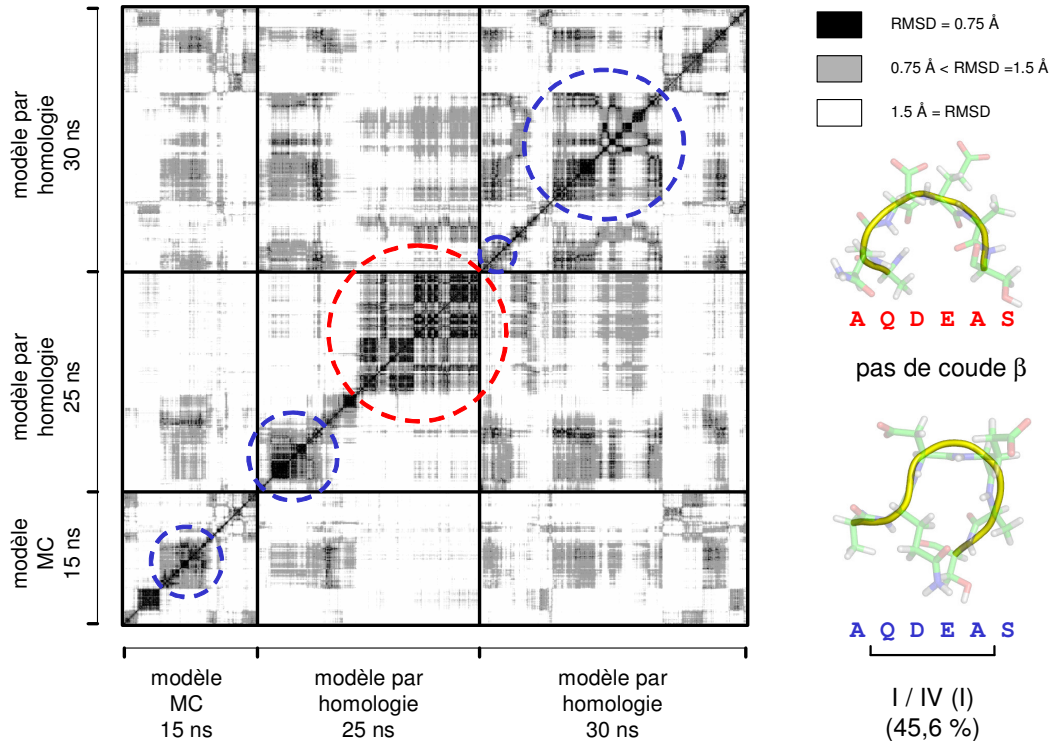


Figure 78 : Carte des RMSD croisés des résidus AQDEAS de Sgal pour les 3 DM dans l'eau. Structure moyenne de l'amas de conformations cerclé de rouge en haut à droite et des amas de conformations cerclés de bleu en bas à droite.

Cette carte des RMSD croisés (**Figure 78**) est obtenue pour la séquence **AQDEAS** définie en ajoutant un résidu de part et d'autre de la séquence **QDEA** pour que les valeurs des RMSD soient bien significatives.

Le coude β de type I sur **QDEA** est présent dans les 3 amas de conformations proches visibles quelques nanosecondes dans les 3 simulations de DM.

(v) DM de 10 ns avec contrainte de distance dans l'eau du peptide Sgal (modèle par homologie)

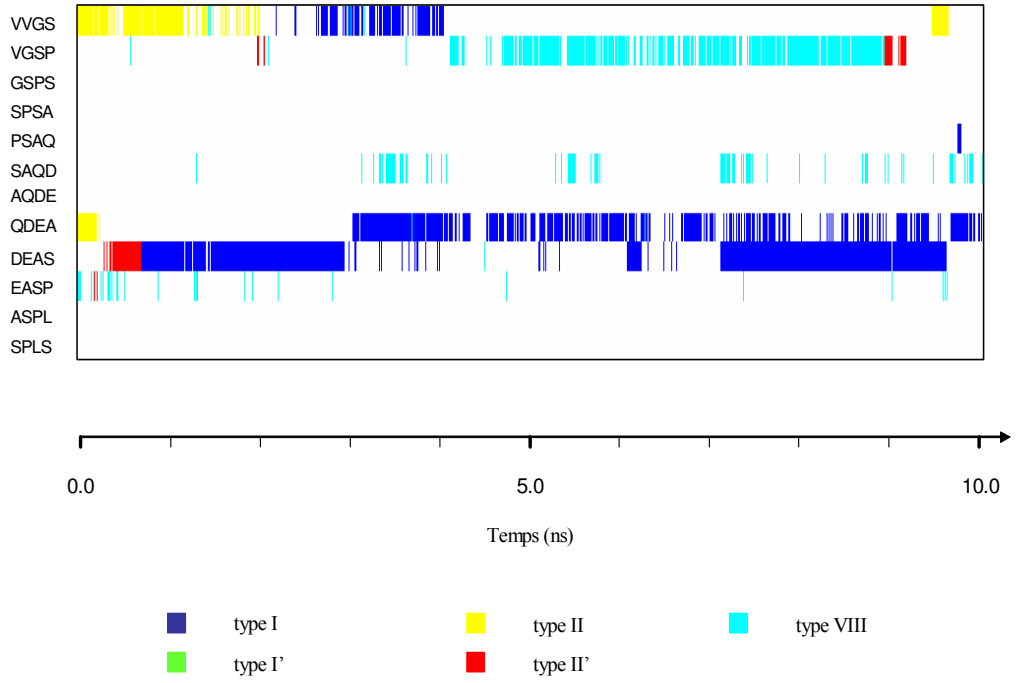
La partie homologue au peptide Sgal dans l'élastase porcine possède une distance C_α-C_α de 13,9 Å entre $\mathbf{V_1}$ et $\mathbf{Q_{15}}$. Pour étudier le comportement du peptide Sgal avec les contraintes que pourrait lui imposer une protéine entière comme l'EBP, nous avons simulé le peptide Sgal en imposant une force de 1000 kcal.mol^{-1} à cette distance C_α-C_α de 13,9 Å entre $\mathbf{V_1}$ et $\mathbf{Q_{15}}$.

Les seules structures ordonnées qui ont été mises en évidence sont là encore les coudes β (**Tableau 19**). Les coudes β, quasi exclusivement de type I, les plus fréquents sont : **QDEA** (42,9 %) et **DEAS** (53,9 %). Du côté N-terminal, **VVGS** est en coude β dans 23,9 % des conformations et **VGSP** est en coude β pratiquement exclusivement de type VIII (35,1 %).

	Types de coudes β (%)					
	canoniques et déformés					**Total**
	I	**I'**	**II**	**II'**	**VIII**	
VVGS	9,1	–	14,3	–	0,5	23,9
VGSP	–	–	–	1,8	35,1	36,9
GSPS	–	–	–	–	–	–
SPSA	–	–	–	–	–	–
PSAQ	0,4	–	–	–	–	0,4
SAQD	–	–	–	–	9,9	9,9
AQDE	–	–	–	–	–	–
QDEA	41,9	–	0,7	–	0,3	42,9
DEAS	50,3	–	–	3,6	–	53,9
EASP	–	–	–	0,2	3,4	3,6
ASPL	–	–	–	–	–	–
SPLS	–	–	–	–	–	–

Tableau 19 : Pourcentages des types de coude β de Sgal dans l'eau avec une contrainte de distance.

Le suivi de la présence des coudes β et de leurs différents types permet de mieux comprendre leur évolution (**Figure 79**).



<u>Figure 79</u> : Présence des différents types de coudes β de Sgal dans l'eau, au cours du temps pour la simulation sous contrainte de distance.

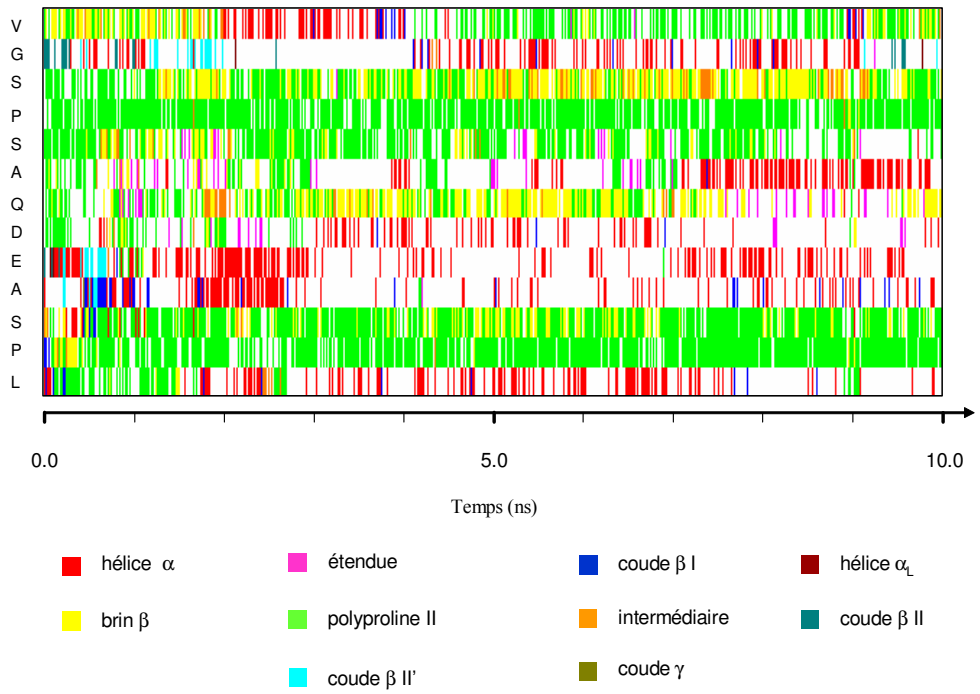
Ainsi, le coude β de type II sur **VVGS** (modèle de départ de la DM) disparaît progressivement au bout de 1,5 ns pour devenir un coude β de type I pendant environ 1,5 ns, puis n'est plus en coude β hormis vers 9,5 ns de trajectoire où il redevient un coude β de type II.

Le coude β de type VIII sur **VGSP** apparaît après 4 ns de trajectoire, reste stable pendant environ 5 ns puis devient un coude β de type I pour finalement disparaître.

QDEA, initialement en coude β de type II dans le modèle de départ, voit cette structure en coude disparaître en moins de 0,3 ns pour se replier à nouveau en coude β, mais cette fois-ci de type I et ce jusqu'à la fin de la trajectoire.

Enfin, le motif **DEAS** qui présente le plus de coudes β au cours de cette trajectoire, adopte, dans un premier temps un coude β de type II' après 0,3 ns puis devient un coude β de type I très stable jusqu'à 3 ns de dynamique moléculaire. De 3 à 7 ns, cette séquence est en coude β de type I de manière transitoire, cette structure redevenant ensuite très stable pendant 2,5 ns.

Nous avons tracé **Figure 80**, l'occupation des différentes zones de la carte de Ramachandran pour chacun des résidus, en fonction du temps. Le **Tableau 20** nous donne les pourcentages d'occupation de chacune des zones par les différents résidus.



<u>Figure 80</u> : Zones de la carte de Ramachandran occupées par les résidus de Sgal dans l'eau et sous contrainte de distance.

La principale zone visitée par les résidus est celle de la polyproline II (29,6 % des résidus) soit en moyenne 4 résidus en polyproline II par conformation échantillonnée.

Zones	Pourcentages	Zones	Pourcentages
hélice α	9,4	**intermédiaire**	1,2
brin β	9,2	**hélice α_G**	0,1
étendue	1,4	**coude β II**	0,3
polyproline II	29,6	**coude β II'**	0,5
coude β I	1,0	**coude γ**	0,5
	hors zone	46,8	

<u>Tableau 20</u> : Pourcentages des zones de la carte de Ramachandran occupées par les résidus de Sgal dans l'eau et sous contrainte de distance.

La carte des RMSD croisés permet de montrer qu'un ou deux amas de conformations proches sont visibles durant la première nanoseconde, puis un autre amas de conformations proches apparaît au bout de 2 ns et se maintient jusqu'à 9 ns (**Figure 81**).

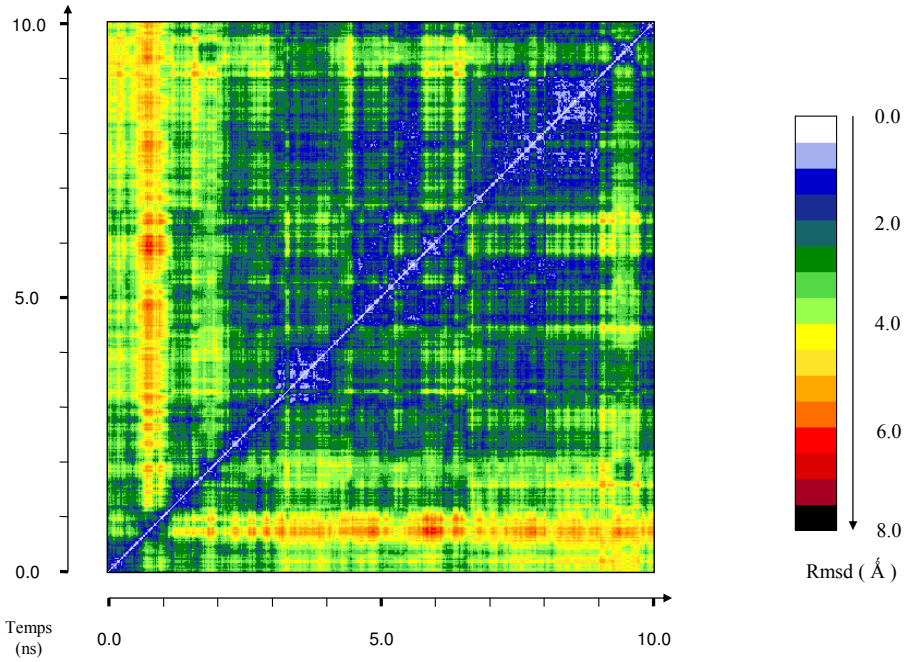


Figure 81 : Carte des RMSD croisés de Sgal dans l'eau et sous contrainte de distance.

Cinq amas de conformations ont été obtenus par ART-2, avec un rayon de troncature de 56,5° (**Figure 82)**.

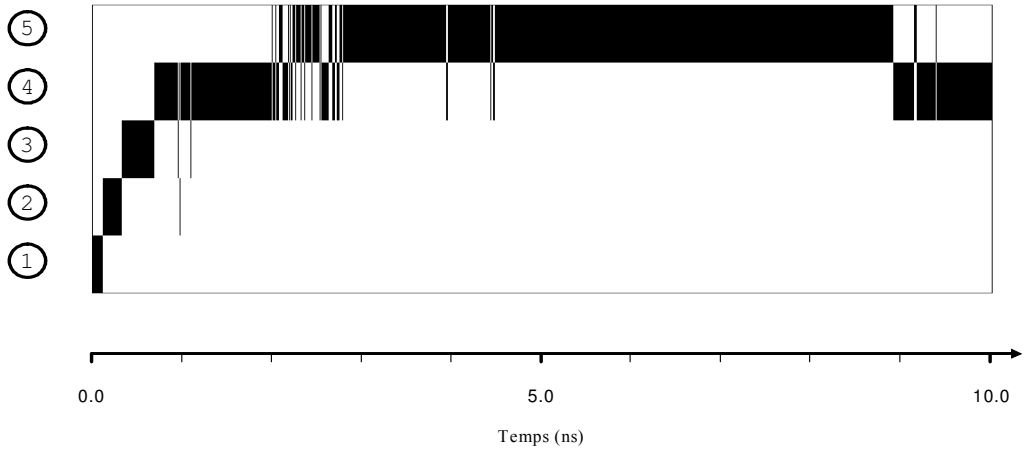


Figure 82 : Amas de conformations au cours du temps calculés par ART-2.

Trois amas de conformations se succèdent dans la première nanoseconde puis l'amas 4 est présent entre 1 et 2 ns puis l'amas 5 jusqu'à 9 ns et enfin les structures redeviennent proches de l'amas 4. On peut donc constater qu'ART-2 donne des résultats proches de ceux obtenus à partir de la carte des RMSD croisés.

Les structures représentatives de chacun des amas sont reportées en **Figure 83**, les coudes β qui les composent, ainsi que l'occurrence de chacun des clusters, étant reportés dans le **Tableau 21**.

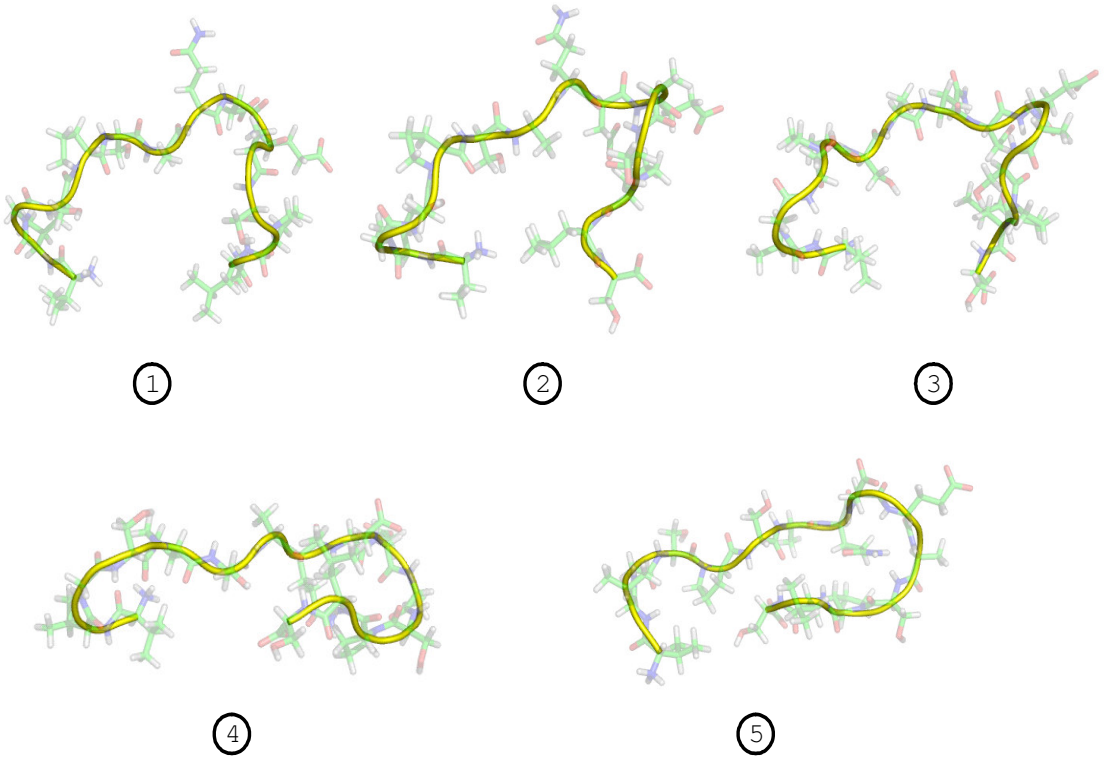


Figure 83 : Structures représentatives de chaque amas de conformations calculés par ART-2 au cours de la trajectoire de Sgal dans l'eau et sous contrainte de distance.

On s'aperçoit ainsi que le coude β sur **QDEA** est présent dans pratiquement tous les amas. Il passe d'un coude β de type II (amas 1 et 2), qui était présent dans le modèle de départ, à un coude β de type I visible dans les amas 4 et 5.

amas	Types de coudes β				Occurrences (%)
	V_1VGS_4	Q_8DEA_{11}	D_9EAS_{12}	$E_{10}ASP_{13}$	
1	II	IV (II)	-	IV (VIII)	1,1
2	IV (II)	IV (II)	IV	VIII	2,4
3	IV	-	IV (II')	-	3,8
4	IV (II)	IV (I)	-	-	27,2
5	-	IV (I)	-	-	65,5

Tableau 21 : Coudes β des amas de conformations proches calculés par ART-2, associés à leur occurrence au cours de la trajectoire de Sgal dans l'eau et sous contrainte de distance.

(vi) Comparaison des résultats expérimentaux et théoriques pour Sgal dans l'eau

L'étude des zones de la carte de Ramachandran visitées par les résidus aux cours des DM permet de bien comparer le comportement théorique et expérimental puisque le spectre de dichroïsme circulaire ne dépend que de la position relative des résidus, c'est-à-dire de leur structure secondaire. Le spectre DCE du peptide Sgal démontre la présence d'une structure en polyproline II en équilibre avec une structure non ordonnée. Or la zone la plus explorée durant toutes nos DM est celle de la polyproline II ce qui démontre un parfait accord entre résultats expérimentaux et théoriques.

La structure non ordonnée en équilibre avec la polyproline II, dont la présence est suggérée par les spectres de DCE, est certainement un coude β sur **PSAQ** et / ou sur **SAQD** ou plus vraisemblablement sur **QDEA** seul, puisqu'il est présent tout au long des trajectoires de DM. De plus, sur l'ensemble des 60 ns de simulations de DM effectuées dans l'eau, le coude β le plus fréquent est bien celui sur **QDEA**, avec 45,6 % de présence. Le coude β dont il est le plus proche est le type I qui donne un signal caractéristique en DCE, mais la DM montre que le coude β n'est pas la structure majoritaire au cours du temps, et qu'il s'agit en outre d'un type I ou d'un type IV proche du type I qui n'est détectable en moyenne que dans 39,4 % des conformations. La présence d'un coude β de type I sur **QDEA** serait donc relativement transitoire et le signal dichroïque serait alors noyé par celui de la polyproline II.

La spectroscopie RMN n'a pas permis de détecter d'autres effets NOE séquentiels ou intra-résidus et le fait que les coefficients de température soient élevés étaye encore notre hypothèse de la présence d'un coude β flexible et faiblement présent comme semble l'être le coude β sur **QDEA**. La spectroscopie suggère comme structure la plus probable une conformation en polyproline II ce qui est en accord avec la spectroscopie de DCE, mais également avec nos résultats de dynamique moléculaire.

Dans le cas de la simulation de dynamique moléculaire dans l'eau sous contrainte de distance, la comparaison entre nos résultats théoriques et nos résultats expérimentaux est difficile. En effet, la contrainte de distance est artificiellement maintenue et empêche un bon reflet du comportement structural du peptide Sgal. Cependant, la zone la plus peuplée est celle de la polyproline II ce qui est en parfaite adéquation avec les résultats de RMN et de DCE qui mettaient en avant la présence de ce type de structure. De plus, certains coudes β (sur **VVGS** et **QDEA**) déjà présents dans les simulations dénuées de contrainte sont également détectables. Le but de cette simulation est de "mimer" l'influence, même de manière très

approximative, de l'EBP sur cette séquence. On se rend donc compte que le comportement du peptide Sgal n'est pas totalement différent que ce qu'il pourrait être au sein de l'EBP.

β. Simulations de dynamique moléculaire du peptide Sgal dans d'autres solvants.

(i) DM dans le cyclohexane du peptide Sgal (modèle MC de plus basse énergie)

Tout comme pour l'eau, nous avons voulu étudier la dynamique moléculaire du peptide Sgal, mais cette fois-ci dans un milieu apolaire, identique à celui utilisé pour les calculs de Monte-Carlo : le cyclohexane. Après avoir été plongé dans une boîte de cyclohexane préalablement équilibrée, le peptide Sgal a été chauffé et équilibré comme pour les simulations précédentes. Malheureusement, durant la phase de production, pratiquement aucun mouvement du peptide Sgal n'a pu être mise en évidence. Il est possible que la grande viscosité du modèle de cyclohexane empêche tout mouvement du peptide. La solution pour remédier à ce problème serait peut être de construire une boîte de solvant mixte cyclohexane-eau, mais l'étude aurait alors perdu de son intérêt à savoir, connaître le comportement du peptide Sgal dans le cyclohexane à partir du modèle obtenu en Monte-Carlo dans ce même type de solvant.

(ii) DM de 40 ns dans le TFE du peptide Sgal (modèle par homologie)

Seules des structures en coudes β ont pu être observées au sein des structures analysées.

	Types de coudes β (%)					
	Canoniques et déformés					**Total**
	I	**I'**	**II**	**II'**	**VIII**	
VVGS	0,2	–	1,4	–	–	1,6
VGSP	–	–	–	–	0,4	0,4
GSPS	–	–	–	–	–	–
SPSA	7,8	–	–	–	48,5	56,3
PSAQ	–	–	–	–	–	–
SAQD	23,6	–	–	–	5,0	28,6
AQDE	–	–	–	–	5,8	5,8
QDEA	0,2	–	0,4	–	3,3	8,3
DEAS	0,4	–	–	–	3,1	3,5
EASP	–	–	–	–	2,0	2,0
ASPL	–	–	–	–	–	–
SPLS	–	–	–	–	–	–

Tableau 22 : Pourcentages des types de coudes β de Sgal dans le TFE.

Pour **SPSA,** 56,8 % des conformations sont en coude β dont 48,5 % en type VIII et pour **SAQD,** 28, 6 % des conformations sont en coude β dont 23,6 % de type I.

Le suivi des coudes β au cours de la trajectoire (**Figure 84**) permet de montrer que la séquence **SPSA** est tout au long de la DM en coude β principalement de type VIII, avec aléatoirement un basculement en type I dont la présence est surtout localisée en fin de la trajectoire, c'est-à-dire à partir de 25 ns.

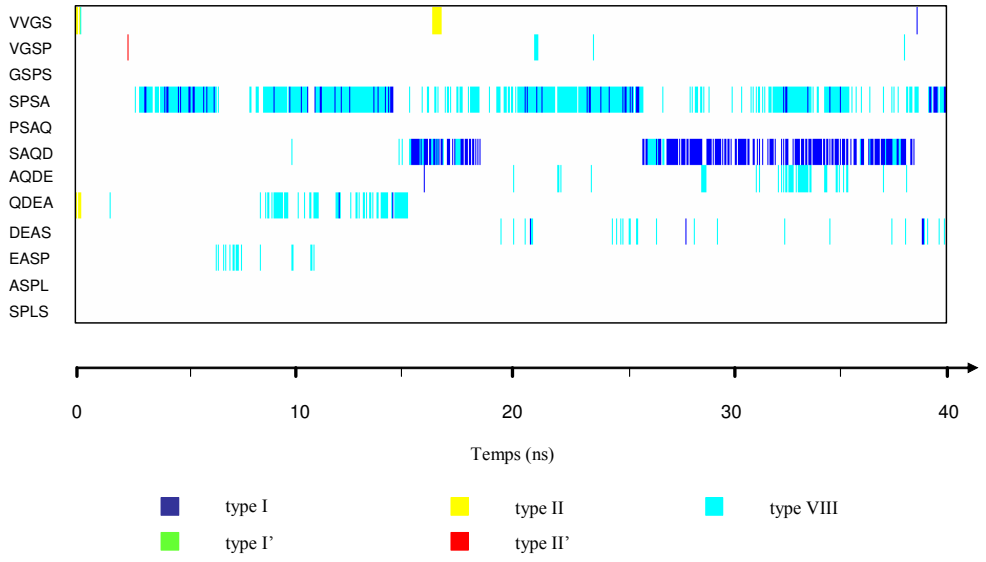


Figure 84 : Coudes β présents au cours du temps pour la DM de Sgal dans le TFE.

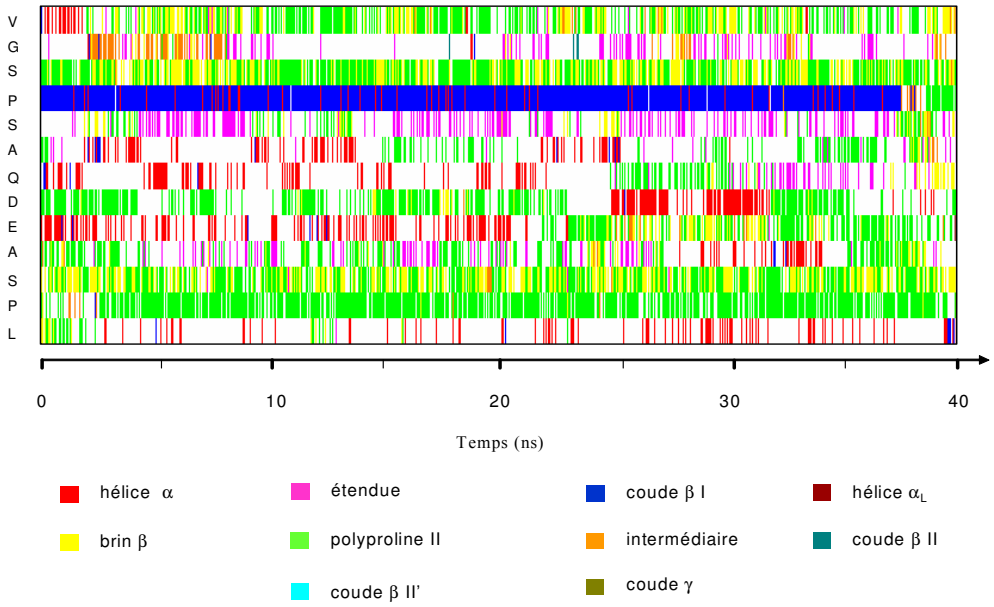


Figure 85 : Zones de la carte de Ramachandran occupées par les résidus de Sgal dans le TFE.

L'évolution des différentes zones de la carte de Ramachandran occupées par chacun des résidus est donnée ci-dessus **Figure 85**. Comme dans l'eau, la zone la plus peuplée par les résidus est celle de la polyroline II, suivie ensuite par celle des brins β, puis celle des coudes β de type I (**Tableau 23**).

Zones	Pourcentages	Zones	Pourcentages
hélice α	6,1	**intermédiaire**	1,2
brin β	9,3	**hélice α_G**	–
étendue	4,4	**coude β II**	–
polyproline II	23,7	**coude β II'**	–
coude β I	7,0	**coude γ**	–
	hors zones	48,1	

Tableau 23 : Pourcentages des zones de la carte de Ramachandran occupées par les résidus de Sgal dans le TFE; – : pourcentage inférieur à 0,1 %.

La carte des RMSD croisés met en évidence que durant les 20 premières nanosecondes, le peptide Sgal passe par différents amas de conformations proches, mais à partir de 20 ns, un, voire deux amas, sont visibles jusqu'à la fin de la simulation de dynamique moléculaire.

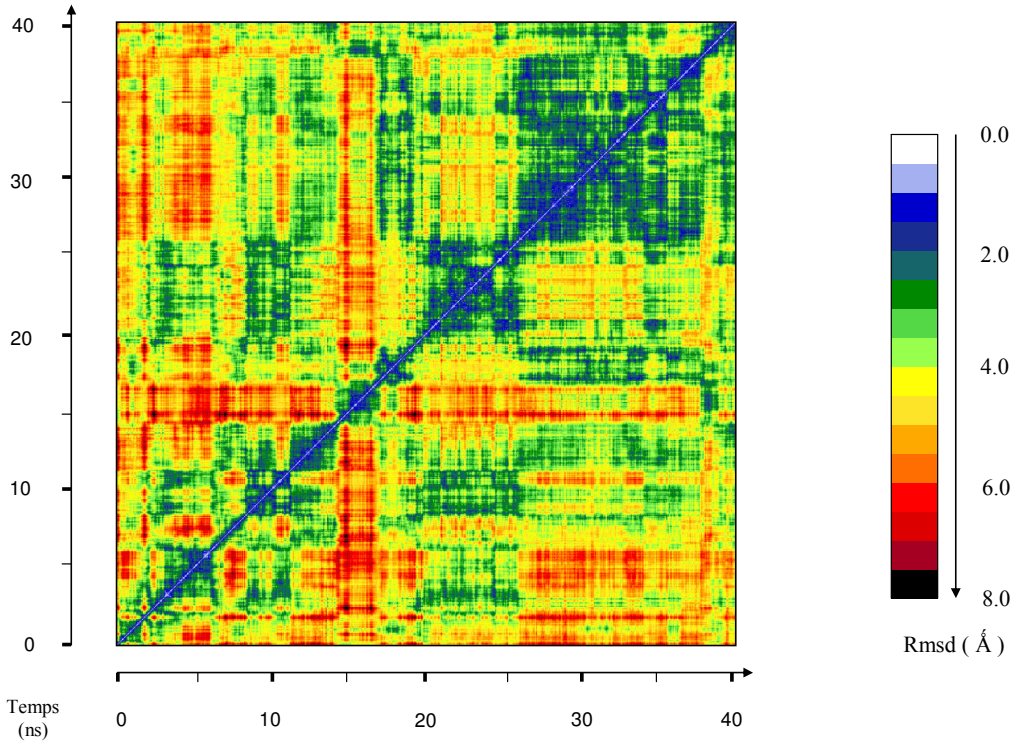


Figure 86 : Carte des RMSD croisés de Sgal dans le TFE.

Le calcul des amas de conformations proches par ART-2 avec un angle de troncature de 62° a donné 5 amas différents (**Figure 87**). Les 4 premiers amas de conformations proches se succèdent en moins de 5 ns, le quatrième se poursuivant jusqu'à 8,5 ns puis, le peptide Sgal retourne dans une conformation proche de l'amas 3 pendant 5 ns. Enfin l'amas 5, le plus peuplé, apparaît à partir de 20 ns jusqu'à la fin des 40 ns de trajectoire.

Les structures représentatives de chacun des amas sont données en **Figure 88** et les types de coude β qui composent ces structures sont reportés dans le **Tableau 24**.

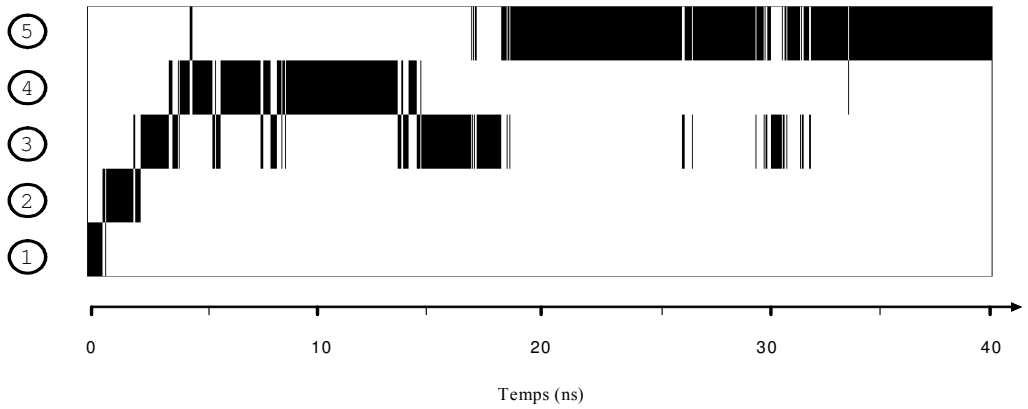


Figure 87 : Amas de conformations au cours du temps calculés par ART-2.

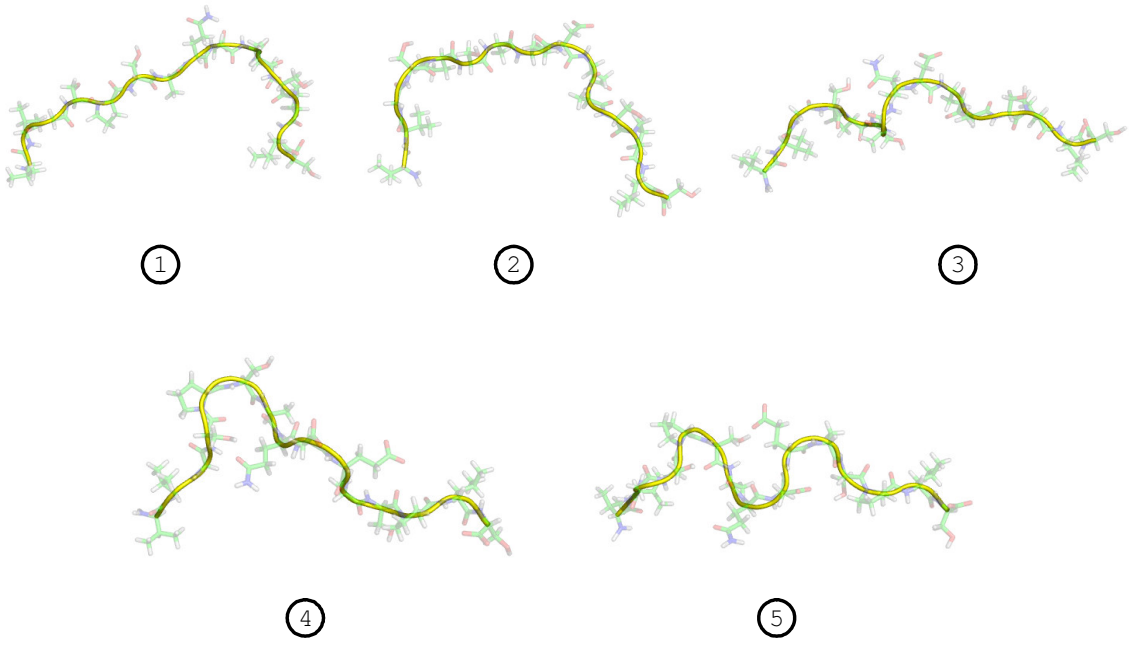


Figure 88 : Structures représentatives de chaque amas de conformations calculé par ART-2 au cours de la trajectoire de Sgal dans le TFE.

amas	Types de coudes β				Occurrences (%)
	S_4PSA_7	S_6AQD_9	Q_8DEA_{11}	D_9EAS_{12}	
1	-	-	IV (I)	-	1,7
2	-	-	-	-	4,1
3	IV (I)	IV (I)	-	-	17,8
4	IV	-	IV	-	24,1
5	IV (I)	IV (I)	-	IV (VIII)	52,3

Tableau 24 : Coudes β des amas de conformations calculés par ART-2, associés à leur occurrence au cours de la trajectoire de Sgal dans le TFE.

(iii) Comparaison des résultats expérimentaux et théoriques pour Sgal dans le TFE

Les spectres de DCE du peptide Sgal dans le TFE permettent de conclure que la structure de ce dernier est en coude β de type I plutôt qu'en hélice α.

Les données préliminaires RMN vont de plus dans le même sens car aucun effet caractéristique de l'hélice α n'est observable. Nos résultats de DM montrent que la structure majoritaire du peptide Sgal est effectivement un coude β de type I sur la séquence **SAQD**. Le nombre de conformations qui possèdent ce type de coude β sur **SAQD** n'est seulement que de 23,6 %. Mais, au vu des structures représentatives des amas 3 et 5, qui possèdent un coude β de type IV proche du type I et qui regroupent à eux deux environ 76 % des conformations, il est très probable que durant la simulation, le motif **SAQD** oscille autour d'une structure en coude β de type I en accord avec les spectres de DCE. Il est d'ailleurs intéressant de remarquer que les résultats préliminaires de RMN laissent à penser qu'un coude β de type I serait présent sur la séquence **SAQD**.

B.2.d. Conclusions

Les résultats expérimentaux de DCE et de RMN, portant sur la caractérisation structurale du peptide Sgal, ont montré la présence de structure polyproline en équilibre avec une conformation sans structure régulière. Cette conformation, au vu du modèle construit par homologie, des résultats de Monte-Carlo et des simulations de dynamique moléculaire obtenus, pourrait être caractérisée par un coude β de type I sur le motif **QDEA**.

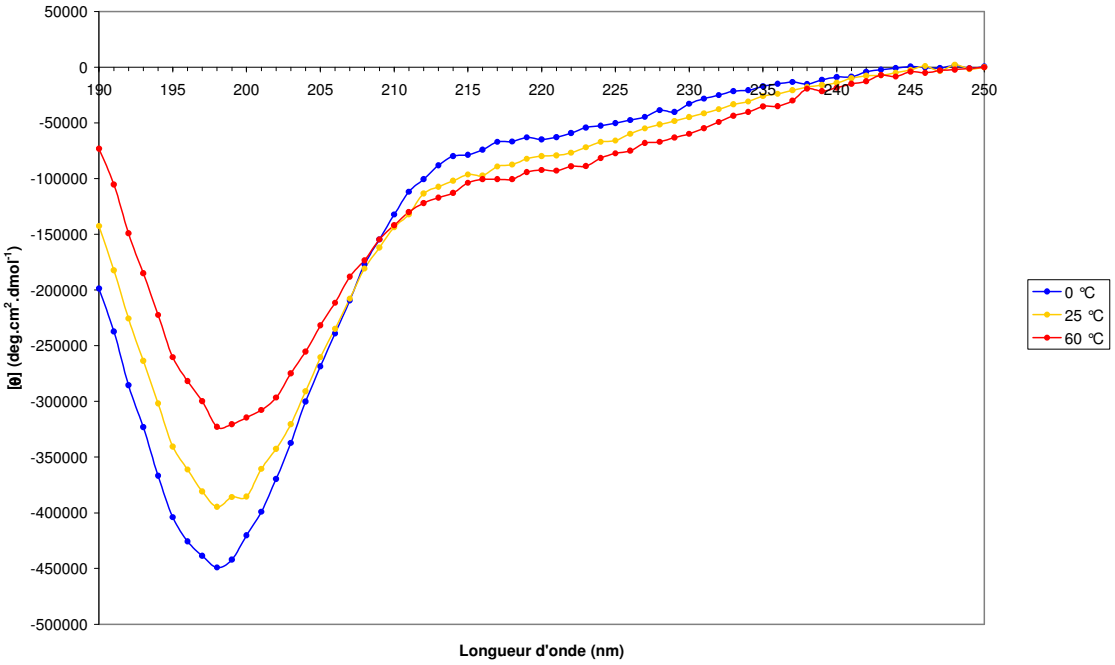
On peut donc conclure que ce coude β sur **QDEA** est certainement très important pour l'interaction entre l'EBP et ses ligands.

C. ETUDE DU PEPTIDE DE 32 RESIDUS DONT LA SEQUENCE EST SPECIFIQUE A L'EBP

L'étude du peptide 32 résidus, spécifique à l'EBP, a été menée en espérant que les résidus flanquant la séquence du peptide Sgal stabiliseraient une structure bien définie notamment grâce à la présence d'éventuelles structures secondaires.

C.1. Etude en spectroscopie de Dichroïsme Circulaire Electronique

Les spectres de DCE ont été enregistrés dans l'eau à 25° et 80° C (**Figure 89**).



<u>Figure 89</u> : Spectres de DCE du peptide de 32 résidus dans l'eau.

Le comportement du peptide de 32 résidus est similaire à celui du peptide Sgal. On remarque la présence d'un point isoelliptique à 208 nm indiquant que le peptide peut prendre 2 types de conformations. Une bande négative à 208 nm, dont l'intensité augmente avec la diminution de la température, associée à une autre petite bande vers 219 nm, qui tend à devenir positive avec la diminution de la température, laisse à penser que comme pour le peptide Sgal, le peptide de 32 résidus présente une conformation en polyproline II en équilibre avec une structure non ordonnée. Cependant, dans le cas du peptide de 32 résidus, à cause de la faible bande à 219 nm, on peut dire que l'équilibre est largement déplacé en faveur de la structure non ordonnée.

Pour analyser les spectres de DCE du peptide de 32 résidus, nous avons également utilisé le package CDPRO (**II.A.2.e.**). Ces programmes de traitements des spectres de DCE permettent de proposer quels sont les différents pourcentages de structures secondaires régulières qui composent une protéine à partir du spectre obtenu (**Tableau 25**).

T (°C)	Structures secondaires (%)						
	Hélice α	**Hélice 3_{10}**	**Brin β**	**coude β**	**PP II**	**Non ordonné**	**RMSD**
				ContinLL			
0 °C	-	3,3	-	21,5	12,8	62,4	0,015
25 °C	-	7,0	4,1	32,8	8,2	47,9	0,057
60 °C	-	4,1	4,8	24,2	4,1	62,8	0,090
				Selcon			
0 °C	1,6	4,3	21,9	14,5	11,7	46,0	0,773
25 °C	1,8	4,3	22,0	14,4	11,6	45,9	0,775
60 °C	-	3,6	23,3	12,7	11,7	48,7	0,490
				CDsstr			
0 °C	20,2	16,1	12,0	17,1	17,6	17,0	0,064
25 °C	22,0	16,4	9,9	16,9	17,9	16,9	0,071
60 °C	25,1	16,6	7,4	16,6	18,1	16,2	0,042

Tableau 25 : Pourcentages des structures secondaires estimées par les 3 programmes de CDPRO du peptide de 32 résidus dans l'eau pour différentes températures.

Le RMSD entre le spectre expérimental et le spectre reconstitué à partir des contributions spectrales des structures secondaires est, quelle que soit la méthode, toujours faible, indiquant ainsi un très bon accord entre ces 2 spectres. Les programmes Selcon et CDsstr proposent pratiquement les mêmes pourcentages de structures secondaires quelle que soit la température à laquelle le spectre a été enregistré. En revanche, le programme ContinLL, considéré comme le meilleur des 3, donne des pourcentages de polyproline II qui diminuent avec l'augmentation de la température, ce qui est le comportement habituel pour ce type de structure. Il faut rester très prudent avec ces résultats car, bien que la taille du peptide soit de 32 résidus, elle est sûrement insuffisante pour pouvoir comparer les spectres de DCE avec ceux des protéines globulaires à partir desquelles ont été développés les programmes utilisés. Dans ce cas l'analyse qualitative comparative visuelle reste encore une solution raisonnable.

C.2. Spectroscopies de Résonance Magnétique Nucléaire

Les spectres de RMN ont été enregistrés par Mademoiselle Elise PROST, au sein de l'Unité CNRS FRE2715, IFR53 de Reims, dirigée par le Dr Jean-Marc NUZILLARD.

L'acquisition des spectres COSY, TOCSY et NOESY a été effectuée sur un spectromètre Brücker Avance 500 MHz dans un mélange H_2O / TFE (90 % / 10 %), à une température de 25° C.

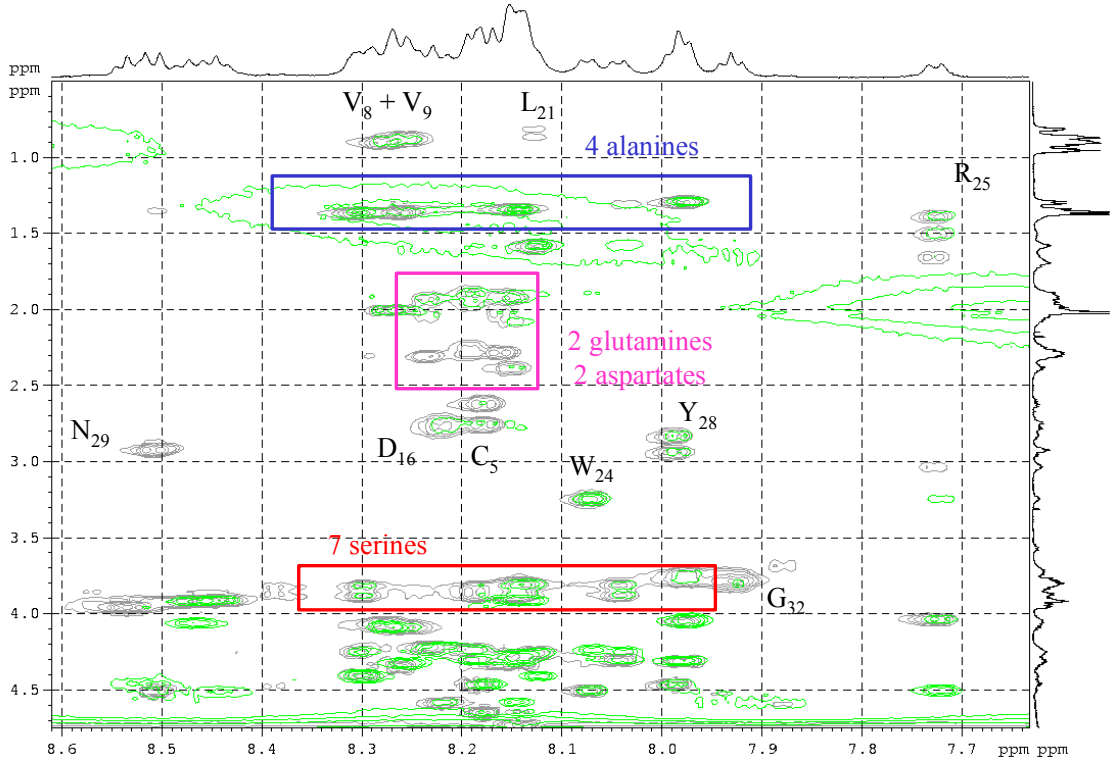


Figure 90 : Spectres COSY (vert) et TOCSY (gris) du peptide de 32 résidus.

Malheureusement, la complexité des spectres RMN obtenus pose des problèmes pour bien différencier les différents résidus (**Figure 90**).

Bien qu'il ne soit pas équivalent à une protéine, un peptide de 32 résidus a déjà une taille conséquente pour la RMN. Cependant, il subsiste encore trop d'incertitudes concernant les résidus à l'origine des différents effets RMN principalement parce que l'on n'arrive pas à distinguer les acides aminés de même nature (particulièrement les résidus Alanine et Sérine).

C.3. Conclusion / perspective

A partir des spectres de DCE, on peut dire que le peptide de 32 résidus présente une structure majoritairement non ordonnée, néanmoins, en équilibre avec une conformation en polyproline II.

La RMN n'a pas pu, pour l'instant, apporter d'informations structurales complémentaires.

La modélisation moléculaire est actuellement également incapable d'apporter des solutions alternatives à ces études expérimentales. En effet, aucune homologie de séquence suffisamment importante, c'est-à-dire supérieure à 20 %, avec une protéine de structure tridimensionnelle connue n'a été trouvée. Pour ce qui est des simulations de dynamique moléculaire, de Monte-Carlo ou de carte adiabatique, la taille relativement importante du peptide exclut tout calcul de ce type sans un minimum d'informations structurales.

Les données expérimentales de RMN restent alors indispensables pour restreindre l'espace conformationnel de ce peptide.

Deux solutions sont alors envisageables :

- marquage isotopique des atomes de carbone et d'azote, comme cela est habituellement fait pour réaliser l'étude RMN structurale de protéines.
- utilisation d'un spectromètre plus puissant de 800 MHz ou plus, ce qui pourrait faciliter l'attribution des différents effets RMN de chacun des résidus.

Partie IV

ETUDE des LIGANDS PEPTIDIQUES de L'EBP

La métaphore habituellement utilisée pour expliquer l'importance de la structure tridimensionnelle d'un récepteur et de son ligand est celle de la clé et de la serrure. En effet pour qu'une clé, le **ligand**, puisse entrer, c'est-à-dire interagir avec une serrure, le **récepteur**, il est nécessaire que la clé et la serrure aient des structures compatibles.

L'étude de la structure des ligands peptidiques de l'EBP est donc cruciale pour comprendre l'interaction qui a lieu, pour éventuellement pouvoir agir sur celle-ci. Nous avons restreint notre étude aux ligands peptidiques contenant une séquence **GXXP** (**X** étant n'importe quel résidu) pour lesquels de nombreuses études expérimentales ont montré leur capacité à se lier à l'EBP, conduisant à induire des modifications du comportement de la cellule (Senior et al. 1984; Grosso et Scott 1993a; Grosso et Scott 1993b; Bisaccia et al. 1994; Hauck et al. 1995; Castiglione Morelli et al. 1997; Brassart et al. 1998; Faury et al. 1998; Fulop et al. 1998; Lograno et al. 1998; Brassart et al. 2001).

Pour connaître la structure nécessaire à un ligand de l'EBP, nous avons tout d'abord étudié l'ensemble des séquences **GXXP**, **GXXPX**, **XGXXP** et **XGXXPX** de l'élastine humaine par simulation de dynamique moléculaire en solvant implicite.

Les peptides contenant la séquence **GXXP** et dont l'activité ou l'inactivité biologique a été mise en évidence ont ensuite été étudiés plus particulièrement par Monte-Carlo et carte adiabatique.

Des simulations de dynamique moléculaire en solvant explicite (eau et TFE) ont été également réalisées pour certains de ces peptides dérivant de l'élastine (**VGVAPG**, **PGAIPG** et **GGVPG**).

Nous avons également étudié par simulations de dynamique moléculaire, 2 peptides dérivant de la fibrilline-1, dont un contenant la séquence **GXXP**.

A. PEPTIDES DERIVANT DE L'ELASTINE HUMAINE

A.1. Etude structurale de l'ensemble des EDP contenant GXXP

A.1.a.. Peptides étudiés

La séquence de la tropoélastine humaine (Indik et al. 1987; Bashir et al. 1990), enregistrée sous la référence P15502 dans la banque de données Swiss-Prot (Bairoch et Apweiler 1997), contient 43 motifs **GXXP**. L'exon 26A n'étant que rarement exprimé et sa présence toujours associée à des pathologies tissulaires (Liptay et al. 1993; Bisaccia et al. 1998; Ostuni et al. 2002), nous avons décidé de ne pas en tenir compte, ce qui permet de faire apparaître un motif **GXXP** supplémentaire. Nous avons donc en tout 45 séquences **GXXP**, dont 25 non redondantes. Si l'on désire étudier l'influence d'un résidu situé avant et / ou après le motif **GXXP**, cela porte le nombre de peptides différents à étudier à 128 (**Figure 91**).

GGVP **GAIP** **GGVP**GGVFYPGAGLGALGGGALGP **GGKP** LKPVPGGLAGAGLGAGL **GAFP** V TFPGALVPGGVADAAAAYKAAKAGAGL **GGVP** GVGGLGVSAGAVVPQPGA **GVKP** **GKVP** GV GLP **GVYP** G **GVLP** GARFPGV **GVLP** GVPTGA **GVKP** KAPGVGGAFAGIP **GVGP** FG **GPQP** GVPLGYPIKAPKLPGGYGLPYTT **GKLP** Y **GYGP** GGVAGAAGKAGYPTGT **GVGP** Q AAAAAAAKAAAKFGAGAA **GVLP** GVGGAGVPGVP **GAIP** GIGGIAGVGTPAAAAAAAAAAKA AKYGAAA **GLVP** GGP **GFGP** GVVGVPGAGVPGVGVPGAGIPVVPGAGIPGAAVPGVVSPEAA AKAAAKAAKY **GARP** GVGV **GGIP** TYGVGA **GGFP** GFGVGV **GGIP** GVAGVPSV **GGVP** GV **GGVP** GV **GISP** EAQAAAAAKAAKYGAAGAGVLG **GLVP** GPQAAVPGVPGT **GGVP** GV GTPAAAAAKAAAKA AQF **GLVP** GV **GVAP** GV **GVAP** GV **GVAP** GV **GLAP** GV **GVAP** GV **GVAP** GV **GVAP** **GIGP** GGVAAAAKSAAKVAAKAQLRAAAGLGAGIPGLGVGVGVPGLGV GAGVPGLGVGAGVPGF **GA** ~~GADEGVRRSLSPELREGDPSSSQHLPSTPSSPR~~ **VP** GALAAAK AAKYGAAVPGVLGGLGALGGVGIPGGVV **GAGP** AAAAAAAKAAAKAAQFGLVGAAGLGGLGVG GLGVPGVGGL **GGIPP** AAAAKAAKYGAAGLGGVLGGA **GQFP** LGGVAARPGF **GLSP** IFPG GACLGKACGRKRK

Figure 91 : Séquence de la tropoélastine humaine, avec, en gras les motifs GXXP, en souligné les motifs GXXP qui apparaissent pour la première fois, la séquence barrée correspond à la séquence de l'exon 26A.

Par convention, nous avons appelé $\mathbf{X_2}$ et $\mathbf{X_3}$, les résidus centraux du motif **GXXP**, $\mathbf{X_1}$, le résidu situé avant et $\mathbf{X_4}$, celui placé après le motif **GXXP**. Nous avons donc étudié par simulations de dynamique moléculaire 26 tétrapeptides $\mathbf{GX_1X_2P}$, 65 pentapeptides (30 $\mathbf{GX_1X_2PX_4}$ et 35 $\mathbf{X_1GX_1X_2P}$) et 37 hexapeptides $\mathbf{X_1GX_1X_2PX_4}$, ce qui représente bien un total 128 peptides $\mathbf{(X_1)GX_1X_2P(X_4)}$ dérivant de la tropoélastine humaine.

A.1.b. Protocole de simulation de dynamique moléculaire :

Les simulations de dynamique moléculaire des 128 peptides ont été effectuées avec la version c27b3 du programme CHARMM (Brooks et al. 1983), et le jeu de paramètres "tout atome" 22 (MacKerell et al. 1998). Les calculs de DM ont été réalisés à une température de 300 K, dans des conditions de solvant implicite, avec une constante diélectrique de 80 pour mimer un environnement aqueux. Les interactions non liantes ont été traitées entre 10 et 12 Å, avec la fonction FORCE SWITCH pour les interactions électrostatiques et la fonction SCHIFT pour les interactions de van der Waals, comme il est recommandé pour ce type de solvant (Steinbach et Brooks 1994).

N'ayant pas d'information structurale suffisamment précise, les structures de départ de tous les EDP ont été construites "toutes étendues", c'est-à-dire que les angles dièdres ϕ et ψ ont été fixées à une valeur de 180°, excepté pour la proline pour laquelle ϕ = -76,1°. Les peptides sont sous forme zwitterion, les extrémités N- et C- terminales étant chargées (NH_3^+ et COO^-).

Pour éliminer les mauvais contacts entre les atomes, les EDP sont minimalisés avec 500 pas de l'algorithme SD et 500 pas de l'algorithme GC. La température du système est augmentée par pas de 3 K tous les 0,5 ps jusqu'à 300 K sur une période de 50 ps.

La phase d'équilibrage dû système se déroule en deux étapes consécutives de 40 ps. Dans la première étape, les vitesses des atomes sont attribuées tous les 50 pas suivant une fonction de distribution gaussienne des vitesses. Dans la deuxième étape, les vitesses des atomes sont modifiées, si besoin, de manière à ce que la température ne sorte pas d'une fenêtre de ± 10 K autour de 300 K.

Enfin, la phase de production a été choisie d'une durée de 3 ns, en accord avec des résultats précédents qui montraient que seulement 0.75 ns était nécessaire au peptide **VGVAPG** pour se replier en coude β de type VIII (Floquet et al. 2004). Durant cette phase, l'équation du mouvement de Newton a été résolue grâce à l'algorithme de Verlet, avec un pas d'intégration de 0,1 fs.

De façon à vérifier que nos résultats ne sont pas biaisés par le choix des conditions de simulation, nous avons effectué d'autres simulations dans d'autres conditions (**Tableau 27**) testées sur les EDPs dont l'activité ou l'inactivité biologique était connue expérimentalement (peptides actifs : **PGAIPG**, **VGVAPG**, **GVLPG** et **GVAPG**; peptide inactif : **GGVPG**).
Nous avons ainsi augmenté les temps de simulation, utilisé des extrémités non chargées, c'est-à-dire l'extrémité N-terminale acétylée (-NH-CO-CH_3) et l'extrémité C-terminale amidée

(-CO-NH_2), mais aussi modifié le rayon de coupure, en le traitant à l' "infini", c'est-à-dire largement supérieur à la taille maximale possible des EDPs (99 Å).

Nous avons également modifié le traitement du solvant, en prenant une constante diélectrique dépendante de la distance entre les atomes ($\varepsilon = f(r_{ij})$), mais également deux modèles de solvant implicite basés sur un continuum électrostatique "Generalized Born", implanté dans CHARMM dans les modules `GBORN` et `GBMV`.

Les simulations de dynamique moléculaire utilisant le module `GBORN` ont fait appel au même protocole précédemment décrit. Pour ces 2 types de méthodes, un rayon de troncature infini et des extrémités non chargées ont été utilisées. Cependant, le modèle de solvant `GBMV` n'ayant été implanté qu'à partir de la version c30b1 de CHARMM, nous avons donc eu recours à cette version pour toutes les simulations dans ce type de solvant.

De plus, la température a du être contrôlée durant l'étape de production à l'aide d'un "bain thermostaté" suivant les méthodes décrites par Nosé (Nosé 1984) et Hoover (Hoover 1985).

Enfin, pour avoir une bonne conservation de l'énergie totale du système, la surface du peptide a été mise à jour à chaque pas d'intégration. Aucun autre paramètre, hormis ceux-ci, n'a été modifié par rapport au protocole utilisé pour l'ensemble des 128 EDPs.

EDP	Conditions de simulation			
	Modèle de solvant implicite	Rayon de coupure	Extrémités chargées	Durées (ns)
128 $(X_1)GX_2X_3P(X_4)$	**$\varepsilon = 80$**	**oui**	**oui**	**3**
PGAIPG VGVAPG GLVPG GVAPG GGVPG	**$\varepsilon = 80$**	**oui**	**oui**	**50**
	$\varepsilon = f(r_{ij})$	**oui**	**oui**	**3**
	$\varepsilon = 80$	**non**	**non**	**3**
	GENBORN	**non**	**non**	**3**
	GBMV	**non**	**non**	**3**

Tableau 26 : Conditions des simulations de dynamique moléculaire.

A.1.c. Résultats

Pour chaque simulation de dynamique moléculaire réalisée, 3.000 conformations ont été dans un premier temps analysées grâce au programme STRIDE.

Aucune structure en hélice α ou brin β n'ayant été détectée, les structures des EDPs étaient alors, soit en random coil, soit en coudes β. STRIDE ne permettant pas de repérer la présence d'hélice polyproline II, nous avons donc utilisé un programme qui, à partir des séries temporelles des différents angles dièdres ϕ et ψ de trois résidus consécutifs, permet de mettre en évidence la présence d'hélice polyproline II. De même, nous avons utilisé le programme, précédemment décrit dans Matériels et Méthodes (**§ II.B.3.c.γ.)** pour analyser les coudes et préciser de quel type de coude β un coude β de type IV est le plus proche.

Nous allons, dans un premier temps, détailler la présence des structures en coudes β, puis celle en hélice polyproline II sur l'ensemble des 128 peptides étudiés.

α Coudes β sur $X_1GX_2X_3$

Comme le montre le **Tableau 27** ci-dessous, le motif **$X_1GX_2X_3$** a de grandes propensions à se replier en coudes β.

	Coudes β canoniques et déformés (% des conformations)					
	Type I	Type I'	Type II	Type II'	Type VIII	Total
Pentapeptides X_1 GX_2X_3P						
P GAIP	20,6			51,7	2,2	74,5
P GKVP	25,5			51,1	1,7	78,3
G GPQP	12,1			46,3	0,9	59,3
T GKLP	26,2			46,1	1,1	73,4
V GLAP	31,6			38,2	1,3	71,1
P GVGP	18,8			36,8	6,9	62,5
A GLVP	28,3			36,6	1,9	66,8
A GVKP	20,8			34,6	7,5	62,9
V GVLP	28,7			33,8	3,0	65,5
Y GARP	28,9			32,5	3,1	64,5
F GAVP	27,1			32,2	2,7	62,0
G GLVP	20,0			31,2	2,4	53,6
A GQFP	25,1			29,9	2,7	57,7
L GAFP	34,7			29,8	4,7	69,2
G GVLP	18,0			29,2	3,2	50,4
F GLSP	31,0			28,4	3,3	62,7
P GIGP	12,2			28,4	5,8	46,4
V GAGP	22,9			27,9	3,6	54,4
Y GYGP	20,9			27,2	1,7	49,8
F GLVP	26,1			26,3	5,2	57,6
A GVYP	15,9			24,8	7,9	48,6
P GFGP	17,1			23,4	3,3	43,8
V GISP	22,5			21,5	6,1	50,1
A GVLP	13,9			21,3	5,0	40,2
V GVAP	20,2			19,8	5,9	45,9
P GGVP	9,4	17,9	16,9	16,7	3,2	64,1
P GGKP	10,1	18,0	21,5	15,9	3,6	69,1
T GVGP	2,7			13,9	14,4	31,0
G GIPP				12,5	12,6	25,1
T GGVP	10,8	16,9	14,0	10,3	3,4	55,4
L GGVP	10,8	16,5	30,6	9,7	2,6	70,2
V GGVP	10,1	18,8	17,4	9,4	2,6	58,3
A GGFP	6,6	12,1	12,6	9,1	3,8	44,2
L GGIP	9,6	16,2	23,6	9,1	4,5	63,0
V GGIP	11,2	17,3	18,3	7,1	3,5	57,4

	Coudes β canoniques et déformés (% des conformations)					
	Type I	Type I'	Type II	Type II'	Type VIII	Total
Hexapeptides X_1 GX_2X_3P X_4						
P GAIP G	27,6			59,1	0,7	87,4
P GKVP G	27,7			54,8	1,0	83,5
F GVAP G	25,3			52,0	0,8	78,1
V GVAP G	24,5			50,9	1,8	77,2
V GVLP G	30,5			50,8	2,0	83,3
V GLAP G	32,9			47,9	0,5	81,3
Y GARP G	38,3			44,0	1,9	84,2
G GPQP G	12,1			40,7	0,7	53,5
A GLVP G	23,1			39,0	2,4	64,5
T GKLP Y	35,9			38,5	0,9	75,3
A GQFP L	30,1			37,9	1,8	69,8
A GVLP G	28,4			37,7	1,4	67,5
F GLVP G	36,1			37,5	2,1	75,7
G GLVP G	27,4			36,7	1,8	65,9
P GFGP G	25,9			36,3	1,6	63,8
F GLSP I	34,5			35,0	2,3	71,8
V GISP E	25,6			30,2	3,8	59,6
P GIGP G	14,8			29,3	7,1	51,2
P GVGP F	10,7			29,2	11,2	51,1
L GAFP A	31,9			29,0	3,0	63,9
A GVKP G	20,6			28,7	4,1	53,4
G GVLP G	15,2			27,5	3,9	46,6
A GVYP G	16,1			25,7	4,8	46,6
A GVKP K	14,3			25,1	4,2	43,6
V GAGP A	24,6			24,4	4,6	53,6
Y GYGP G	25,5			23,4	2,2	51,1
T GVGP Q	19,6			22,6	3,4	45,6
P GGKP L	12,0	22,2	20,6	19,4	2,7	76,9
P GGVP G	11,6	22,7	19,3	18,6	1,7	73,9
T GGVP G	9,7	16,8	13,0	14,9	3,1	57,5
V GGIP T	10,3	21,7	18,5	12,0	2,6	65,1
L GGVP G	12,1	17,7	23,4	11,9	2,3	67,4
L GGIP P	12,2	20,2	25,5	10,6	2,7	71,2
A GGFP G	8,7	12,0	16,4	10,0	3,1	50,2
V GGIP G	11,6	24,4	22,4	9,2	1,5	69,1
V GGVP G	13,7	21,9	26,7	8,0	1,9	72,2
G GIPP A				6,9	8,0	14,9

Tableau 27 : Pourcentages des différents types de coudes β sur la séquence $X_1GX_2X_3$ dans les peptides $X_1GX_2X_3P(X_4)$. Les glycines sont en rouge et les prolines en jaune. La séquence des 4 résidus étudiés est mise sur fond gris.

La grande majorité des coudes β observés sont de type IV. Nous avons donc appelé "coude β déformé" de type I, un coude β de type IV proche du type I.

Seul un pourcentage faible de coudes β canoniques sont observables, c'est-à-dire ayant des valeurs d'angles dièdres des résidus **G** et $\mathbf{X_2}$ adéquates permettant de classer le coude β dans un type bien défini (I, I', II, II' et VIII).

Le plus fort pourcentage de coudes β canoniques a été obtenu pour l'hexapeptide **GGPQ**PG avec 2,8 % de coudes β de type II'. Pour les autres peptides, plus de 93,0 % des coudes β sont des coudes de type IV. Les peptides qui ont en position $\mathbf{X_2}$ un résidu différent de la glycine se replient essentiellement en coude β déformé de type I, II' et parfois VIII, mais jamais en type I' ou II. A l'inverse, la glycine, en position $\mathbf{X_2}$, permet au motif $\mathbf{X_1GGX_2}$ de présenter de manière égale chacun des types de coudes β.

β. Coudes β sur $X_2X_3PX_4$:

Contrairement à la séquence $\mathbf{X_1GX_2X_3}$, le motif $\mathbf{X_2X_3PX_4}$ n'est jamais en coude β, excepté si le résidu $\mathbf{X_3}$ est une glycine (**Tableau 28**). En effet pour tout autre peptide, la distance C_α-C_α entre les résidus $\mathbf{X_2}$ et $\mathbf{X_4}$ est toujours supérieure à 7,0 Å. Lorsque le résidu $\mathbf{X_3}$ est une glycine, les coudes β sont majoritairement de types IV. Par exemple, le peptide PG**VGPF** possède 89,8 % de coudes β de type IV ou bien PG**FGPG**, qui présente 98,6 % de coudes β de type IV. En fait, ces peptides $(X_1)G\mathbf{X_2GPX_4}$ peuvent être en coude β déformé de type I, II' ou VIII ou alors en structure random coil, plutôt étendue comme le sont les autres peptides n'ayant pas une glycine comme résidu $\mathbf{X_3}$.

	Coudes β canoniques et déformés (% des conformations)					
	Type I	Type I'	Type II	Type II'	Type VIII	Total
Pentapeptides $GX_2X_3P\ X_4$						
GVGP F	22,5			30,5	2,4	55,4
GIGP G	7,5			26,7	2,4	36,6
GFGP G	15,1			26,2	3,1	44,4
GYGP G	13,6			24,0	3,7	41,3
GVGP Q	12,6			23,6	6,5	42,7
GAGP A	9,0			20,0	4,7	33,7
Hexapeptides $X_1\ GX_2X_3P\ X_4$						
P GVGP F	6,3			38,5	1,1	45,9
V GAGP A	10,1			19,0	3,7	32,8
T GVGP Q	5,3			18,5	5,4	29,2
P GIGP G	2,1			12,6	3,5	18,2
P GFGP G	14,2			12,5	7,8	34,5
Y GYGP G	7,7			11,6	7,3	26,6

Tableau 28 : Pourcentages des différents types de coudes β sur la séquence $X_2X_3PX_4$ dans les peptides $(X_1)GX_2X_3PX_4$. Les glycines sont en rouge et les prolines en jaune. La séquence de 4 résidus étudiés est mise sur fond gris.

χ Coudes β sur GX_2X_3P :

L'ensemble des pourcentages des différents types de coudes β obtenus après analyse des 128 motifs GX_2X_3P est reporté ci-dessous **Tableau 29**.

		Coudes β (% de conformations)					
		Coudes β canoniques et déformés					
		Type VIII canonique	Type VIII	Type I	Type II	Type II'	Total
Tetrapeptides GX_2X_3P							
GVKP		15,4	50,5				50,5
GISP		23,9	46,7				46,7
GKVP		13,6	46,6				46,6
GLAP		13,7	46,3				46,3
GLSP		13,8	46,2				46,2
GLVP		11,0	45,0				45,0
GAVP		14,8	44,4				44,4
GQFP		14,1	40,5				40,5
GAIP		6,8	36,0				36,0
GAFP		12,8	34,9				34,9
GARP		10,1	34,4				34,4
GKLP		10,4	31,7				31,7
GVLP		7,7	22,6				22,6
GVAP		5,6	18,3				18,3
G*QP		13,3	16,9				16,9
GGVP		1,1	10,7			7,3	18,0
GGKP		1,1	9,8			7,2	17,0
GAGP		1,2	8,8	5,2	2,7		16,7
GGIP		0,5	8,2			4,3	12,5
GFGP		0,7	6,5	5,8	2,8		15,1
GYGP		0,6	5,7	4,9	1,3		11,9
GGFP		0,5	5,5			2,8	8,3
GVYP		2,3	5,3				5,3
GIGP		1,1	5,0	2,9	5,7		13,6
GVGP		0,3	2,0	1,3	8,0		11,3
GI*P							0,0
Pentapeptides $X_1GX_2X_3P$							
PGAIP		33,5	76,5				76,5
TGKLP		20,6	66,1				66,1
FGLVP		24,4	59,9				59,9
VGLAP		16,3	57,2				57,2
PGKVP		11,4	56,7				56,7
AGLVP		14,8	56,1				56,1
VGVLP		20,5	55,5				55,5
FGAVP		14,5	49,2				49,2
GGLVP		14,6	48,8				48,8
LGAFP		15,3	47,4				47,4
GGVLP		15,3	45,5				45,5
AGQFP		13,8	44,9				44,9
FGLSP		11,1	44,8				44,8
AGVKP		7,8	44,3				44,3
YGARP		13,3	44,0				44,0
GG*QP		27,7	43,5				43,5
VGISP		16,7	33,0				33,0
AGVLP		14,8	30,8				30,8
VGVAP		11,6	30,6				30,6
AGVYP		12,0	30,5				30,5
VGGVP		4,5	24,4			24,5	48,9
PGGVP		3,1	24,0			13,4	37,4
LGGVP		2,0	20,6			17,6	38,2
PGGKP		2,7	19,8			13,2	33,0
YGYGP		1,1	18,5	4,5	2,0		25,0
TGGVP		2,0	18,4			12,3	30,7
VGGIP		2,1	17,8			13,9	31,7
LGGIP		1,8	17,3			11,9	29,2
PGVGP		2,5	16,7	5,5	10,1		32,3
VGAGP		2,9	15,1	4,5	3,1		22,7
PGIGP		2,9	15,0	5,9	5,9		26,8
PGFGP		1,4	12,7	7,6	3,5		23,8
AGGFP		1,1	11,6			7,1	18,7
TGVGP		0,5	3,5	1,0	14,6		19,1
GGI*P		0,0	0,0				0,0

			Coudes β (% de conformations)					
			Coudes β canoniques et déformés					
			Type VIII canonique	Type VIII	Type I	Type II	Type II'	Total
Pentapeptides $GX_2X_3P\ X_4$								
GLVPG		+	11,1	52,8				52,8
GAVPG			17,1	50,0				50,0
GKLPY			13,1	49,9				49,9
GAIPG			18,4	47,6				47,6
GVYPG			15,9	47,1				47,1
GQFPL			14,9	46,0				46,0
GLSPI			10,9	40,5				40,5
GARPG			13,9	37,5				37,5
GLAPG			9,0	37,4				37,4
GAFPA			9,7	30,8				30,8
GKVPG			8,8	29,7				29,7
GVKPK			7,0	29,7				29,7
GVLPG			8,2	24,3				24,3
GVAPG		+	6,1	20,1				20,1
GVKPG			6,2	14,6				14,6
GGIPP			1,8	12,6			12,5	25,1
GGIPG			1,7	11,2			8,0	19,2
GGIPT			1,1	10,7			7,6	18,3
GGKPL			1,0	10,5			5,3	15,8
GGVPG		-	0,6	10,3			6,7	17,0
GVGPF			0,4	9,6	5,5	3,2		18,3
GIGPG			0,8	9,3	8,2	4,7		22,2
GAGPA			0,5	7,9	8,0	3,8		19,7
GVGPQ			0,3	7,1	12,9	2,9		22,9
GGFPG			0,6	6,6			3,6	10,2
G*QPG			4,2	6,2				6,2
GFGPG			0,2	4,8	4,6	1,4		10,8
GISPE			2,4	4,6				4,6
GYGPG			0,2	2,9	5,8	2,3		11,0
GI*PA			0,0	0,0				0,0
Hexapeptides $X_1GX_2X_3P\ X_4$								
PGAIPG		+	26,2	77,2				77,2
TGKLPY			14,6	73,5				73,5
VGVLPG			22,8	73,2				73,2
AGLVPG			24,0	70,7				70,7
VGLAPG			19,4	69,0				69,0
PGKVPG			18,7	68,9				68,9
VGVAPG		+	22,4	67,6				67,6
AGVLPG			21,5	65,5				65,5
GGLVPG			21,8	62,5				62,5
FGAVPG			11,5	60,4				60,4
AGQFPL			18,6	60,0				60,0
FGLVPG			11,0	58,9				58,9
FGLSPI			15,5	57,3				57,3
YGARPG			18,1	56,6				56,6
LGAFPA			19,7	52,3				52,3
VGISPE			24,6	47,7				47,7
AGVKPG			12,0	41,8				41,8
GG*QPG			24,5	41,7				41,7
GGVLPG			15,0	37,7				37,7
AGVKPK			10,7	37,5				37,5
PGGVPG			8,8	37,0			26,8	63,8
AGVYPG			12,6	31,8				31,8
TGGVPG			4,3	26,8			18,8	45,6
PGGKPL			3,8	25,0			13,8	38,8
VGGIPG			3,4	24,6			23,3	47,9
VGGVPG			3,5	24,6			21,3	45,9
LGGIPP			3,1	23,7			19,8	43,5
LGGVPG			1,8	22,7			13,7	36,4
VGGIPT			2,6	21,6			16,9	38,5
PGFGPG			1,7	21,1	13,9	3,3		38,3
TGVGPQ			0,9	16,6	9,3	8,4		34,3
AGGFPG			2,0	16,2			12,5	28,7
PGIGPG			1,7	15,4	7,4	8,2		31,0
VGAGPA			1,1	14,2	6,0	4,6		24,8
PGVGPF			1,2	13,5	0,9	21,1		35,5
YGYGPG			0,6	13,5	11,2	3,3		28,0
GGI*PA			0,0	0,0				0,0

<u>Tableau 29</u> : Pourcentages des différents types de coudes β sur la séquence GX_2X_3P dans les peptides $(X_1)GX_2X_3P(X_4)$. Les glycines sont en rouge et les prolines en jaune. La séquence de 4 résidus étudiés est mise sur fond gris. L'activité (+) ou l'inactivité (-) biologique expérimentale est indiquée.

Comme pour les tableaux précédents, les pourcentages de coudes β canoniques et déformés pour chacun des types de coude β y sont détaillés. Pour simplifier le Tableau, puisque aucun type I' n'est détectable, la colonne correspondante a été supprimée. A l'inverse, du fait de son importance, il a été nécessaire de préciser quand le coude β de type VIII était canonique ou un type IV proche du type VIII (type VIII déformé).

Les résultats indiquent que le motif **GX_2X_3P** a une grande propension à se replier en coude β et plus particulièrement en type VIII. Le repliement en type VIII est favorisé par la proline car elle impose, de par sa nature particulière, au résidu la précédent des valeurs d'angles dièdres proches de celles du brin β (ϕ = - 120° et ψ = - 120°) et donc proches de celles nécessaires pour former un coude β de type VIII (MacArthur et Thornton 1991). En fait, il y a un équilibre entre les conformations repliées en coude β et les conformations non ordonnées plutôt étendues.

Habituellement, les coudes β sont stabilisés par une liaison hydrogène, au sein du squelette peptidique, entre le groupement carboxyle du premier résidu, ici la glycine, et le groupement amide du dernier résidu, ici la proline. Cette liaison hydrogène ne peut malheureusement pas exister dans ce motif **GX_2X_3P**, puisque la proline est le seul résidu à ne pas comporter de groupement amide. L'absence de la liaison hydrogène explique, en partie, la présence à la fois des conformations en coudes β et celles désordonnées.

(i) Influence des résidus centraux X_2 et X_3 sur le repliement en coudes β :

Deux résidus particuliers, lorsqu'ils se trouvent en position centrale **X_2** ou **X_3**, semblent avoir une importance prépondérante sur le repliement en coudes β du motif **GX_2X_3P** : la proline et la glycine.

La proline en position **X_2** a un effet stabilisateur sur la conformation en coude β de type VIII du motif **GX_2X_3P**. Les peptides **G<u>P</u>QP**, G**G<u>P</u>QP**, **G<u>P</u>QP**G et G**G<u>P</u>QP**G ont, pour trois d'entre eux, d'importants pourcentages de coudes β de type VIII canoniques: respectivement 13,3 %, 27,7 %, 4,2 % et 24,5 %. De plus, ces quatre peptides ont les plus hauts rapports nombre de coudes β de type VIII canonique sur nombre total de coudes β (respectivement 0,8; 0,6; 0,7 et 0,6). La proline est le seul résidu à être composé d'une chaîne latérale cyclique, un hétérocyclique pyrrolidine, qui lui impose des valeurs d'angles dièdres particulières.

D'après une étude portant sur l'importance de la proline dans les structure référencées dans la PDB : base de données Protein Data Bank (Berman et al. 2000), la proline est limitée à deux zones de la carte de Ramachandran : 44 % des prolines sont observables dans la zone

($\phi \approx -61°$; $\psi \approx -35°$) et 56 % des prolines dans la zone ($\phi \approx -65°$; $\psi \approx +150°$) (MacArthur et Thornton 1991). La première région est très proche de celle que doit prendre le résidu $\mathbf{X_2}$ ($\phi \approx -60°$; $\psi \approx -30°$) pour que le motif $\mathbf{GX_2X_3P}$ soit en coude β de type VIII. C'est donc la nature cyclique et très contrainte de la proline, lui imposant des valeurs d'angles dièdres favorable à celles nécessaires en position $\mathbf{X_2}$, qui explique son rôle stabilisant sur une conformation en coude β de type VIII.

Au contraire, lorsqu'une proline est située en position $\mathbf{X_3}$ du motif $\mathbf{GX_2X_3P}$, il n'y a plus aucun coude β, de quelque type que ce soit. Les peptides **GIPP**; G**GIPP**, **GIPP**A et G**GIPP**A adoptent une structure étendue tout au long des trajectoires de DM. Il a aussi été montré que deux prolines consécutives ne pouvaient se replier en coude β que si l'une des deux était une *cis*-proline ($\omega = 0°$). Puisque 90 % des prolines composant les protéines sont des *trans*-prolines (Fischer et al. 1994), les modèles de départ de toutes les prolines ont donc tous été construits dans leur forme *trans* ($\omega = +180°$). Pour obtenir un coude β, il aurait donc fallu qu'une des deux prolines subisse une isomérisation *trans-cis* dont la barrière d'énergie est de l'ordre de 20 $kcal.mol^{-1}$ (Wedemeyer et al. 2002). Cette barrière d'énergie d'isomérisation est malheureusement trop importante pour pouvoir être franchie à 300 K, c'est pour cette raison qu'aucun coude β n'a pu être observé sur le motif $\mathbf{GX_2X_3P}$ ayant une proline en position X_3.

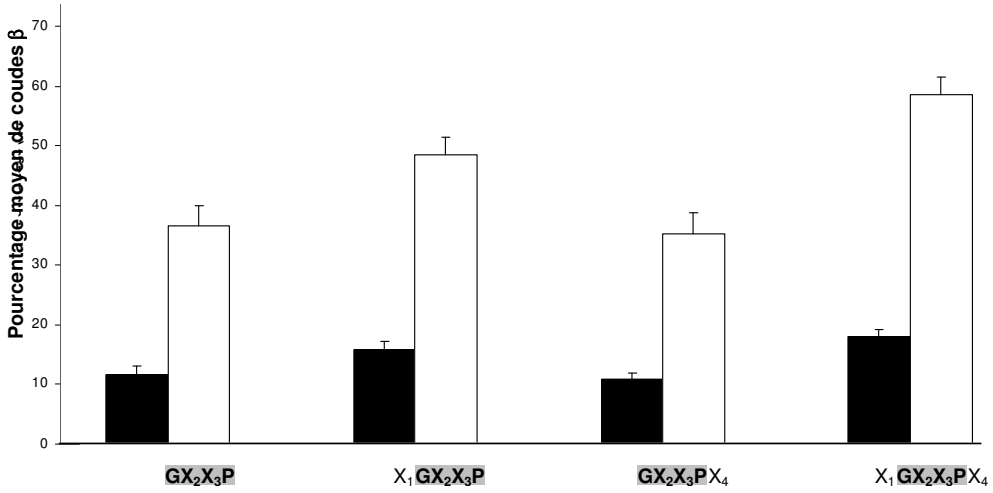
De même, une glycine en position centrale ($\mathbf{X_2}$ ou $\mathbf{X_3}$) diminue grandement la présence de coudes β de type VIII sur le motif $\mathbf{GX_2X_3P}$. Les peptides concernés par ces observations sont **GAGP**, **GGVP**, **GGKP**, **GIGP**, **GFGP**, **GGFP**, **GYGP**, **GVGP** et tous les autres peptides additionnés des résidus $\mathbf{X_1}$ et/ou $\mathbf{X_4}$ (voir plus loin **Figure 92**).

Le plus haut pourcentage de coudes β de type VIII canoniques observés sur le motif $\mathbf{GX_2X_3P}$, avec une glycine en position centrale, est de 8,8 % pour le peptide P**GGVP**G.

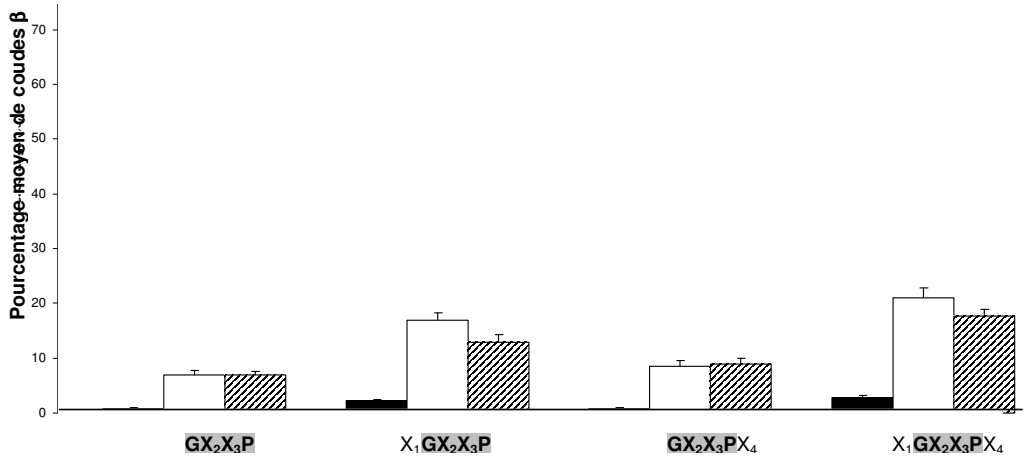
Par contre, d'autres types de coudes β sont observables suivant la position de la glycine. Ainsi, lorsque la glycine est en position $\mathbf{X_2}$, en plus du type VIII, le type II' est également détecté. Par contre, en position $\mathbf{X_3}$, la glycine induit le repliement en plusieurs types de coude β : VIII, I et II, mais jamais comme précédemment de type II'.

La glycine est le seul résidu capable d'explorer quasiment toute la carte de Ramachandran. Sa chaîne latérale n'étant constituée que d'un atome d'hydrogène, ses angles dièdres ϕ et ψ ne connaissent pratiquement pas de contrainte. Ainsi, en position centrale du motif $\mathbf{GX_2X_3P}$, la glycine a ce même comportement, ce qui explique que ses angles dièdres ne sont pas contraints dans la zone nécessaire à la formation d'un coude β de type VIII et que sa présence induit l'observation d'autres types de coudes β contrairement aux autres résidus.

(a)



(b)



<u>Figure 92</u> : Pourcentages moyens de conformations contenant un coude β sur le motif GX_2X_3P a) sans ou b) avec une glycine en position centrale (X_1 ou X_2).
Les coudes β de type VIII canoniques sont représentés en noir, les déformés (types IV proches du VIII) ajoutés à ceux de type VIII canonique en blanc ; les autres types de coudes β étant représentés en hachuré.
La barre représente la déviation standard à la moyenne (SEM).

Il est difficile de conclure quoique ce soit sur l'influence des autres résidus situés en position centrale du motif **GX_2X_3P**.

Aucun n'a apparemment d'influence très remarquable sur le repliement en coude β, mais pour bien étudier cela, il serait nécessaire de réaliser des séries de permutations des résidus de manière à tester tous les couples possibles de résidus centraux (20 x 20).

(ii) Influence des résidus X_1 et X_4 sur le repliement en coudes β :

La comparaison, entre les pourcentages de coudes β de type VIII observés pour les tétrapeptides **GX_2X_3P** et ceux obtenus pour les EDPs possédant un résidu **X_1** (pentapeptides $\underline{X_1}GX_2X_3P$ et hexapeptides $\underline{X_1}GX_2X_3PX_4$), montre qu'un résidu **X_1** augmente pratiquement de moitié les pourcentages de coudes β de type VIII sur le motif **GX_2X_3P** (**Tableau 29**) par rapport respectivement aux tétrapeptides GX_2X_3P et aux pentapeptides $GX_2X_3PX_4$ (**Figure 92)**. Cette observation est valable quel que soit le résidu **X_1** et même si les résidus centraux **X_3** et/ou **X_4** sont des glycines. Il semble par conséquent que ce soit seulement la présence du résidu, plutôt que la nature de sa chaîne latérale, qui exerce une influence favorisant la stabilisation du coude β de type VIII.

Il est difficile d'expliquer cette influence. Les peptides **$X_1GX_2X_3P$** et **$X_1GX_2X_3PX_4$** ont eux-mêmes d'importants pourcentages de coudes β sur la séquence **$X_1GX_2X_3$**, avec par exemple le pourcentage le plus élevé pour **PGAI**PG avec 87,4% des conformations en coudes β. Les coudes β observés sont de type IV, parce que la glycine de la séquence $X_1\underline{G}X_2X_3$ a ses valeurs d'angles dièdres qui fluctuent énormément ; les types les plus proches étant les types I, II' et plus faiblement le type VIII.

L'observation de types I et II' sur **$X_1GX_2X_3P$**(X_4) est logique puisque ce sont les seuls types dont les angles dièdres sont compatibles avec la présence d'un coude β de type VIII sur les quatre résidus suivants qui sont dans notre cas X_1**GX_2X_3P**(X_4).

Il est également intéressant de remarquer qu'en comparant les **Tableaux 27 et 29**, on se rend compte que les pourcentages de coudes β sur la séquence **$X_1GX_2X_3$** sont plus importants que ceux sur la séquence **GX_2X_3P**. Il est donc possible que les coudes β sur **$X_1GX_2X_3$** influent en faveur du repliement en coude β de type VIII du motif X_1**GX_2X_3P** et permette sa stabilisation par des effets hydrophobes ou électrostatiques.

En revanche, l'ajout d'un résidu **X_4** après le motif **GX_2X_3P**, et quelle que soit la nature de sa chaîne latérale, n'a aucune influence significative sur le repliement en coude β de type VIII de ce motif. Il a pourtant été démontré que la présence d'une glycine après le motif **GX_2X_3P**, c'est-à-dire en position **X_4**, était indispensable à l'activité biologique des EDPs (Brassart et al. 2001). La glycine ne semblant pas avoir un rôle structural, on peut donc penser que sa présence est indispensable au phénomène d'ancrage à l'EBP, en se fixant par exemple dans une poche dont seule la petite taille de la glycine puisse lui permettre de s'insérer.

δ Présence d'hélice polyproline II

En l'absence de connaissance de la tolérance dont les angles dièdres de l'hélice polyproline II peuvent faire preuve, nous avons décidé que pour la détecter, trois résidus consécutifs doivent avoir une valeur d'angle dièdre ϕ de - 75° ± 30° et ψ de - 145° ± 30°.

Il est ainsi impossible d'avoir trois résidus en hélice polyproline II sur **GX_2X_3P** car le premier et le dernier résidu ne possèdent qu'un seul angle dièdre. De plus, la glycine du motif **GX_2X_3P** possède extrêmement rarement (< 1% des structures) les valeurs d'angles dièdres caractéristiques de l'hélice polyproline II, engendrant ainsi l'absence d'hélice polyproline II sur la séquence de trois résidus **GX_2X_3** et par conséquence, sur la séquence de quatre résidus **GX_2X_3P**.

En fait, il n'est possible d'observer des conformations en hélice polyproline II que sur les séquences **X_2X_3P** et ce uniquement dans les séquences comprenant au moins un résidu de part et d'autre de **X_2X_3P** comme dans les pentapeptides **$GX_2X_3PX_4$** et les hexapeptides **$X_1GX_2X_3PX_4$** (**Tableau 30**).

Peptide $GX_2X_3PX_4$		Hélice PPII
G **IPP** A		23,8
G **PQP** G		20,6
G **ISP** E		16,6
G **VAP** G	+	12,7
G **KVP** G		10,8
G **VLP** G		10,8
G **VKP** K		10,2
G **LSP** I		6,1
G **VYP** G		5,9
G **LAP** G		5,6
G **AIP** G		4,4
G **KLP** Y		4,4
G **AFP** A		4,3
G **QFP** L		4,1
G **ARP** G		3,2
G **AVP** G		2,4
G **LVP** G	+	1,3
G **GIP** P		1,1
G **GIP** T		0,8
G **GVP** G	-	0,8
G **GFP** G		0,6
G **GIP** G		0,6
G **IGP** G		0,4
G **VGP** Q		0,3
G **FGP** G		0,2
G **GKP** L		0,1
G **YGP** G		0,1
G **AGP** A		0,0
G **VGP** F		0,0
G **VKP** G		0,0

Peptide $X_1GX_2X_3PX_4$		Hélice PPII
GG **IPP** A		20,7
GG **PQP** G		7,9
AG **VKP** K		5,8
GG **VLP** G		4,9
AG **VKP** G		4,2
AG **VYP** G		3,9
GG **LVP** G		3,7
VG **VAP** G	+	2,8
VG **ISP** E		2,3
LG **AFP** A		2,0
AG **VLP** G		1,7
FG **LSP** I		1,5
AG **QFP** L		1,2
FG **LVP** G		1,2
PG **KVP** G		1,1
AG **LVP** G		1,0
VG **VLP** G		0,9
TG **KLP** Y		0,8
VG **LAP** G		0,7
FG **VAP** G		0,5
AG **GFP** G		0,4
LG **GIP** P		0,3
PG **IGP** G		0,3
YG **ARP** G		0,3
TG **GVP** G		0,2
VG **GIP** G		0,2
LG **GVP** G		0,1
PG **AIP** G	+	0,1
VG **GIP** T		0,1
VG **GVP** G		0,1
PG **FGP** G		0,0
PG **GKP** L		0,0
PG **GVP** G		0,0
PG **VGP** F		0,0
TG **VGP** Q		0,0
VG **AGP** A		0,0
YG **YGP** G		0,0

<u>Tableau 30</u> : Pourcentages de conformations en hélice polyproline II observées après analyse des DM des pentapeptides $GX_2X_3PX_4$ et des hexapeptides $X_1GX_2X_3PX_4$. L'activité (+) ou l'inactivité (-) biologique expérimentale est indiquée.

D'une manière générale, d'après le **Tableau 30**, on se rend compte que les EDPs étudiés n'ont que de faibles pourcentages de conformations en hélice polyproline II. Ce résultat est dû au fait que ces EDP se replient préférentiellement en coudes β de type VIII ou IV proche du VIII, comme nous l'avons vu précédemment (Cf. **§ A.1.c.γ.**).

Cependant si les valeurs des angles dièdres du résidu $\mathbf{X_3}$ en coude β de type VIII (ϕ = - 120° ; ψ = + 120°) ou en hélice polyproline II (ϕ = - 75°; ψ = + 145°) sont relativement assez proches et peuvent permettre, grâce à la tolérance de ± 30°, que ce résidu soit classé simultanément dans ces deux sortes de conformations.

Ceci n'est absolument pas possible pour le résidu $\mathbf{X_2}$. En effet, les angles dièdres du résidu $\mathbf{X_2}$ en coude β de type VIII (ϕ = - 60° ; ψ = - 30°) et en hélice polyproline II (ϕ = - 75° ; ψ = + 145°), malgré la tolérance de ±30°, ne peuvent coexister dans une même conformation, à cause de l'angle ψ. En fait, la présence d'un coude β de type VIII ou IV proche du VIII sur $\mathbf{GX_2X_3P}$ interdit la présence d'un tour d'hélice polyproline II sur $\mathbf{X_2X_3P}$.

Ainsi, l'hélice polyproline II ne peut être observée qu'en l'absence de coude β de type VIII ou IV proche du VIII.

Le rôle déstabilisateur de la glycine est à nouveau confirmé puisque sa présence, en position $\mathbf{X_2}$ ou $\mathbf{X_3}$, engendre de faibles taux d'hélices polyproline II : le peptide G**GIP**G possède le plus grand pourcentage avec seulement 1,1 % des conformations en hélice polyproline II.

A l'inverse, la proline stabilise la présence d'hélice polyproline II, car ce sont les peptides qui en possèdent une en position $\mathbf{X_2}$ ou $\mathbf{X_3}$ qui ont les plus hauts pourcentages de sa présence : GG**IPP**A, G**IPP**A, GG**PQP**G et G**PQP**G. La proline explorant quasi exclusivement deux zones de la carte de Ramachandran (MacArthur et Thornton 1991), la principale qui regroupe environ 56 % des prolines, correspond à celle de la polyproline II ($\phi \approx$ - 65° ; $\psi \approx$ + 150°). De plus, il a été démontré que la proline forçait le résidu la précédant à rester dans une zone de la carte de Ramachandran correspondant à celle du brin β et de la polyproline II (MacArthur et Thornton 1991). Ces deux caractéristiques structurales de la proline permettent donc d'expliquer son rôle stabilisateur dans la formation d'une l'hélice polyproline II.

Le pourcentage moyen d'hélice polyproline II sur la séquence $\mathbf{X_2X_3P}$ est de 1,9 % pour les hexapeptides $\mathbf{X_1GX_2X_3PX_4}$, alors qu'il est de 5,1 % pour les pentapeptides $\mathbf{GX_2X_3PX_4}$. Cette différence est due au fait que la présence du résidu $\mathbf{X_1}$ augmente la stabilité du coude β de type VIII et diminue donc celle de l'hélice polyproline II.

On peut remarquer également que les EDPs biologiquement actifs (**PGAIPG**, **VGVAPG**, **GLVPG** et **GVAPG**) ne présente pas de différence de pourcentage significatif avec **GGVPG** qui est le seul EDP démontré comme étant inactif. L'hélice polyproline II n'est donc pas la structure qui permet à un peptide d'interagir avec l'EBP.

Dans le cadre de l'étude de la structure de la tropoélastine par l'intermédiaire de certaines séquences répétées ou de certains exons, il a été montré expérimentalement que la structure de ces peptides présentait un mélange de conformations en polyproline II, en coudes β et en structure non ordonnée (Tamburro et al. 2003; Bochicchio et al. 2004a; Bochicchio et al. 2004b). Certaines de ces séquences étudiées contenaient un motif $\mathbf{GX_2X_3P}$. Nos résultats théoriques qui montrent la quasi-absence de structures de type polyproline II, ne sont pas en contradiction avec les résultats de ces travaux. En effet, les tailles des peptides concernés sont très différentes : les motifs $\mathbf{GX_2X_3P}$ étaient inclus dans des peptides relativement longs, allant de 15 à 55 résidus, contrairement aux nôtres qui ne sont constitués que de 5 résidus. L'enchaînement et la nature des résidus des peptides expliquent ces différences de comportement structural. Il est, de plus, probable qu'un hexapeptide est de taille trop petite pour avoir une conformation en hélice polyproline II parfaitement stable.

ε. Relations structure - activité biologique

Parmi les 128 EDP que nous avons étudiés, 17 $\mathbf{GX_2X_3PG}$ et 24 $\mathbf{X_1GX_2X_3PG}$ possèdent une glycine après le motif $\mathbf{GX_2X_3P}$, c'est-à-dire en position $\mathbf{X_4}$, qui semble être indispensable pour que le peptide soit actif biologiquement (Brassart et al. 2001). Cependant, seuls quatre de ces peptides ont été étudiés dans la littérature, et ont montré une activité biologique : **VGVAPG** (Senior et al. 1984), **PGAIPG** (Grosso et Scott. 1993a), **GVLPG** et **GVAPG** (Bisaccia et al. 1994). Un seul est connu pour ne pas avoir d'effet biologique : **GGVPG** (Castiglione-Morelli et al. 1997).

Sur la base de nos résultats de DM, nous avons pu corréler la présence d'un coude β de type VIII avec l'activité biologique de l'EDP. Pour confirmer ces résultats, nous avons effectué des DM supplémentaires (**Tableau 31**). Les résultats obtenus montrent également une corrélation entre une forte présence de coudes β de type VIII et l'activité biologique des EDPs. De même, la conformation en hélice polyproline II est toujours à peine détectable avec des pourcentages très faibles. Ces observations sont valables quel que soit le type de solvant implicite utilisé. On peut remarquer également que les extrémités chargées N- et C-terminales ne sont pas

responsables du repliement en coudes β puisque les pourcentages de coudes β obtenus sont très proches de ceux obtenus avec des extrémités neutres. Enfin, il n'y a pas de changements notables des pourcentages de coudes β obtenus lorsque l'on augmente la durée de la DM de 3 à 50 ns, ce qui laisse penser qu'un temps de simulation de 3 ns est suffisant dans ces conditions pour étudier ce type de peptides.

EDP	Conditions de simulation			Coudes β sur GX_2X_3P (% des conformations)				PPII sur X_2X_3P (% des conformations)
	Modèle de solvant implicite	Rayon de coupure extrémités chargées	Temps (ns)	Type II' canonique et déformé	Type VIII canonique	Type VIII canonique et déformé	Total	
P GAIP G +	$\varepsilon_{=80}$	y	3		26.2	77.2	77.2	0.1
	$\varepsilon_{= f(r)}$	y	3		20.8	72.3	72.3	0.4
	$\varepsilon_{=80}$	y	50		27.1	77.8	77.8	0.4
	$\varepsilon_{=80}$	n	3		36.2	87.0	87.0	0.1
	GENBORN	n	3		36.6	87.7	87.7	0.3
	GBMV	n	3		35.7	87.7	87.7	0.4
V GVAP G +	$\varepsilon_{=80}$	y	3		22.4	67.6	67.6	2.8
	$\varepsilon_{= f(r)}$	y	3		19.8	68.9	68.9	4.8
	$\varepsilon_{=80}$	y	50		20.1	64.8	64.8	2.4
	$\varepsilon_{=80}$	n	3		22.1	64.9	64.9	2.4
	GENBORN	n	3		21.7	72.5	72.5	1.0
	GBMV	n	3		22.4	63.5	63.5	3.3
GLVP G +	$\varepsilon_{=80}$	y	3		11.1	52.8	52.8	1.3
	$\varepsilon_{= f(r)}$	y	3		10.4	40.5	40.5	6.8
	$\varepsilon_{=80}$	y	50		12.3	47.8	47.8	4.1
	$\varepsilon_{=80}$	n	3		20.1	59.7	59.7	2.0
	GENBORN	n	3		18.1	57.0	57.0	2.8
	GBMV	n	3		18.0	49.4	49.4	0.5
GVAP G +	$\varepsilon_{=80}$	y	3		6.1	20.1	20.1	12.7
	$\varepsilon_{= f(r)}$	y	3		13.1	42.4	42.4	1.8
	$\varepsilon_{=80}$	y	50		12.2	40.2	40.2	7.1
	$\varepsilon_{=80}$	n	3		19.7	57.8	57.8	4.6
	GENBORN	n	3		18.1	57.0	57.0	5.2
	GBMV	n	3		18.0	49.4	49.4	5.1
GGVP G -	$\varepsilon_{=80}$	y	3	6.7	0.6	10.3	17.0	0.8
	$\varepsilon_{= f(r)}$	y	3	11.2	1.8	12.1	23.3	0.8
	$\varepsilon_{=80}$	y	50	6.1	1.0	10.4	16.5	0.6
	$\varepsilon_{=80}$	n	3	10.6	2.0	17.2	27.8	0.5
	GENBORN	n	3	12.2	2.1	20.2	32.4	0.3
	GBMV	n	3	19.1	4.1	23.9	43.0	0.1

Tableau 31: Pourcentages de coudes β observés sur le motif GX_2X_3P et d'hélices polyproline II observés sur X_2X_3P obtenus dans différentes conditions de simulation. L'activité (+) ou l'inactivité (-) biologique expérimentale est indiquée.

ϕ *Conclusions*

Cette étude par simulation de dynamique moléculaire de 128 peptides dérivant de l'élastine et contenant le motif **GX_2X_3P** a permis de mieux comprendre leur comportement structural en termes de coudes β et d'hélices polyproline II.

Très peu de conformations en structure polyproline II ont pu être mises en évidence sur la séquence **X_2X_3P**, excepté si une proline est présente en position centrale **X_2** ou **X_3**. Cette structure n'est observée que transitoirement au cours des simulations, très certainement à cause de la petite taille des peptides étudiés qui ne permet pas de la stabiliser. De ce fait, on peut logiquement penser que la polyproline II n'est pas la structure permettant la fixation à l'EBP et n'est donc pas à l'origine de l'activité biologique de ces EDP.

Les conformations prises par les 128 EDPs sont soit en structure non ordonnée, soit en coude β. Plus précisément, le motif **GX_2X_3P** se replie essentiellement, au cours des DM, en coude β de type VIII, si aucun des résidus centraux n'est une glycine ou une proline en position **X_3**. Ce résultat est en parfait accord avec l'hypothèse, émise au sein de notre laboratoire, que c'est le coude β de type VIII qui serait responsable de l'activité biologique de ces peptides (Brassart et al. 2001; Fuchs et al. 2001). La glycine en position centrale du motif **GX_2X_3P** est à l'origine du très faible taux de coude β de type VIII obtenu. De plus elle induit une plus grande instabilité que les autres résidus, provoquant ainsi la formation d'autre types de coudes β (I, II et II'). Elle peut aussi, à cause de la plus grande entropie qu'elle génère, diminuer les chances d'avoir la bonne conformation pour se fixer à l'EBP.

La présence d'un coude β de type VIII peut donc être corrélé à l'activité biologique des peptides actifs **PGAIPG**, **VGVAPG**, **GLVPG** et **GVAPG**, qui ont de hauts pourcentages élevés de coudes β de type VIII, alors que **GGVPG**, dont l'absence d'effet biologique a été démontrée, n'a qu'un très faible taux de coudes β de type VIII.

Nous avons pu également démontrer que les coudes β sur le motif **GX_2X_3P** sont favorisés par la présence d'un résidu **X_1** situé avant ce motif. Aucune étude expérimentale n'a encore, pour l'instant confirmé ce résultat.

Par contre, un résidu, **X_4**, situé après ce motif n'a aucune influence sur le repliement du motif **GX_2X_3P** en coude β. Il avait été démontré qu'une glycine après le motif **GX_2X_3P** était indispensable pour induire une activité biologique (Brassart et al. 2001). D'après nos résultats, cette glycine ne semble pas avoir de rôle structural, mais peut donc jouer un rôle dans la fixation de l'EDP sur l'EBP.

Ce travail se place dans l'étude de la relation structure - activité de l'élastine et des peptides provenant de sa dégradation. Pour le moment, 5 EDPs contenant la séquence **GX_2X_3P** ont vu leur activité biologique étudiée et deux seulement parmi ceux-ci ont été observés réellement après la dégradation de l'élastine dans la MEC (**VGVAPG** et **GVLPG**). Parmi les EDPs étudiés dans ce travail, 15 hexapeptides de séquence **$X_1GX_2X_3PG$** et 10 pentapeptides **GX_2X_3PG** ont de forts pourcentages de coudes β de type VIII sur le motif **GX_2X_3P**, ce qui en fait des bons candidats pour tester *in vitro* leur capacité à se fixer à l'EBP.

A.2. Etude structurale des EDPs d'activité biologique connue

A.2.a. Etude par simulation de Monte-Carlo

α. Protocole

Les calculs de Monte-Carlo ont été effectués avec la version 4.0 du programme FANTOM (Schaumann et al. 1990). Tous les peptides ont été construits avec les extrémités N- et C-terminales chargées. Le protocole a été adapté à la taille des peptides : seule la structure de plus basse énergie à la fin de chaque cycle servira de modèle de départ pour le cycle suivant. Pour chaque cycle, si le modèle de plus basse énergie obtenu a une énergie plus importante que le modèle de départ, alors, le cycle suivant voit son nombre d'angles dièdres modifiables diminuer (5, 3, 1 puis encore 1 mais avec non plus 90 pas de minimalisation, mais 120). Pour étudier le comportement structural des peptides, les solvants utilisés sont l'eau et le cyclohexane pour mimer un environnement apolaire semblable à celui au voisinage de l'EBP. Les peptides étudiés, contenant un motif **GXXP**, ont été choisis en fonction de leur activité (**VGVAPG**, **PGAIPG**, **GLVPG** et **GVAPG**) ou inactivité biologique (**GGVPG**) ainsi qu'un peptide dérivant de la laminine capable de se fixer à l'EBP et d'induire expérimentalement des effets biologiques (Liotta et al. 1986).

β. Résultats

L'étude, pour chacun des peptides d'activité biologique connue, a nécessité de 6 à 9 cycles de 5.000 tirages pour un total de 10.000 à 23.000 structures acceptées selon les critères définis.

Dans **l'eau**, les structures de plus basses énergies des hexapeptides actifs sont caractérisées par un coude β de type IV sur **$X_1GX_2X_3$** et un coude β de type VIII ou IV proche du type VIII sur **GX_2X_3P** (voir plus loin **Tableau 32**). Aucun coude β n'est visible sur la séquence **X_2X_3PG**. Les structures obtenues sont toutes stabilisées par une liaison hydrogène entre le

groupement amide de $\mathbf{G_2}$ et le groupement carboxyle du résidu $\mathbf{X_3}$. (**Figure 93**). Un coude β de type IV est également présent sur la séquence **XGXX**.

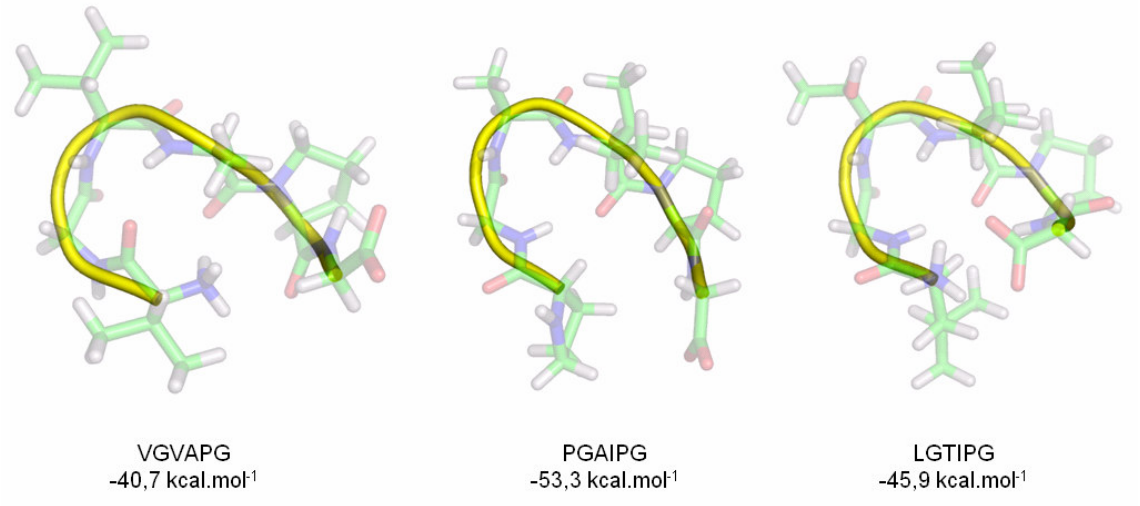


Figure 93 : Structures MC dans l'eau d'hexapeptides actifs.

Dans le **cyclohexane**, environnement apolaire, les structures de plus basse énergie des hexapeptides actifs présente une conformation légèrement différente de celles obtenues dans l'eau. La séquence $\mathbf{X_1GX_2X_3}$ est trouvée en coude β de type II' et est stabilisée par une liaison hydrogène entre le groupement amide de $\mathbf{X_1}$ et le groupement carboxyle de $\mathbf{X_3}$. Le motif $\mathbf{GX_2X_3P}$ est toujours en coude β de type VIII ou de type IV proche du VIII. Aucun des motifs $\mathbf{X_2X_3PG}$ n'est replié en coude β (**Tableau 32 et Figure 94**). Comme pour l'eau, un coude β de type IV sur **XGXX** est observable sur les motifs **XGXX.**

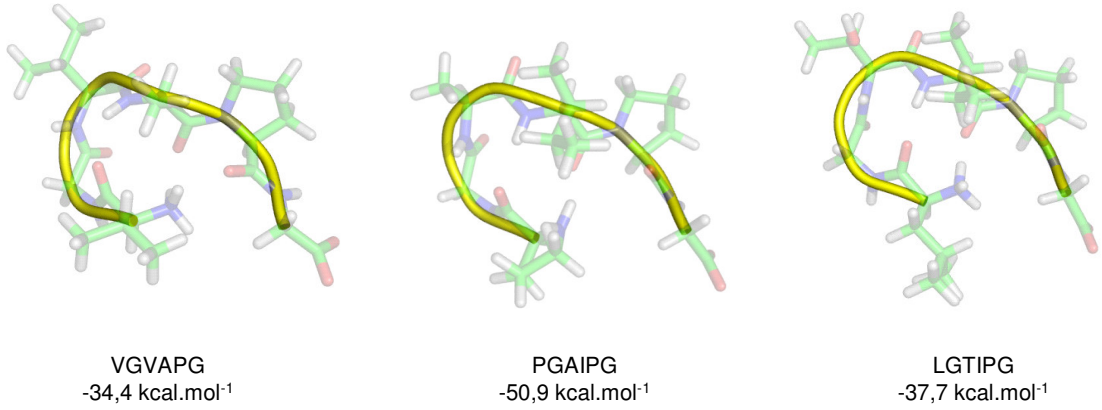
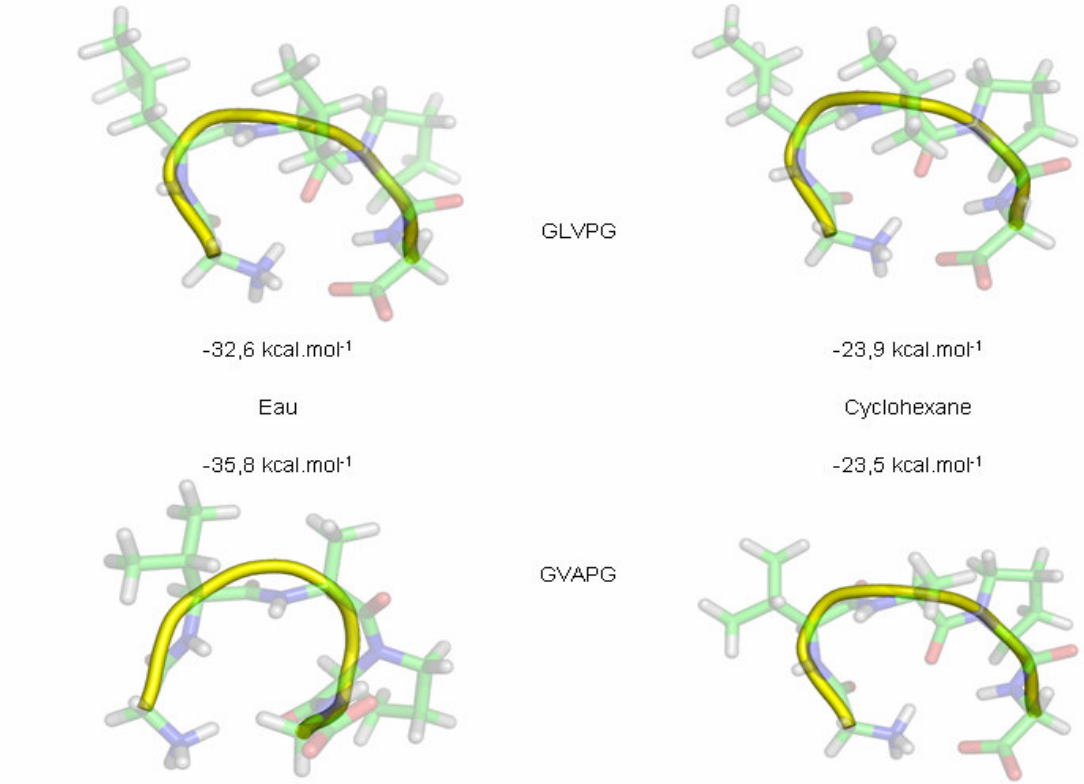
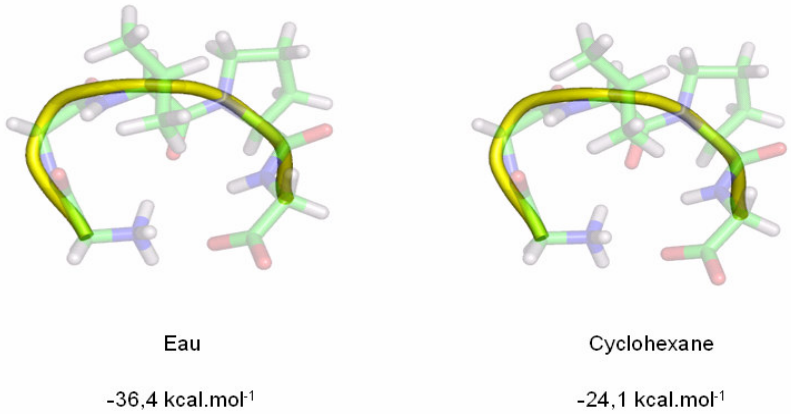


Figure 94 : Structures MC dans le cyclohexane d'hexapeptides actifs.

Les pentapeptides actifs **GLVPG** et **GVAPG** sont caractérisés par la présence d'un coude β de type IV, proche du VIII sur **GX_2X_3P** et par l'absence de coude β sur **X_2X_3PG** dans l'**eau** ou le **cyclohexane** (**Figure 95 et Tableau 32**). Hormis, la structure de plus basse énergie obtenue dans l'eau du peptide **GVAPG**, les 3 autres structures sont stabilisées par un pont salin entre les extrémités N- et C-terminales.



<u>Figure 95</u> : Structures MC dans l'eau et le cyclohexane des pentapeptides GLVPG et GVAPG.



<u>Figure 96</u> : Structures MC dans l'eau et le cyclohexane du peptide inactif GGVPG.

Peptides	Activité	Solvant	Types de coudes β			Energie (kcal.mol^{-1})
			$X_1GX_2X_3$	GX_2X_3P	X_2X_3PG	
VGVAPG	+	**Eau**	**IV**	**IV (VIII)**	-	**- 40,7**
		Cyclohexane	**IV**	**IV (VIII)**	-	**- 34,4**
PGAIPG	+	**Eau**	**IV**	**VIII**	-	**- 53,3**
		Cyclohexane	**IV**	**VIII**	-	**- 50,9**
LGTIPG	+	**Eau**	**IV**	**VIII**	-	**- 45,9**
		Cyclohexane	**IV**	**IV (VIII)**	-	**- 37,7**
GLVPG	+	**Eau**		**IV (VIII)**	-	**- 32,6**
		Cyclohexane		**IV (VIII)**	-	
GVAPG	+	**Eau**		**IV (VIII)**	-	**- 35,8**
		Cyclohexane		**IV (VIII)**	-	**- 23,5**
GGVPG	-	**Eau**		**IV (II')**	-	**- 36,4**
		Cyclohexane		**IV (II')**	-	**- 24,1**

<u>Tableau 32</u> : Types de coudes β des structures MC des peptides contenant GXXP. L'activité (+) ou l'inactivité (-) biologique est indiquée. (en gris : absence de la séquence indiquée dans le peptide).

χ Conclusions

Les peptides actifs présentent tous dans l'eau et le cyclohexane un coude β de type VIII ou IV proche du type VIII sur le motif **GXXP**. Le seul peptide inactif, présente un coude β non plus de type VIII, mais de type IV proche du II'. La différence d'activité entre les peptides peut ainsi s'expliquer par la différence de repliement du motif **GXXP**.

Les résultats des calculs de Monte-Carlo confirment ainsi les résultats des simulations de dynamique moléculaire précédents : le coude β de type VIII sur **GXXP**, ainsi que l'absence de coude β sur le motif **XXPG**, semblent être requis pour que le peptide puisse interagir avec l'EBP et induire une activité biologique.

A.2.b. Etude par carte adiabatique

La structure des peptides d'élastine d'activité biologique connue a été étudiée par carte adiabatique dans l'eau.

Nous avons utilisé le modèle de solvant implicite (Ferrara et al. 2000; Ferrara et al. 2002) implémenté dans le module SASA de CHARMM (Brooks et al. 1983).

Les simulations ont été effectuées avec la version c27b3 de CHARMM et le jeu de paramètres PARAM19.

Les interactions non liantes électrostatiques et de van der Waals ont été calculées avec la fonction SHIFT, sans rayon de troncature.

Seuls les atomes lourds et les atomes d'hydrogène liés à un atome d'oxygène ou d'azote des peptides ont été construits explicitement, en accord avec les paramètres PARAM19 et le modèle de solvant SASA, avec des extrémités N- et C-terminales chargées.

Un couple d'angles dièdres ϕ et ψ a été imposé parmi les 10 zones de référence précédemment définies pour chaque résidu sauf pour la proline qui n'a que 2 zones imposées (**§ II.B.2.a.**).

Ainsi, 200 structures seront calculées pour un pentapeptide, et 2.000 pour un hexapeptide.

Une contrainte a été appliquée sur les angles dièdres pour maintenir leurs valeurs durant les phases de minimalisation pour que seuls les autres paramètres soient minimalisés. Chaque phase de minimalisation a été effectuée de la même manière : le peptide est d'abord minimalisé avec 1.000 pas de SD puis 500 pas de ABNR.

Trois phases de minimalisation ont été utilisées :

- dans la première, la valeur des angles dièdres est maintenue avec une force de 50 $kcal.mol^{-1}$, puis
- dans la deuxième phase la valeur des angles dièdres est maintenue avec une force de 5 kcal.mol-1 et
- dans la dernière phase aucune contrainte n'est appliquée.

Les structures de plus basses énergies obtenues sont regroupées par amas de conformations proches, c'est-à-dire présentant un RMSD sur le squelette protéique inférieur à 1,5 Å. Les amas ont été classés en fonction de leur structure de plus basse énergie.

α. Peptide VGVAPG

La carte adiabatique permet de dégager 4 amas principaux de conformations proches, dont les conformations de plus basses énergies sont représentées plus bas **Figure 97**. L'amas numéro 1 est celui de plus basse énergie et possède un coude β de type II sur **VGVA**, stabilisé par une liaison hydrogène entre **V_1-C=O** et **H-N-A_3**. Les angles dièdres des résidus **V** et **G** du motif **GXXP** ont des valeurs qui permettraient de les classer en coude β de type VIII, mais la distance C_α (**G_2**) - C_α (**P_5**) 7,5 Å ne le permet pas.

L'amas 2, en coude β de type IV sur **VGVA** est en coude β de type VIII sur **GVAP**.

L'amas 3 est une structure de type étendue, tandis que l'amas 4 est caractérisé par un coude β de type IV proche du type VIII sur **GVAP**. Les autres conformations n'étant pas classées dans ces amas, possèdent entre autres des angles dièdres ω cis.

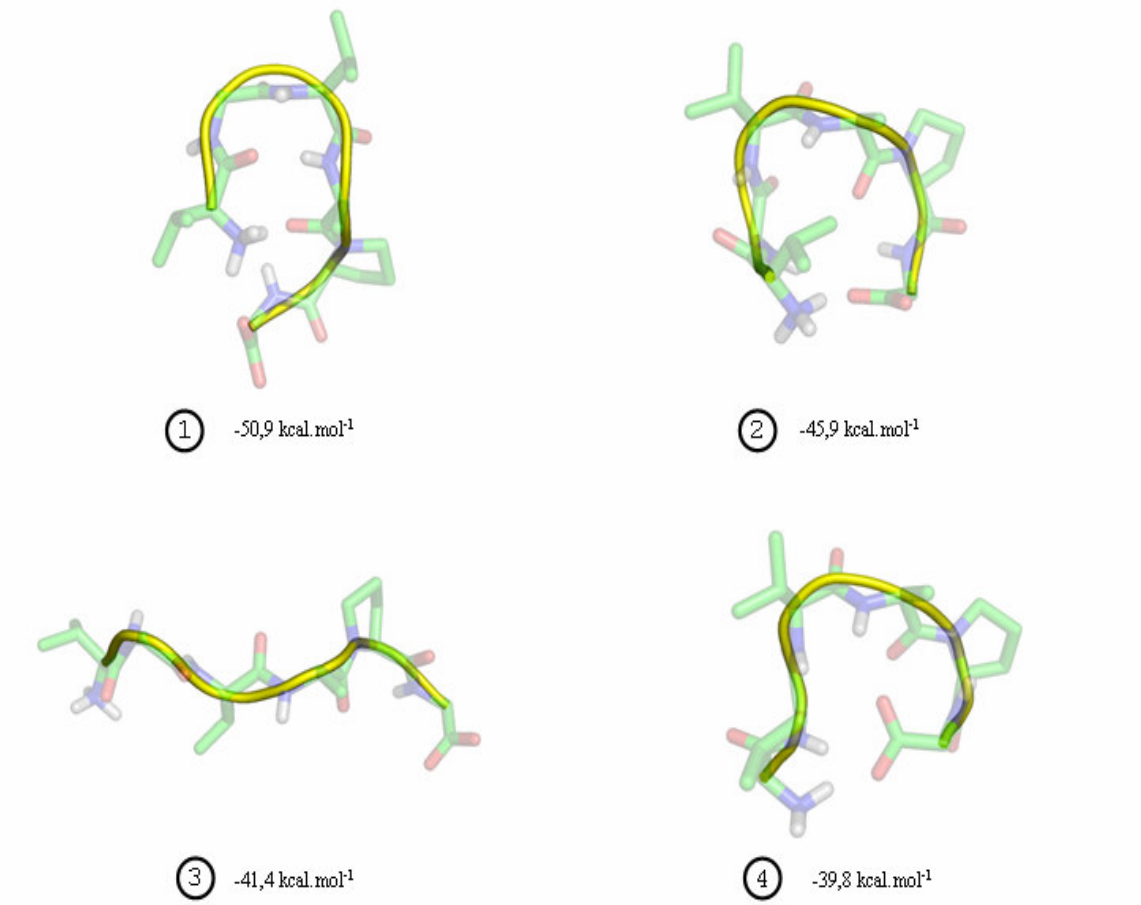


Figure 97 : Structures représentatives des amas de conformations proches de VGVAPG (associées à leur énergie, obtenues par carte adiabatique, pour le peptide).

β. Peptide PGAIPG

Trois amas principaux de structures proches ont été obtenus pour le peptide **PGAIPG** (**Figure 98**). L'amas 1 a sa conformation de plus basse énergie caractérisée par un coude β de type IV sur **PGAI** et de type VIII sur **GAIP** ; l'amas 2 possède seulement un coude β de type IV sur **PGAI et** l'amas 3 englobe les structures étendues. Comme pour le peptide **VGVAPG**, d'autres conformations ne pouvant être regroupées dans ces amas existent et présentent des angles ω cis.

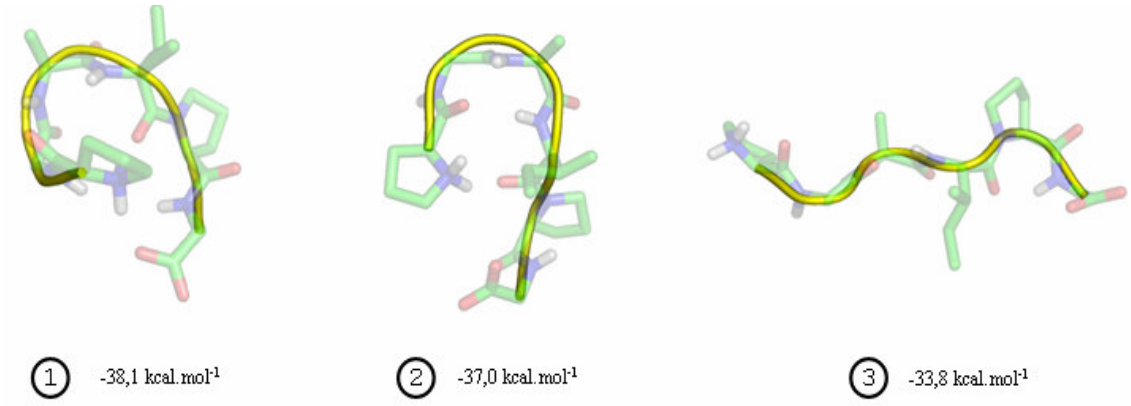


Figure 98 : Structures représentatives des amas de conformations proches de PGAIPG.

γ. Peptide GVAPG

Le peptide **GVAPG** présente 3 amas de conformations proches. La conformation de plus basse énergie de l'amas 1 est en coude β de type VIII sur **GVAP**, la proline étant en conformation cis. L'amas 2 est constitué également d'un coude β de type VIII sur **GVAP**. L'amas 3 enfin est en structure étendue. Les autres conformations ne pouvant être classées parmi les amas précédent ne sont pas représentées.

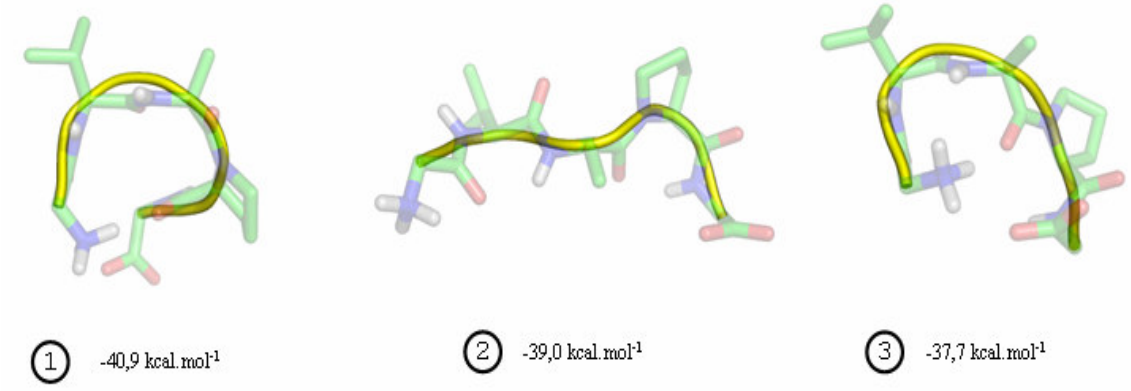


Figure 99 : Structures représentatives des amas de conformations proches de GVAPG.

δ. Peptide GLVPG

Deux amas de conformations proches sont mis en évidence : le premier amas est caractérisé par la présence de coude β de type VIII sur **GLVP**, le deuxième amas comprend des structures étendues (**Figure 100**). D'autres conformations existent mais ne peuvent être regroupées en amas et ne sont pas représentées.

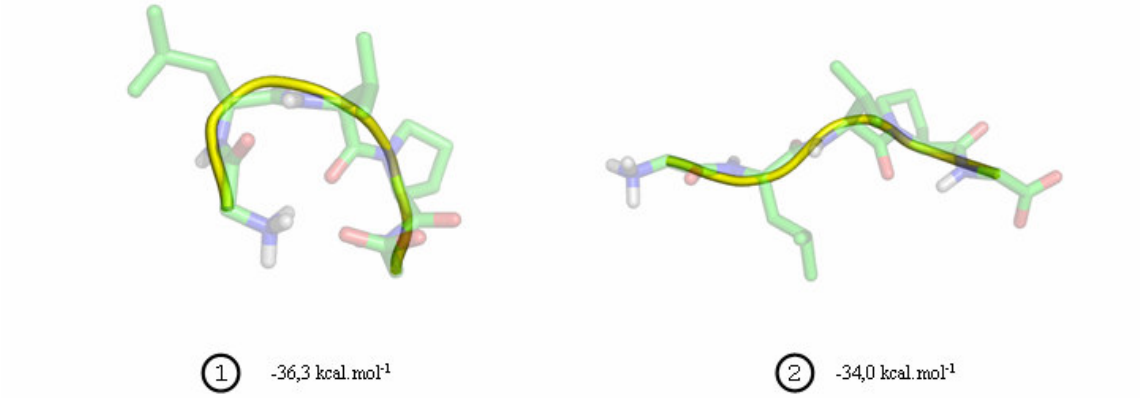


Figure 100 : Structures représentatives des amas de conformations proches de GLVPG.

ε. Peptide GGVPG

Ce peptide inactif est le seul à présenter un coude β de type II' sur **GGVP** (amas 1), l'amas 2 étant constitué de structures étendues et l'amas 3 présentant des structures repliées en coude β de type VIII sur **GGVP** (**Figure 101**).

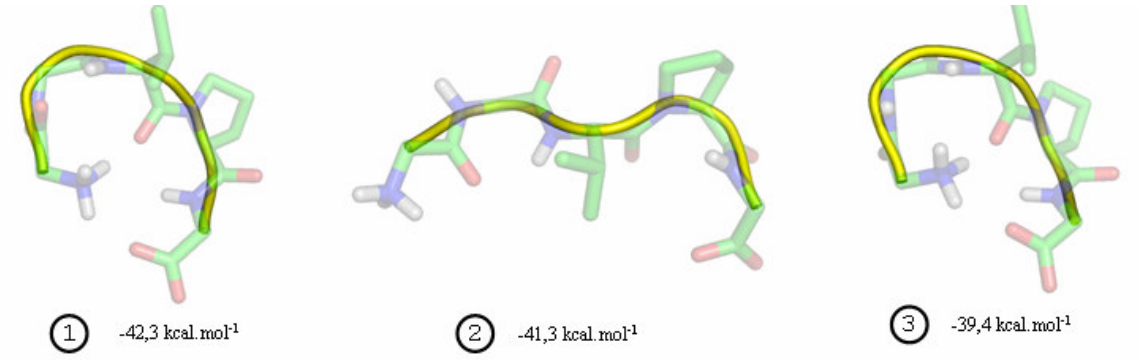


Figure 101 : Structures représentatives des amas de conformations proches de GGVPG.

ϕ. Conclusions

La méthode des cartes adiabatiques a permis de montrer que les peptides dérivant de l'élastine et contenant le motif **GXXP** se replient de manière préférentielle en coude β de type VIII sur **GXXP**. En effet, les structures de plus basses énergie représentatives des amas de conformations proches possèdent un coude β de type VIII sur **GXXP**, excepté pour le

peptide **VGVAPG** pour lequel l'amas 1 rassemble des conformations présentant un coude β de type II sur **XGXX**. Toutefois, le coude β de type VIII sur **GXXP** n'est absent dans cet amas qu'à cause de la contrainte de distance entre le C_α de $\mathbf{G}_2$ et le C_α de $\mathbf{P}_5$, les angles dièdres correspondant parfaitement aux valeurs canoniques. De plus, l'amas 2 est possède bien un coude β de type VIII sur **GVAP**.

Le peptide inactif **GGVPG**, quant à lui, est bien le seul à avoir un comportement structural atypique avec son amas de plus basse énergie adoptant un coude β de type II' sur **GXXP**.

A.2.c. Etude par dynamique moléculaire

Les hexapeptides actifs, **VGVAPG** et **PGAIPG**, induisant une activité biologique par leur interaction avec l'EBP, et **GGVPG** biologiquement inactif ont été étudiés par simulations de dynamique moléculaire en condition de solvant explicite : l'eau, avec le modèle TIP3P (Jorgensen et al. 1983) et le TFE avec des conditions périodiques aux limites. Pour les mêmes raisons que pour le peptide Sgal, l'utilisation du cyclohexane, par souci de cohérence avec les calculs de Monte-Carlo, n'a pas été possible. Les dimensions des boîtes cubiques de solvant ont été choisies de manière ce qu'il y ait une couche de solvant de 10 Å entre le peptide et la limite de la boîte, permettant ainsi d'avoir une distance minimale de 20 Å entre le peptide de la boîte principale et ceux des boîtes images.

Les interactions à "longue distance" ont été traitées avec la fonction FORCE-SWITCH pour les interactions électrostatiques et la fonction VSHIFT pour les interactions de van der Waals avec un rayon de troncature de 14 Å.

Les peptides ont été construits en conformation étendue ($\phi = 180$; $\psi = 180°$, la proline ayant un $\psi = - 76{,}1°$), avec les extrémités N- et C-terminales chargées (NH_3^+ et COO^-). L'équation du mouvement est résolue avec l'algorithme de Verlet avec un pas d'intégration de 1 fs.

Pour éliminer les mauvais contacts entre les atomes, les peptides seuls sont minimalisés avec 500 pas de SD et 500 pas de GC, puis les molécules de solvants seules sont minimalisées de la même manière. Ce cycle est répété une seconde fois et enfin l'ensemble du système est minimalisé selon ces mêmes paramètres. La température du système a été augmentée jusqu'à 300 K durant 15 ps.

La phase d'équilibrage du système s'est déroulée en deux étapes consécutives de 20 ps. Dans la première étape, les vitesses des atomes ont été attribuées tous les 50 pas suivant une fonction de distribution gaussienne.

Dans la seconde étape, les vitesses des atomes ont été modifiées de manière à ce que la température ne sorte pas d'une fenêtre de ± 10 K autour de 300 K.

Pour la phase de production dans le TFE, la température de 300 K a été régulée par couplage avec un bain externe, la constante de couplage étant de 5 ps (Berendsen et al. 1984), alors que dans l'eau, aucune contrainte n'a été imposée pour contrôler la température de 300 K.

peptides	Solvants		Durées (ns)	Dimensions de la boîte (Å x Å x Å)
	Eau	TFE		
VGVAPG	**X**		**40**	**40 x 40 x 40**
		X	**20**	**41 x 41 x 41**
PGAIPG	**X**		**40**	**40 x 40 x 40**
		X	**20**	**41 x 41 x 41**
GGVPG	**X**		**20**	**38 x 38 x 38**
		X	**20**	**38 x 38 x 38**

Tableau 33 : Conditions de simulation en solvant explicite des EDPs.

α. Peptide VGVAPG

DM dans l'eau

La trajectoire de dynamique moléculaire du peptide **VGVAPG** dans l'eau nous montre qu'il se replie essentiellement en coudes β (**Tableau 34**). Comme précédemment, le motif **GXXP** adopte une et une seule conformation, un coude β de type VIII, associé à un coude β, essentiellement de type IV sur le motif **XGXX** qui est ici composé de la séquence **VGVA**.

	Types de coudes β (%)						
	canoniques	canoniques et déformés					
	VIII	VIII	I	I'	II	II'	Total
VGVA	–	1,5	6,6	–	–	7,5	15,6
GVAP	5,5	15,8	–	–	–	–	15,8
VAPG	–	–	–	–	–	–	–

Tableau 34 : Pourcentages des types de coudes β de VGVAPG dans l'eau.

L'analyse de la présence des différents types de coudes β au cours du temps nous indique que le motif **GVAP** ne se structure en coude β de type VIII qu'après 20 ns de simulation (**Figure 102**).

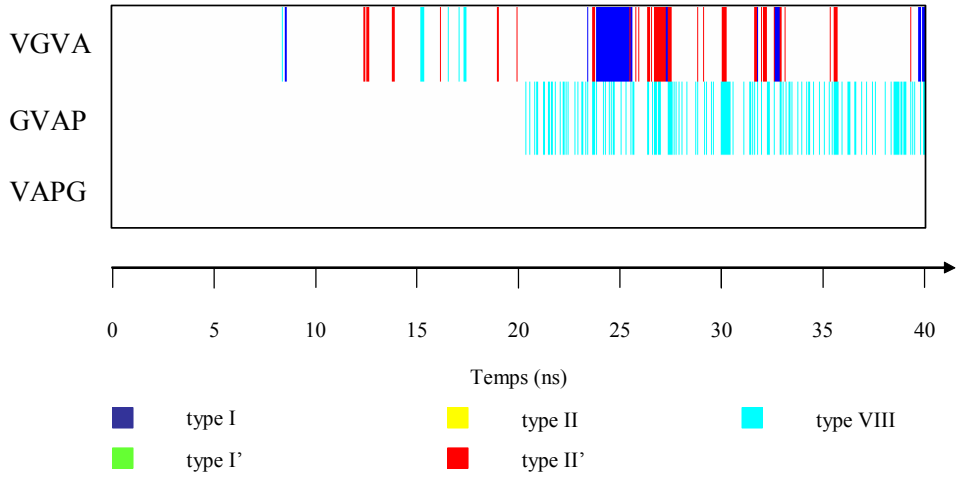


Figure 102 : Types de coudes β au cours de la DM de VGVAPG dans l'eau.

Le suivi de la présence de ce tour d'hélice polyproline II est reporté **Figure 103**. On observe 16,8 % des conformations avec un tour d'hélice polyproline II sur la séquence **VAP**. A l'inverse du coude β sur **GVAP**, l'hélice polyproline II n'est observable que durant la première partie de la trajectoire. Ces 2 structures ne sont pas compatibles en même temps, et apparaissent donc de manière alternée.

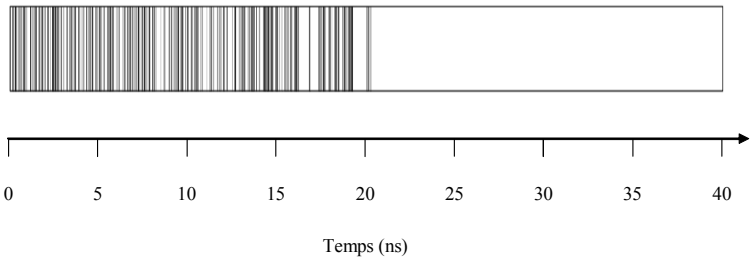


Figure 103 : Hélice polyproline II (sur VAP) de VGVAPG dans l'eau.

La carte des RMSD croisés montre également qu'il y a deux amas de structures proches, dont la limite se situe vers 20 ns (**Figure 104**). Jusqu'à 20 ns, **VGVAPG** est en structure étendue et / ou en hélice polyproline II (amas A) tandis qu'après 20 ns, le peptide adopte une structure plus compacte, en coude β de type VIII sur **GVAP** (amas B).

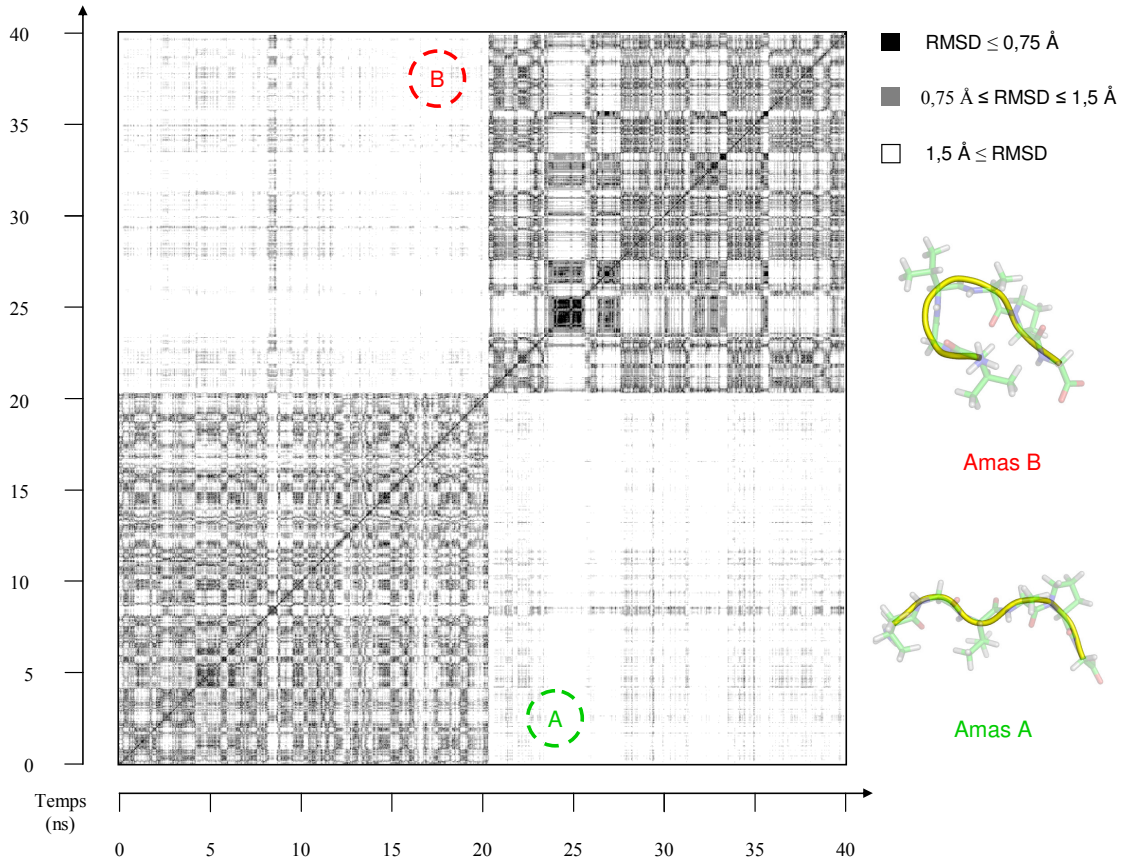


Figure 104 : Carte des RMSD croisés de VGVAPG dans l'eau.
Les 2 amas de conformations sont notés respectivement A et B.

La transition entre ces 2 amas de structure correspond à la variation de l'angle dièdre ψ de la **V_3** qui passe de + 130° à - 60° (**Figure 105**).

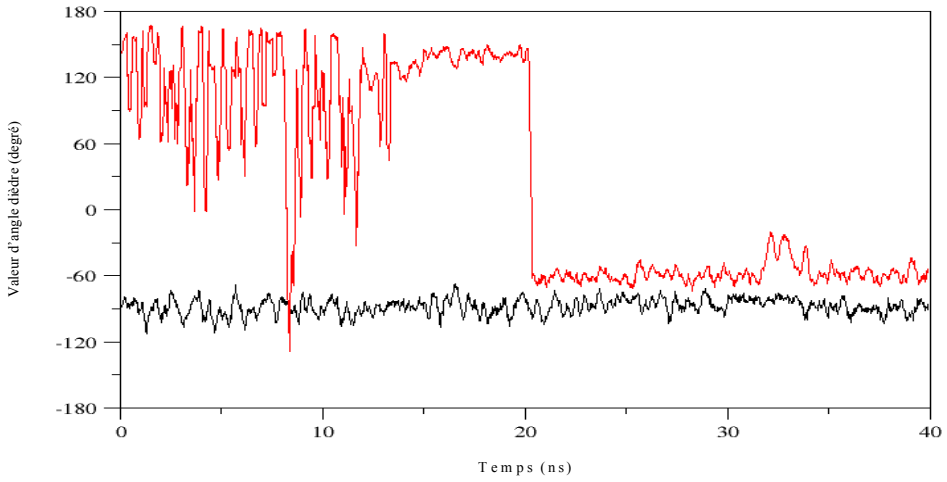


Figure 105 : **Angles dièdres ϕ (noir) et ψ (rouge) de V_3 dans la DM de VGVAPG.**

Cependant, le fait que **VGVAPG** présente deux amas de conformations dont la durée de la présence est équivalente (environ 20 ns) laisse à penser que la DM devrait être prolongée. Ce constat est renforcé par le calcul d'énergie libre pour le coude β de type VIII sur **GVAP** (**Figure 106**). On peut notamment voir que c'est à partir de 20 ns que le ΔG diminue, mais surtout qu'il ne peut pas encore être considéré comme stable à la fin de la DM de 40 ns.

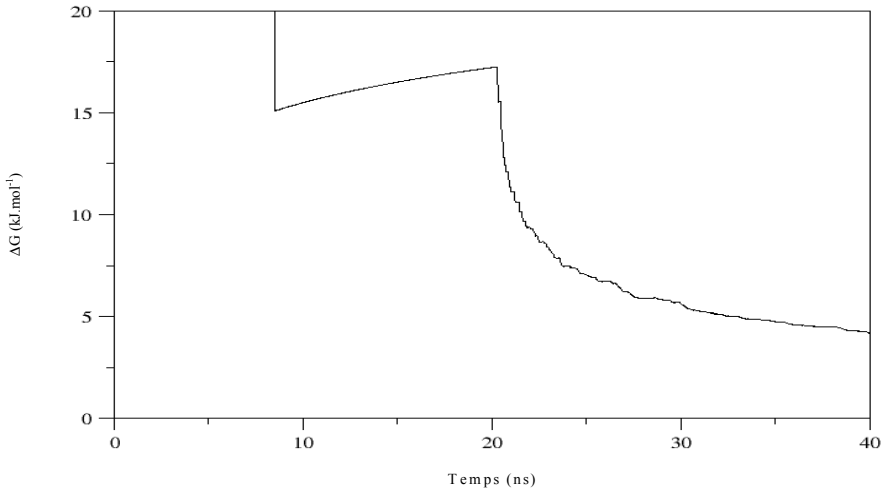


Figure 106 : ΔG du coude β-VIII sur GVAP dans la DM de VGVAPG dans l'eau.

DM dans le TFE

Le comportement du peptide **VGVAPG** dans le TFE est légèrement différent de celui qu'il manifeste dans l'eau puisqu'il présente de plus forts pourcentages de coudes β (**Tableau 35**). Il présente en effet un coude β de type VIII sur **VGVA** dans 43,9% des conformations analysées, ainsi qu'un coude β de type II' sur **GAIP** dans 53,3 % des conformations. De plus, la structure en hélice polyproline II est à peine détectable car seules 1,4 % des conformations analysées possèdent ce type de structure.

	Types de coudes β (%)						
	canoniques	canoniques et déformés					
	VIII	VIII	I	I'	II	II'	Total
VGVA	–	0,6	2,7	–	–	53,3	56,6
GVAP	17,8	43,9	–	–	–	–	43,9
VAPG	–	–	–	–	–	–	–

Tableau 35 : Pourcentages des types de coudes β de VGVAPG dans le TFE.

Le suivi de la présence des coudes β au cours du temps nous montre que le type VIII sur **GVAP** est détectable tout au long de la trajectoire (**Figure 107)**.

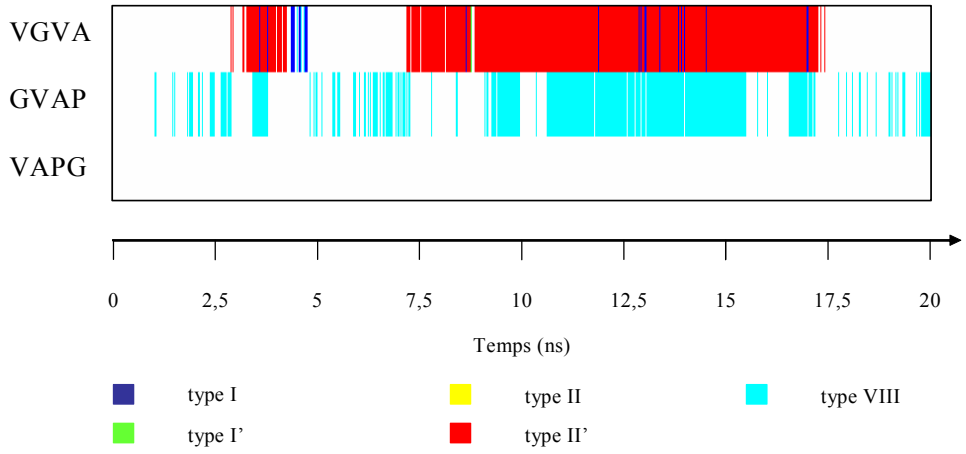


Figure 107 : Types de coudes β au cours de la DM de VGVAPG dans le TFE.

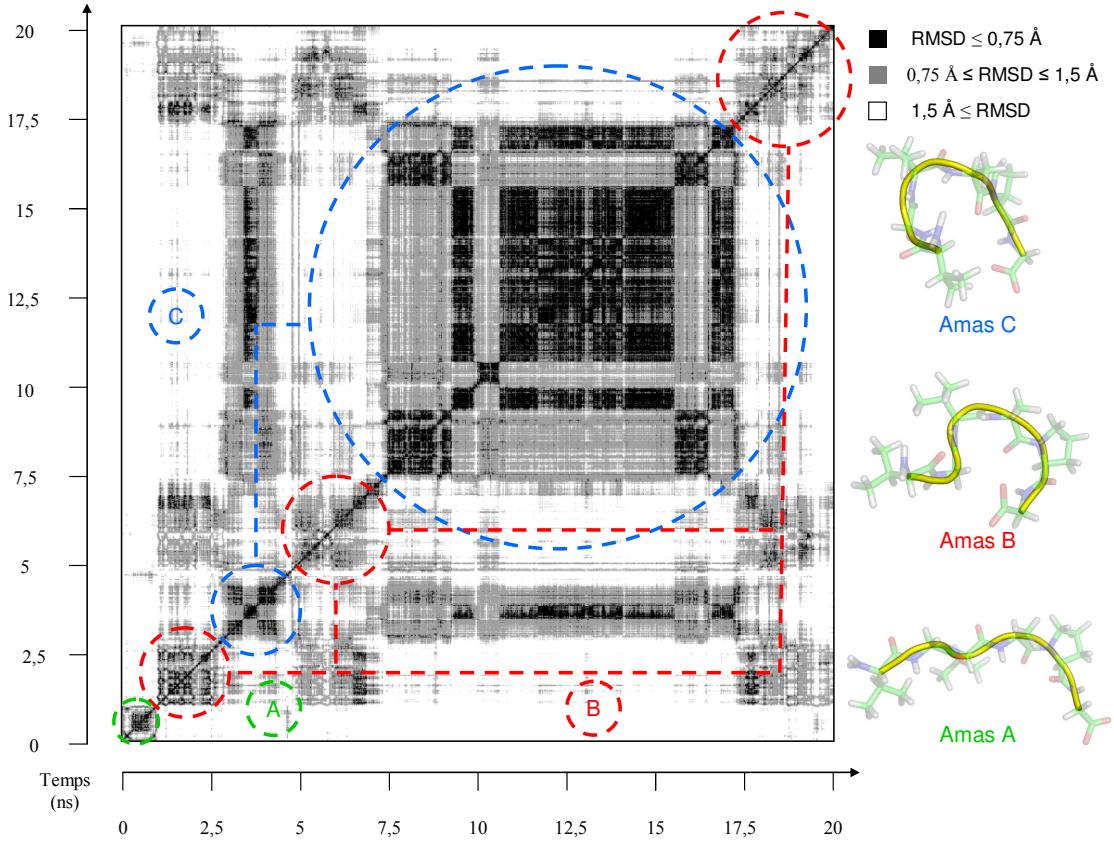


Figure 108 : Carte des RMSD croisés du peptide VGVAPG dans le TFE. On distingue 3 amas de conformations notés respectivement A, B et C.

La carte des RMSD croisés (ci-dessus **Figure 108**) montre qu'en début de trajectoire, un petit amas A est formé de structures étendues et / ou en hélice polyproline II sur **VAP**. L'amas B est formé de structures ayant un coude β de type VIII sur **GVAP**, mais pas de coude β sur **VGVA**, cette séquence présentant plutôt une structure étendue. Enfin, l'amas C, le plus peuplé, est composé de structures en coudes β de type VIII sur la séquence **GVAP** et de coudes de type II' β sur la séquence **VGVA**.

Une liaison hydrogène entre **H-N-G_2**-et **C=O-A_4**, présente dans 26,9 % des conformations analysées, stabilise les conformations de l'amas de structures C (**Figure109**).

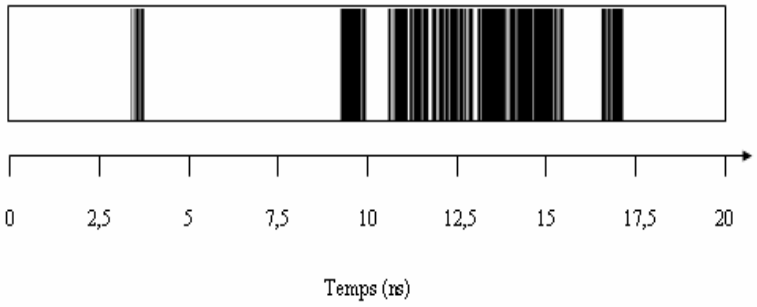


Figure 109 : Liaison hydrogène au cours de la DM de VGVAPG dans le TFE.

L'évolution de la variation de l'énergie libre ΔG de formation du coude β de type VIII sur **GVAP** permet de voir que l'équilibre entre les conformations ayant un coude β et celle n'ayant en pas est atteinte après 15 ns de trajectoire, avec une valeur de ΔG moyenne de 0,6 kJ.mol^{-1} (**Figure 110**).

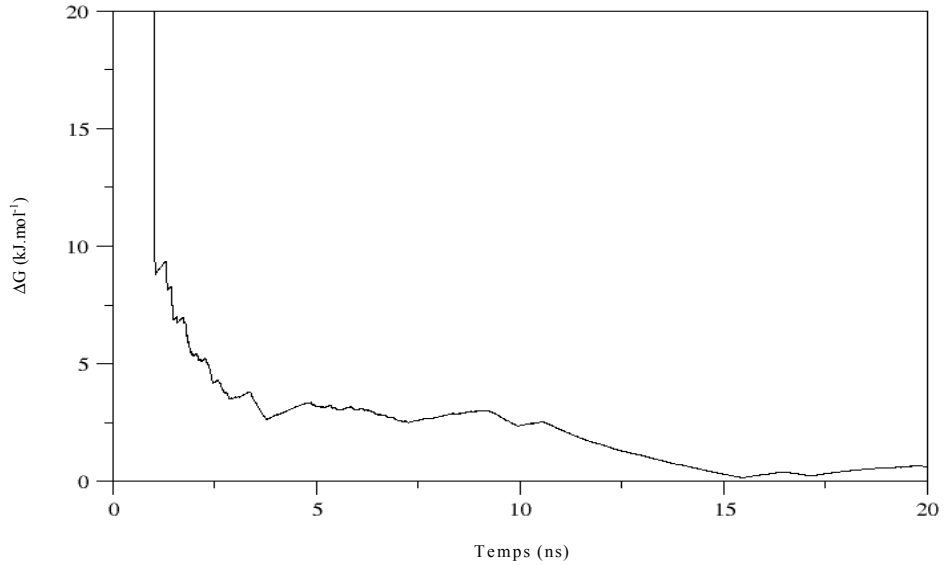


Figure 110 : ΔG du coude β-VIII (sur GVAP) dans la DM de VGVAPG dans le TFE.

Comparaison avec les données expérimentales

Si le peptide **VGVAPG** a été intensivement étudié du point de vue de ses effets biologiques par l'intermédiaire de son interaction avec l'EBP, sa structure a cependant fait l'objet de bien moins d'études.

La spectroscopie RMN du peptide **VGVAPG** a été faite dans l'eau, dans le DMSO et dans un mélange 80 % de TFE / 20 % d'H_2O (Bisaccia et al. 1994; Castiglione Morelli et al. 1997). Dans ces solvants, **VGVAPG** possède une structure plutôt étendue. Toutefois, la RMN donnant une structure moyennée dans l'espace et le temps, ce résultat n'interdit donc pas la présence de manière minoritaire et / ou transitoire d'une structure plus repliée. Les constantes de couplage de **VGVAPG** sont les mêmes dans le DMSO et le mélange 80 % de TFE / 20 % d'H_2O. Les solutions possible de l'équation de Karplus (Karplus 1959) pour les constantes de couplages NH-H_α de **V_3** et **A_4** obtenues, nous donnent ainsi les valeurs possibles de l'angle dièdre ϕ (**Tableau 36**).

	Solvant		RMN		
	Eau	DMSO ou 80 % TFE - 20 % H_2O	$J_{NH-H\alpha}$	Solutions possibles pour ϕ (°)	
Valine 3	**X**		**7,6**	**– 87°**	**– 153°**
		X	**9,1**	**– 103°**	**– 137°**
Alanine 4	**X**		**5,9**	**– 74°**	**– 166°**
		X	**6,8**	**– 81°**	**– 159°**

Tableau 36 : Résultats obtenus par RMN pour VGVAPG suivant le solvant utilisé.

La simulation de dynamique moléculaire dans l'eau est en très bon accord avec les données RMN car les valeurs de ϕ de **V_3** et **A_4** obtenues durant la trajectoire sont très proches d'une des solutions déduites de la résolution de l'équation de Karplus à partir des données expérimentales (Cf. **Figure 111)**.

La DM permet ainsi de pouvoir privilégier une des deux solutions de ϕ pour chacun des résidus : - 87° pour **V_3** et - 74° pour **A_4**.

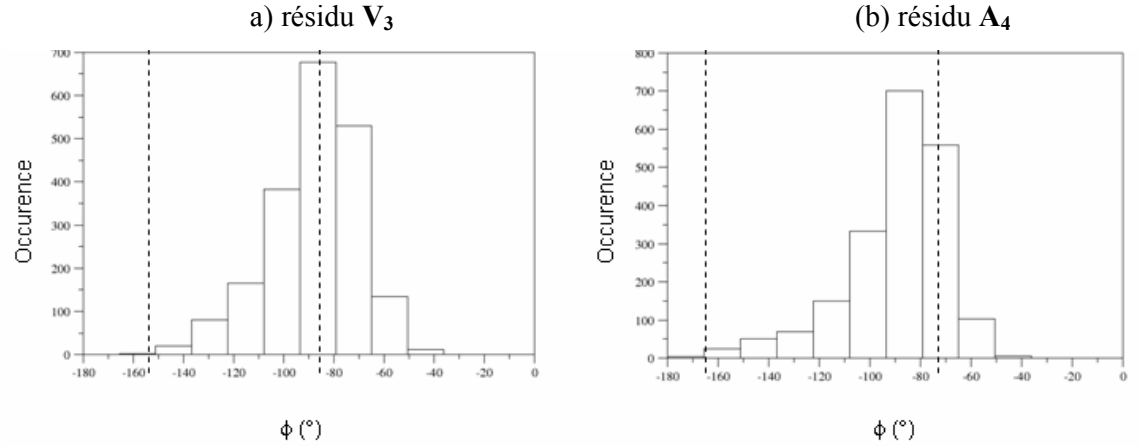


Figure 111 : Histogrammes des valeurs de ϕ durant la DM de VGV_3A_4PG dans l'eau. En pointillé : les solutions possibles des valeurs de ϕ à partir des données RMN.

De même façon que pour l'eau, la simulation de dynamique moléculaire du peptide **VGVAPG** dans le TFE reproduit bien son comportement expérimental : l'une des valeurs de ϕ déduite de la RMN de chacun des résidus est assez proche de celles obtenues durant la trajectoire (**Figure 112**).

On peut tout comme précédemment privilégier une solution pour l'angle ϕ de la $\mathbf{V_3}$ avec une valeur de - 103° et pour $\mathbf{A_4}$ avec une valeur de - 81°.

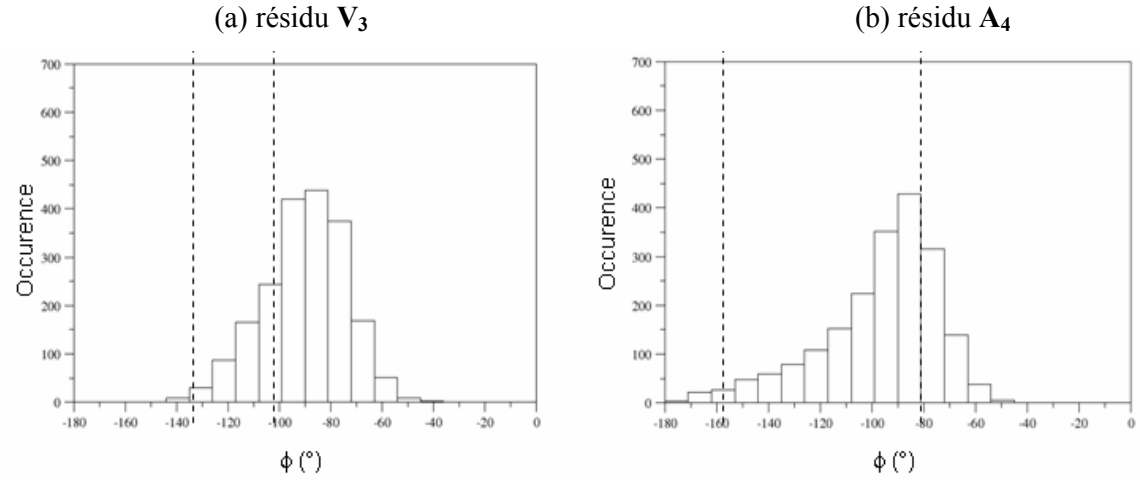


Figure 112 : Histogrammes des valeurs de ϕ durant la DM de VGV_3A_4PG dans le TFE. En pointillé : les solutions possibles des valeurs de ϕ à partir des données RMN.

L'hypothèse a été émise que puisque $\mathbf{G_2}$ avait une faible valeur de coefficient de température, son proton amide pourrait être impliqué dans une liaison hydrogène avec le groupement carbonyle de $\mathbf{A_4}$ (Castiglione-Morelli et al. 1997). Il est intéressant de remarquer

que, d'après nos résultats, la présence d'un coude β de type VIII sur **GVAP** est corrélée avec la présence de cette même liaison hydrogène.

Le peptide **VGVAPG** a également été étudié par spectroscopie de DCE dans l'eau (Bisaccia et al. 1994). Les auteurs concluent que **VGVAPG** se replie en une structure très stable car le signal n'est pas modifié par l'augmentation de la température. La présence d'une forte bande négative autour de 200 nm leur fait dire que **VGVAPG** n'a pas de structure ordonnée. Cependant, on peut remarquer la présence dans leur spectre DCE d'un épaulement à 220 nm, qui, à la lumière d'une étude récente (Fuchs et al. 2005) pourrait s'avérer être le spectre d'un peptide présentant une conformation contenant un coude β de type VIII en équilibre avec une structure non ordonnée comme nous le proposent nos propres résultats de dynamique moléculaire.

Les simulations de dynamique moléculaire du peptide **VGVAPG** sont donc en accord, tant du point de vue de la valeur des angles dièdres ϕ de $\mathbf{V_3}$ et $\mathbf{A_4}$, de la présence d'une liaison hydrogène entre $\mathbf{G_2}$ et $\mathbf{A_4}$ ou de la présence d'un coude β de type VIII, avec les données expérimentales provenant de la RMN ou du DCE.

On peut donc penser que nos simulations de dynamique moléculaire sont bien représentatives du comportement structural du peptide **VGVAPG** en solution.

β. Peptide PGAIPG :

<u>DM dans l'eau</u>

Tout comme pour le peptide **VGVAPG**, la simulation de dynamique moléculaire montre que la séquence **PGAI** se replie en coude β, mais surtout que le motif **GAIP** adopte une structure en coude β de type VIII, alors que la séquence **AIPG** ne présente aucune structure particulière.

	Type de coudes β (%)						
	Canoniques	Canoniques et déformés					
	VIII	VIII	I	I'	II	II'	Total
PGAI	**0,2**	**1,1**	**4,7**	**–**	**–**	**8,6**	**14,4**
GAIP	**12,5**	**22,5**	**–**	**–**	**–**	**–**	**22,5**
AIPG	**–**	**–**	**–**	**–**	**–**	**–**	**–**

<u>Tableau 37</u> : Pourcentages des types de coudes β du peptide PGAIPG dans l'eau.

Le suivi de la présence des types de coudes β au cours du temps montre que les coudes β sur **PGAI** et **GAIP** sont observables de façon transitoire tout au long de la trajectoire (**Figure 113**).

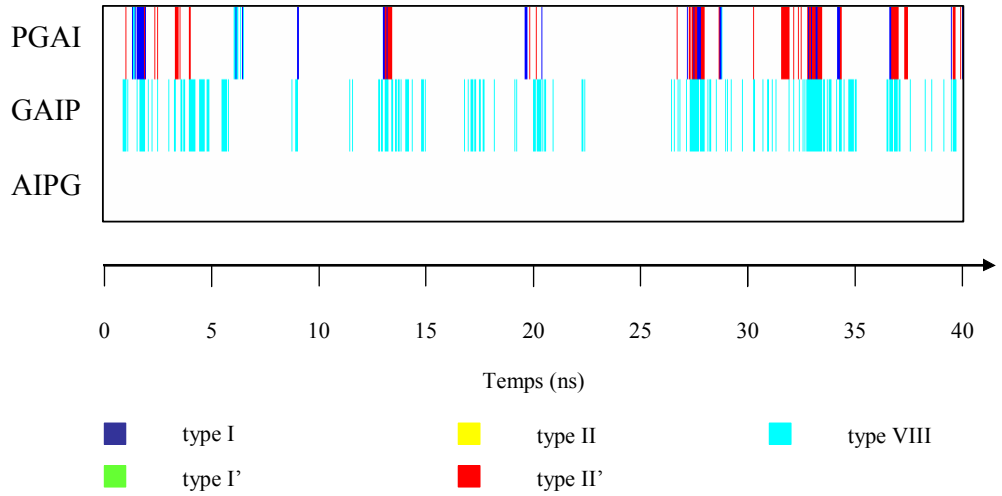


Figure 113 : Types de coudes β au cours de la DM du peptide PGAIPG dans l'eau.

Les structures possédant une hélice polyproline II sur la séquence **AIP** adoptent un comportement similaire puisque les 8,8 % des conformations, présentant ce type de structure, sont visibles de manière très transitoire au cours de la trajectoire (**Figure 114**).

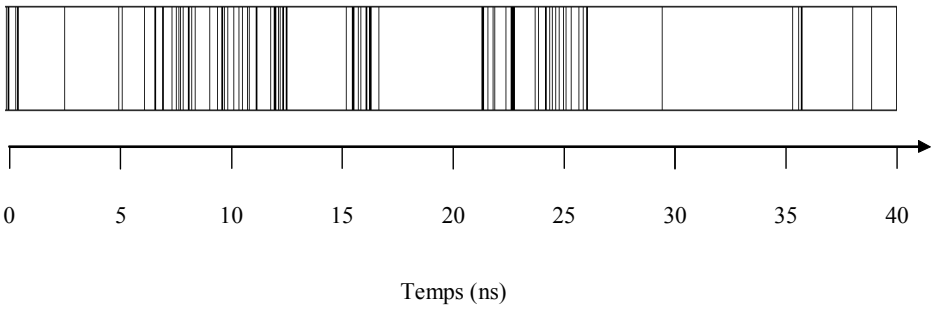


Figure 114 : Hélice polyproline II (sur AIP) au cours de la DM de PGAIPG dans l'eau.

Tout comme **VGVAPG** dans l'eau, l'hélice polyproline II sur **AIP** et le coude β de type VIII sur **GAIP** ne sont pas visibles au sein de la même conformation de **PGAIPG**.

La carte des RMSD croisés confirme ce résultat puisque nous pouvons y voir deux amas de conformations qui apparaissent de manière alternée au cours de la trajectoire (**Figure 115**). Les deux amas se succèdent rapidement ce qui empêche de pouvoir matérialiser sur la figure leur présence.

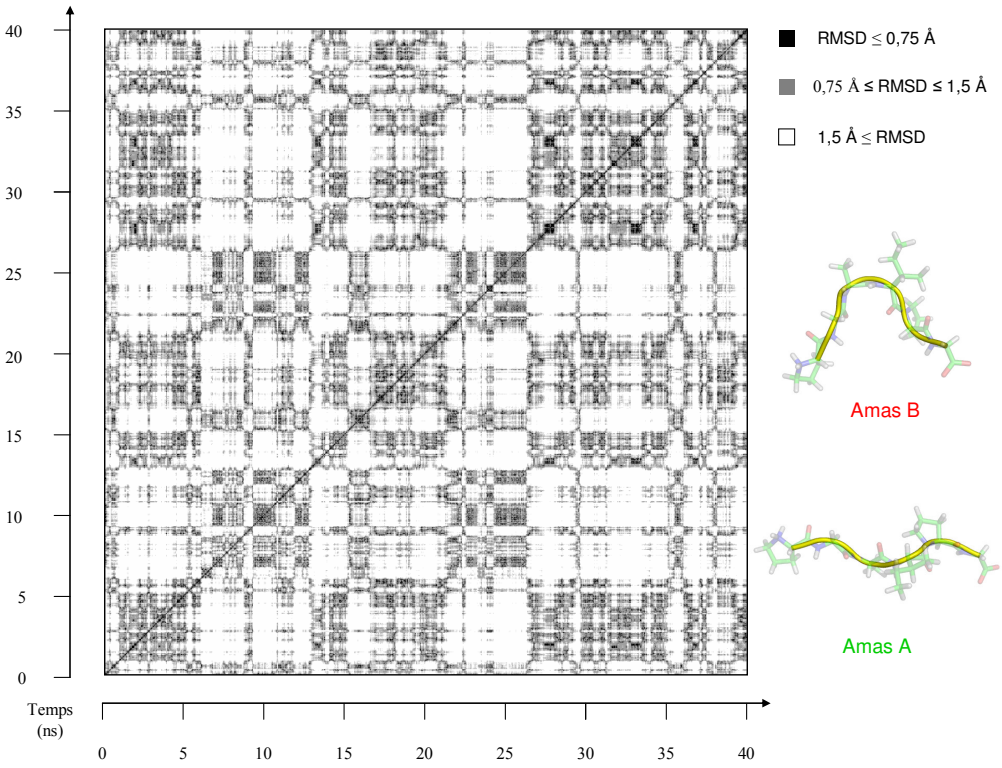


Figure 115 : Carte des RMSD croisés du peptide PGAIPG dans l'eau.

Le premier amas (A) de conformations proches est caractérisé par une structure étendue pouvant contenir un tour d'hélice polyproline II sur **AIP**. Le deuxième amas (B) contient des structures possédant un coude β de type VIII sur **GAIP** et / ou un coude β sur **PGAI**.

Le suivi du ΔG au cours de la DM montre qu'un équilibre est atteint, entre la forme repliée en coude β de type VIII sur **GAIP** et celle ne présentant pas de coude β de type VIII sur cette séquence, à partir de 15 ns, avec une valeur moyenne de ΔG de 3,0 kJ.mol^{-1}.

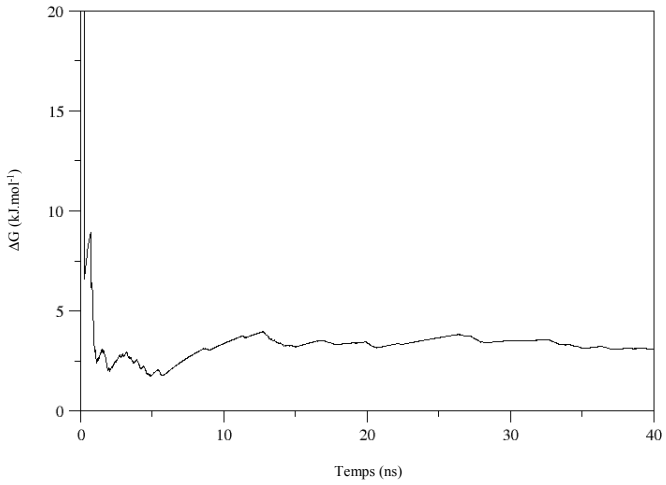


Figure 116 : ΔG du coude β -VIII (sur GAIP) dans la DM de PGAIPG dans l'eau.

DM dans le TFE

Comme pour la simulation de dynamique moléculaire dans l'eau, le peptide **PGAIPG** se replie en coude β de type VIII sur la séquence **GAIP**, de même la séquence **PGAI** adopte un coude β majoritairement de type II'.

	Types de coudes β (%)						
	Canoniques	Canoniques et déformés					
	VIII	VIII	I	I'	II	II'	Total
PGAI	0,1	0,2	6,5	–	–	49,5	56,2
GAIP	14,7	35,0	–	–	–	–	35,0
AIPG	–	–	–	–	–	–	–

Tableau 38 : Pourcentages de coudes β du peptide PGAIPG dans le TFE.

La présence d'un coude β du type VIII sur **GAIP** est stable à partir de 6,5 ns de dynamique moléculaire jusque 14 ns puis à nouveau au bout de 16 ns. Le coude β sur **PGAI** apparaît vers 6 ns jusque 14 ns et sa présence est à nouveau détectable à partir de 16 ns (**Figure 117**). Peu de conformations présentent une structure en hélice polyproline II : seules 3,4 % des conformations ont une hélice polyproline II sur **AIP**.

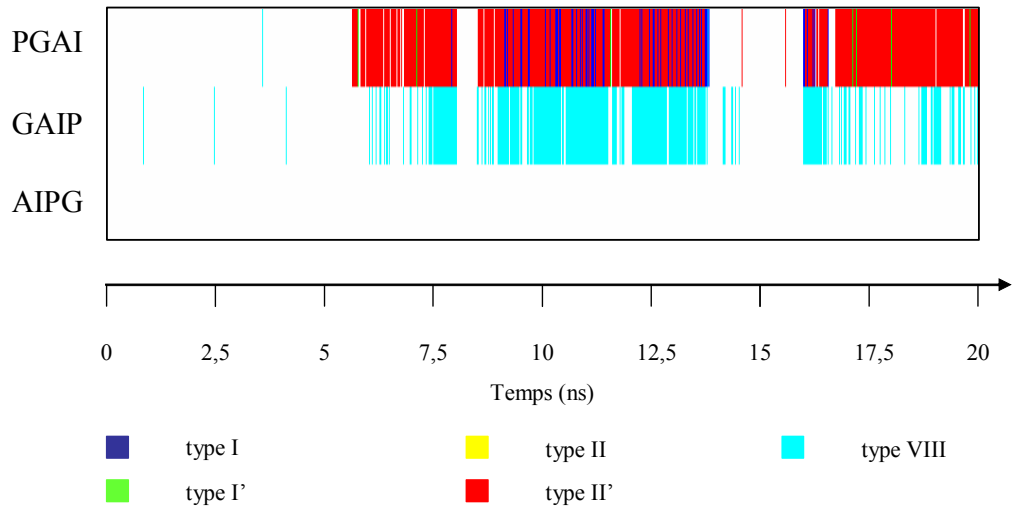


Figure 117 : Types de coude β au cours de la DM de PGAIPG dans le TFE.

La carte des RMSD croisés montre que 3 amas de conformations proches sont visibles au cours de la trajectoire de dynamique moléculaire (**Figure 118**).

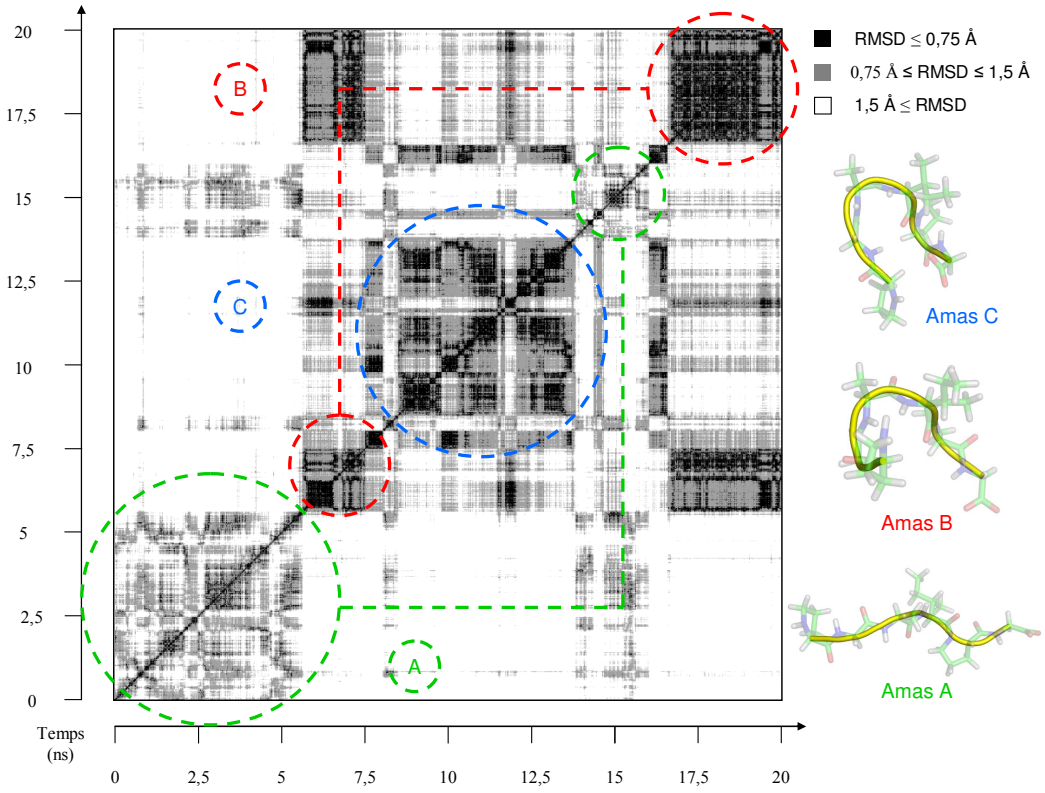


Figure 118 : Carte des RMSD croisés du peptide PGAIPG dans le TFE. 3 amas de conformations proches sont visibles (A, B et C).

L'amas A est composé de structures étendues, l'amas B et C sont caractérisés par la présence d'un coude β de type VIII sur **GAIP**, la différence étant que sur la séquence **PGAI** le coude β de type IV est proche du type I pour l'amas B et que le coude β de type IV est proche du type II' pour l'amas C. De plus l'amas C est stabilisé par une liaison hydrogène entre le **H-N-G_2** et le **O=C-A_4** : 18,4 % de l'ensemble des conformations de la trajectoire de dynamique présente cette liaison hydrogène (**Figure 119**).

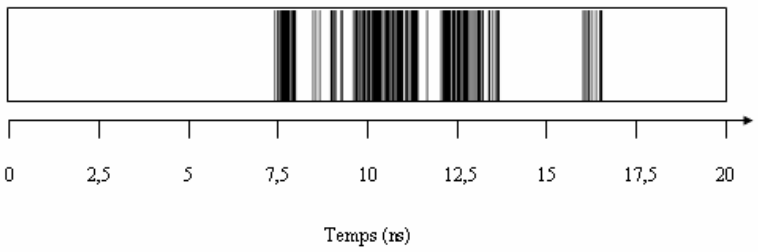


Figure 119 : **Liaison hydrogène au cours de la DM de PGAIPG dans le TFE.**

Un équilibre est atteint entre la forme repliée en coude β de type VIII sur **GAIP** et celle non repliée en coude β de type VIII sur cette séquence à partir de 13 ns de dynamique moléculaire, avec un ΔG moyen de 1,5 kJ.mol^{-1} (**Figure 120**).

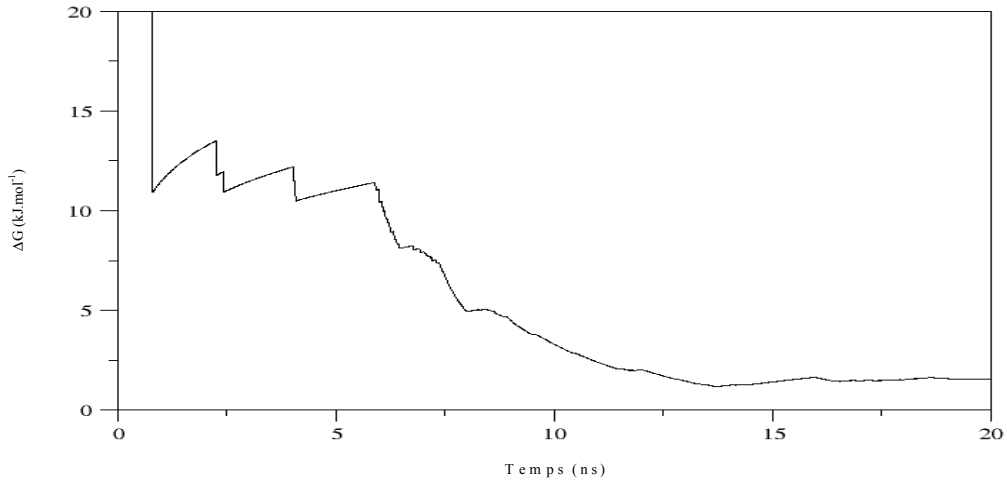


Figure 120 : ΔG du coude β-VIII (sur GAIP) dans la DM de PGAIPG dans le TFE.

Comparaison avec les données expérimentales

Malheureusement, le peptide **PGAIPG** n'a pratiquement jamais été étudié expérimentalement. Aucune donnée RMN n'est ainsi disponible.

Le seul spectre de DCE disponible n'a été enregistré qu'à une seule température. Toutefois, un épaulement à 220 nm, ainsi qu'une bande négative à 200 nm sont visibles, tout comme le signal du peptide **VGVAPG** pris dans les mêmes conditions (Brassart et al. 2001; Fuchs et al. 2001). Le spectre DCE de **PGAIPG** possède les caractéristiques d'une structure en coude β de type VIII (Fuchs et al. 2005).

Nos résultats de dynamique moléculaire sont donc en accord avec les rares résultats expérimentaux disponibles portant sur la structure de **PGAIPG**.

γ. Peptide GGVPG

DM dans l'eau

Le peptide **GGVPG** dans l'eau donne très peu de coudes β, ainsi le motif **GGVP** n'est en coude β de type VIII que dans 2,1 % des conformations, mais également en coude β de type II' dans 2,3 % des conformations. Comme pour les autres peptides, la séquence **GVPG** n'est jamais en coude β (**Tableau 39**).

	Types de coudes β (%)						
	Canoniques	Canoniques et déformés					
	VIII	VIII	I	I'	II	II'	Total
GGVP	0,7	2,1	–	–	–	2,3	4,4
GVPG	–	–	–	–	–	–	–

Tableau 39 : Pourcentages des types de coudes β du peptide GGVPG dans l'eau.

La présence des coudes β, que ce soit des coudes β de type VIII ou de type II' sur la séquence **GGVP** se répartissent tout au long de la trajectoire, de manière ponctuelle et rare (**Figure 121**).

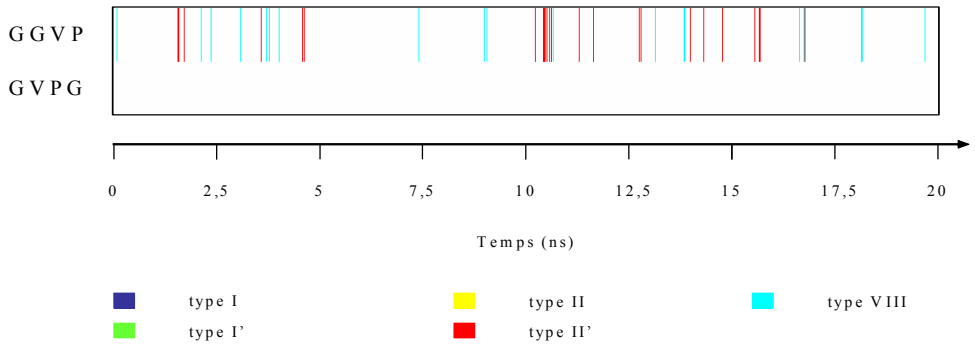


Figure 121 : Types de coudes β au cours de la DM de GGVPG dans l'eau.

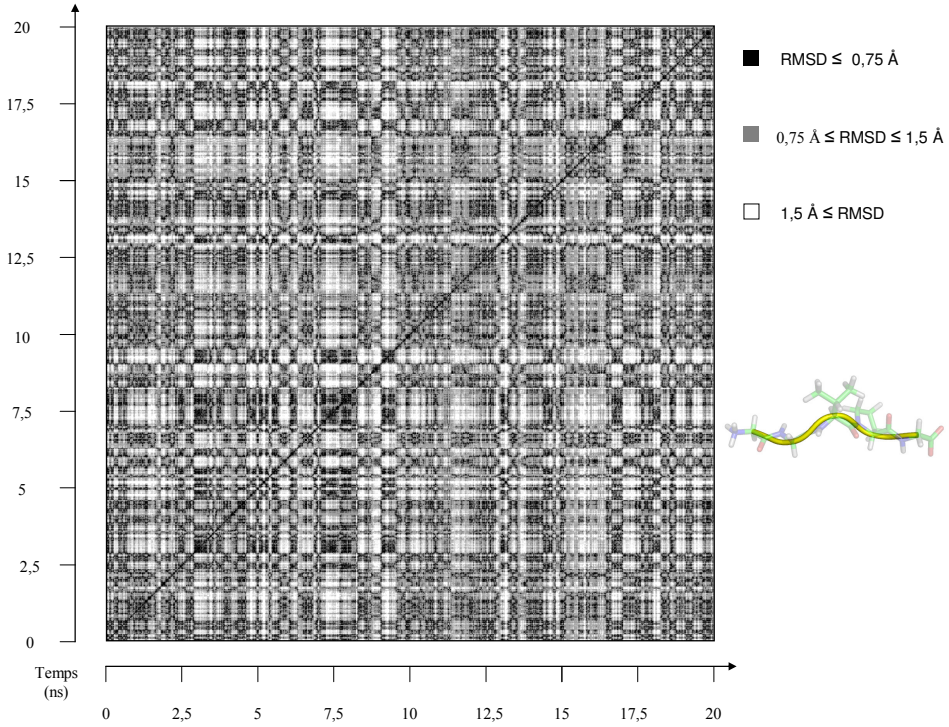


Figure 122 : Carte des RMSD croisés du peptide GGVPG dans l'eau.

La carte des RMSD du peptide de **GGVPG** nous montre que les structures que prend ce peptide sont très homogènes au cours du temps, aucun changement important dans la structure du peptide **GGVPG** n'est visible (voir plus haut **Figure 122**).

Durant toute la trajectoire de dynamique moléculaire, le peptide **GGVPG** et principalement le motif **GGVP** reste en conformation étendue, comme le montre la **Figure 123**, la distance C_α-C_α entre $\mathbf{G_1}$ et $\mathbf{P_4}$ étant en moyenne de 8,6 Å. La conformation est différente de la polyproline II puisque un tour de ce type d'hélice n'est observable que dans 1,3 % des conformations sur la séquence **AIP**.

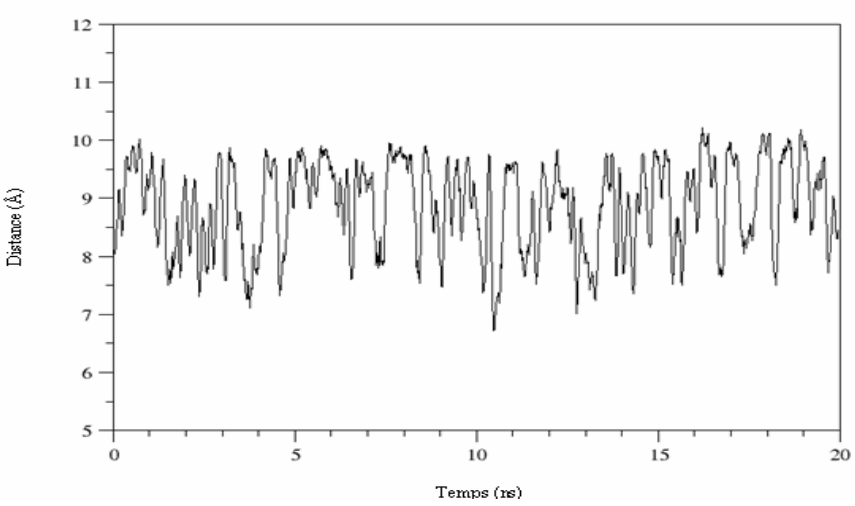


Figure 123 : Distance $C_\alpha(G_1)$-$C_\alpha(P_4)$ en fonction du temps de GGVPG dans l'eau.

L'équilibre entre les conformations coudées et non coudées est atteint à partir de 12 ns de dynamique moléculaire (**Figure 124**).

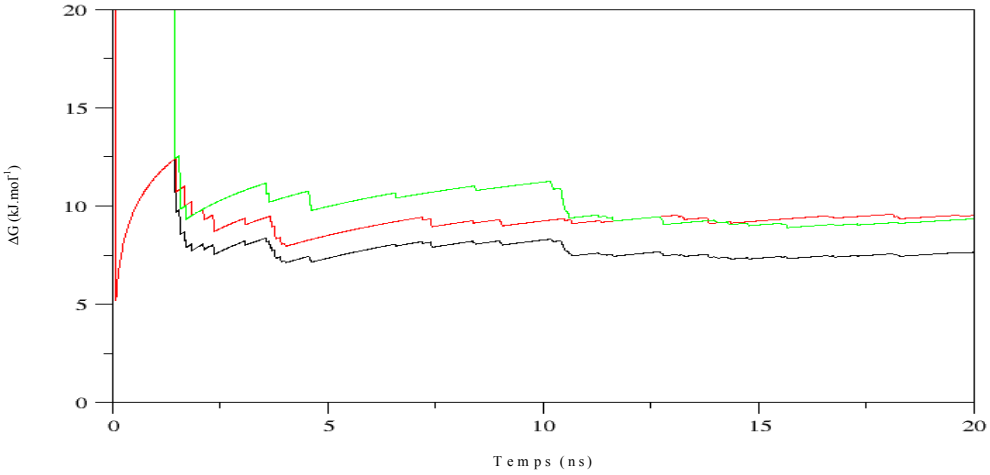


Figure 124 : ΔG du coude GGVP dans la DM de GGVPG dans l'eau.
ΔG des coudes : β, en noir; β-VIII en rouge et β-II' en vert.

Dans le TFE

Tout comme dans l'eau, le peptide **GGVPG** donne très peu de coudes β de type VIII dans le TFE (0,8 % des conformations).

Le type de coude β majoritaire est le type II' avec seulement 2,2 % des conformations présentes au cours de la trajectoire (**Tableau 40**).

	Types de coudes β (%)						
	Canoniques	Canoniques et déformés					
	VIII	VIII	I	I'	II	II'	Total
GGVP	0,2	0,8	–	–	–	2,2	3,0
GVPG	–	–	–	–	–	–	–

Tableau 40 : Pourcentages des types de coudes β du peptide GGVPG dans le TFE.

Les coudes β sur la séquence **GGVP** sont très peu fréquents et transitoires au cours de la simulation de dynamique moléculaire (**Figure 125**).

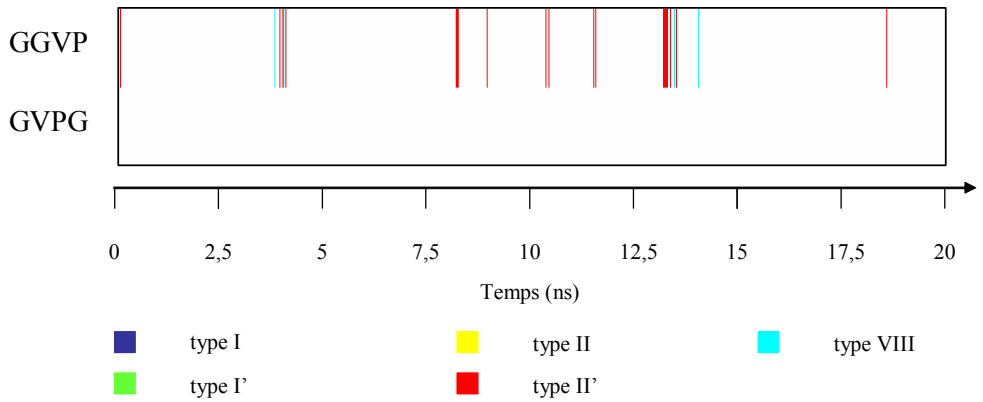


Figure 125 : Types de coudes β au cours de la DM de GGVPG dans le TFE.

La carte de RMSD ne met en évidence qu'un seul amas de conformations proches les unes des autres est observable durant les 20 ns de simulations (voir plus loin **Figure 126**).

La conformation représentative de la dynamique moléculaire est une structure étendue. En effet, la distance moyenne C_α (**G_1**) - C_α (**P_4**) est de 8,7 Å et descend extrêmement rarement en dessous des 7 Å, nécessaires à la caractérisation d'un coude β sur **GGVP** (voir plus loin **Figure 127**).

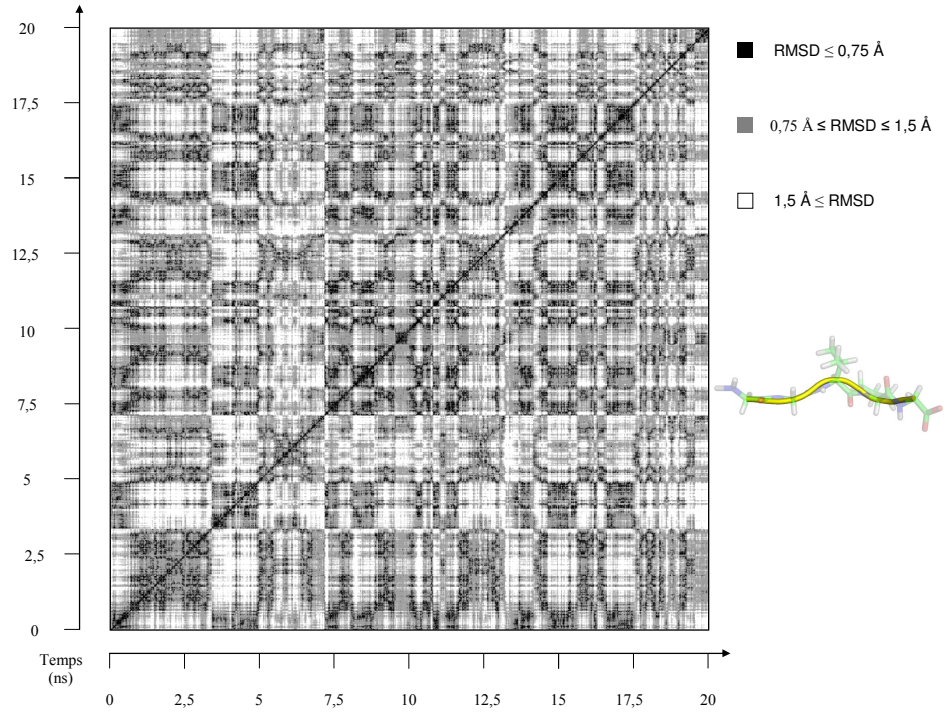


Figure 126 : Carte des RMSD croisés du peptide GGVPG dans le TFE.

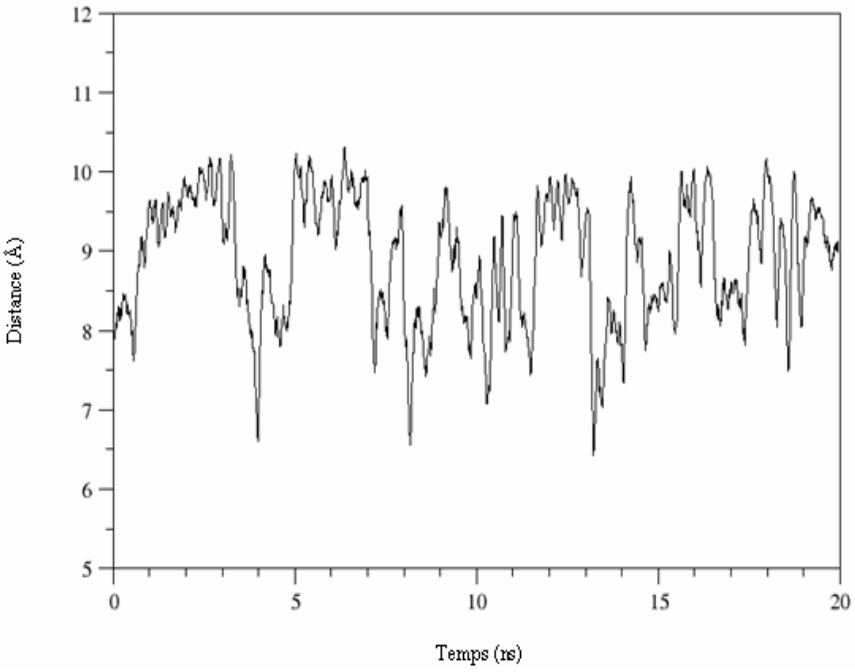


Figure 127 : Distance Cα(G_1)-Cα(P_4) en fonction du temps de GGVPG dans le TFE.

L'équilibre entre les conformations en coudes β et celle qui ne le sont pas, est atteint à partir de 14 ns de dynamique moléculaire (**Figure 128**).

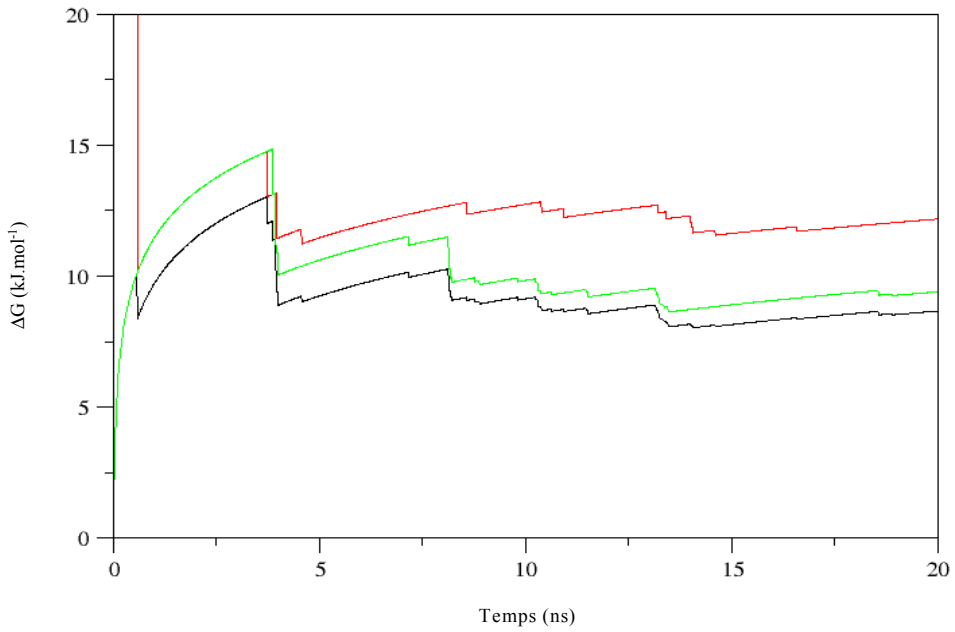


Figure 128 : ΔG du coude GGVP dans la DM de GGVPG dans le TFE. ΔG des coudes : β, en noir; β-VIII en rouge et β-II' en vert.

Comparaison avec les données expérimentales:

La structure du peptide **GGVPG** a été étudiée par spectroscopie RMN dans le DMSO et par spectroscopie DCE dans l'eau et le TFE (Castiglione Morelli et al. 1997). La RMN dans le DMSO donne une constante de couplage $J_{NH\text{-}H\alpha}$ de 7,5 Hz pour $\mathbf{V_3}$ ce qui donne un angle dièdre ϕ de - 154° ou - 86 ° (Figure **129**). Bien que le TFE et le DMSO soient tout deux des solvants ayant des propriétés physico-chimiques différentes, le peptide **VGVAPG** semblait adopter une structure suffisamment proche pour avoir des déplacements chimiques identiques dans ces deux solvants. On peut donc raisonnablement penser qu'il est de même pour le peptide **GGVPG**.

Il est donc possible de favoriser la valeur de ϕ de - 86° pour $\mathbf{V_3}$ car cette valeur est la plus fréquemment observée au cours de la trajectoire.

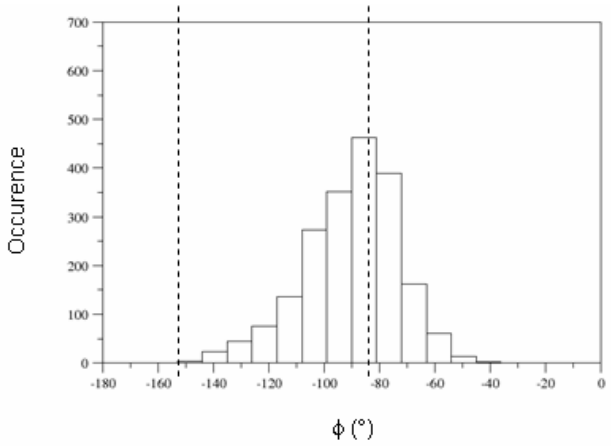


Figure 129 : Histogramme des valeurs de $\phi(V_3)$ durant la DM de GGV_3PG dans le TFE. En pointillés : les solutions possibles des valeurs de ϕ à partir des données RMN.

L'étude par DCE de **GGVPG** dans l'eau et le TFE laissait à penser qu'il était majoritairement en structure non ordonnée mais la présence de structures repliées semblait tout de même possible (Castiglione Morelli et al. 1997).

Nos résultats sont donc en adéquation avec ces résultats expérimentaux puisque nous avons également une structure majoritairement non ordonnée, plutôt étendue, avec la présence transitoire de coudes β au cours de la simulation de dynamique moléculaire (4,4 % des conformations dans l'eau et 3,0 % dans le TFE).

δ. Conclusions

Les simulations de dynamique moléculaire en solvant explicite ont permis de démontrer que les peptides biologiquement actifs, **VGVAPG** et **PGAIPG**, adoptaient une structure en coude β de type VIII sur **GXXP** associé à un coude β surtout de type II' sur **XGXX**, contrairement au peptide inactif **GGVPG**. Les formes coudées ne sont pas seules présentes, car ces peptides peuvent aussi être en conformation non structurée plutôt étendue. Ce résultat est valable aussi bien dans l'eau que dans un solvant moins polaire comme le TFE qui pourrait mimer l'environnement au voisinage de l'EBP.

Le comportement structural des peptides actifs est légèrement différent suivant le solvant. Les pourcentages de coudes β sont plus importants dans le TFE que dans l'eau. Dans l'eau, les peptides présentent également plus de structure polyproline II (16,8 % des conformations pour **VGVAPG** et 8,8 % des conformations pour **PGAIPG**) que dans le TFE. Cette présence de la structure en polyproline II est possible grâce aux molécules d'eau qui stabilisent ce type d'hélice (Sreerama et Woody 1999). L'autre différence remarquable est que, dans le TFE, la

conformation des peptides actifs est stabilisée par une liaison hydrogène entre le **H-N-G_2** et le **O=C-X_4.**

Le temps de nos simulations semble judicieusement choisi puisqu'il permet d'atteindre un équilibre entre les formes coudées et non coudés, hormis pour la simulation dans l'eau du peptide **VGVAPG** qui devrait être prolongée.

Enfin nos résultats sont en accord avec les études expérimentales précédentes portant sur la structure des EDPs du point de vue des données de RMN et de DCE qui proposaient une structure non ordonnée majoritaire, mais avec des conformations coudées possibles qui serait donc, d'après nos résultats, des coudes β de type VIII sur la séquence **GXXP**.

A.3. Etude d'un peptide inactif non naturel

Une des méthodes les plus courantes pour bloquer un peptide dans une conformation particulière est de réduire son espace conformationnel. Pour cela il est possible de le "cycliser" par une liaison peptidique entre les résidus N et C-terminaux (Kessler 1982; Weisshoff et al. 1999).

Il a donc été choisi au laboratoire d'étudier le peptide **VGVAPG** additionné de part et d'autre d'une cystéine de manière à cycliser le peptide par un pont disulfure. L'avantage du pont disulfure est de permettre de garder une flexibilité suffisante pour éventuellement toujours permettre sa fixation à l'EBP.

Le peptide **C-VGVAPG-C** cyclique a été étudié en particulier à l'état solide par spectroscopie Raman à température ambiante.

Un premier spectre balayant le domaine 400-2000 cm^{-1} a été enregistré de manière à obtenir le maximum d'informations vibrationnelles. Puis 99 spectres ont été accumulés pour augmenter le rapport signal sur bruit, avec un pas de 2 cm^{-1} et un temps d'intégration de 1 s (**Figure 130**).

La bande amide I, principalement due à la vibration d'élongation de la fonction carbonyle **C=O**, est bien visible centrée autour d'une fréquence de 1.674 cm^{-1}. Cette position de bande amide I peut donc caractériser soit un brin β (1.665-1672 cm^{-1}) soit une structure non ordonnée (1660-1670 cm^{-1}) soit même un coude β, auquel une fréquence de 1.680 cm^{-1} à été associée (Krimm et al. 1986). La bande amide III, provenant de l'élongation **C-N** et de la déformation angulaire **C-N-H** de la liaison peptidique, est constituée de deux maxima à 1256 et 1294 cm^{-1}. Elles peuvent être attribuées respectivement à une structure non ordonnée (1243-1255 cm^{-1}) et de type hélice α (1270-1300 cm^{-1}).

La bande visible à 510 cm^{-1} est caractéristique de la présence d'un pont disulfure de conformation GGG.

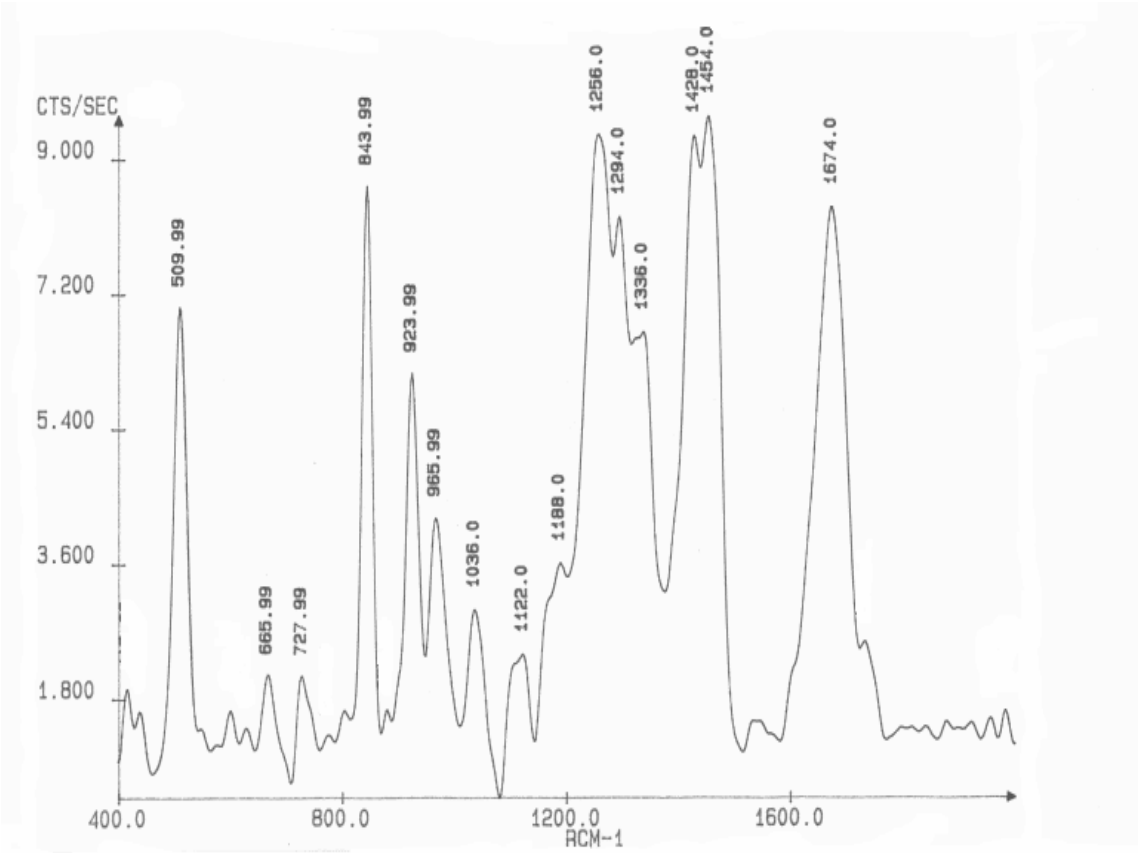
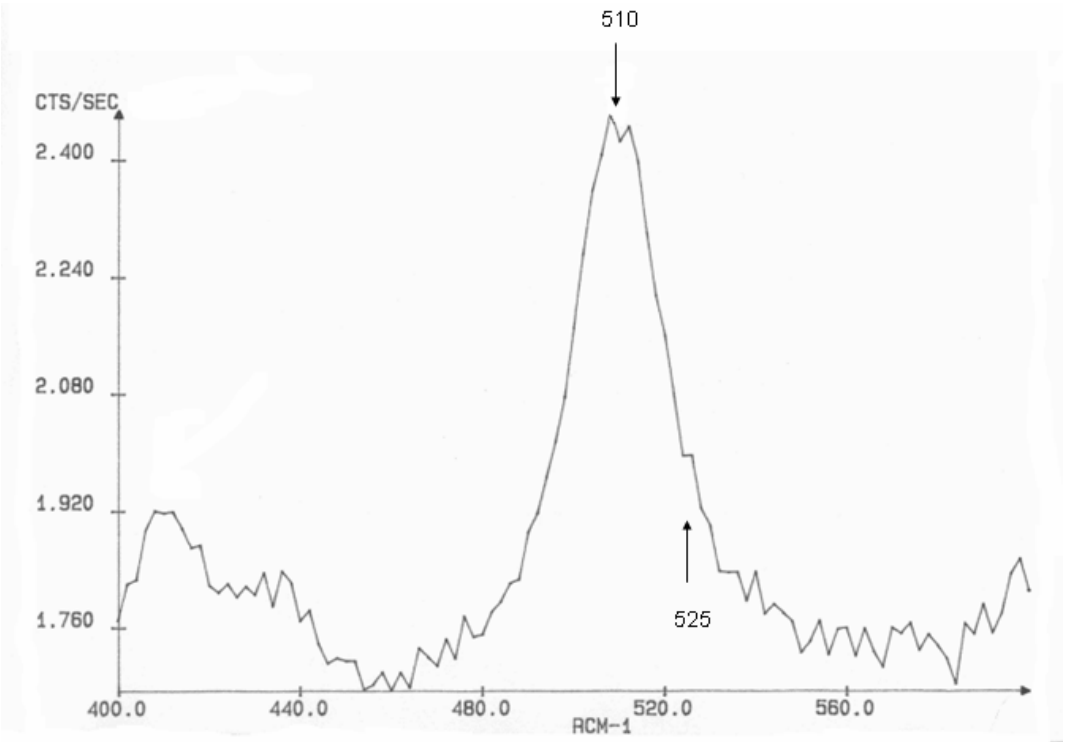


Figure 130 : Spectre Raman normal du peptide cyclique C-VGVAPG-C. En abscisse R : nombre d'onde relatif en cm^{-1}.

Un second spectre a été enregistré dans les mêmes conditions expérimentales, mais en se restreignant au voisinage de la zone caractéristique des ponts disulfures, c'est-à-dire entre 400 et 600 cm^{-1} (**Figure 131**).

Le spectre obtenu confirme la présence quasi exclusive d'un pont disulfure en conformation GGG (510 cm^{-1}). A cela s'ajoute également un faible épaulement à 520 cm^{-1}, caractéristique d'un pont difulfure en conformation GGT ou TGG.
L'autre résultat d'importance est l'absence de bande à 540 cm^{-1} indiquant ainsi que le pont disulfure n'adopte pas de structure de type TGT.



<u>Figure 131</u> : Spectre Raman du peptide cyclique C-VGVAPG-C dans le domaine caractéristique des liaisons -S-S (ponts disulfures).
En abscisse R : nombre d'onde relatif en cm^{-1}.

Les autres études de ce peptide (tests biologiques, spectroscopies de DCE, Monte-Carlo et dynamique moléculaire) ont permis de montrer que contrairement à ce que l'on pouvait espérer, le peptide cyclique **C-VGVAPG-C** ne se replie pas en coude β de type VIII et n'avait aucune activité biologique (Floquet 2004).

Ce résultat renforce l'hypothèse que le coude β sur le motif **GXXP** des peptides d'élastine est nécessaire à l'interaction entre les peptides et l'EBP.

B. AUTRES LIGANDS PEPTIDIQUES DE L'EBP : LES PEPTIDES DERIVANT DE LA FIBRILLINE-1 :

Ils existent plusieurs protéines de la matrice extracellulaire (MEC), autres que la tropoélastine, qui contiennent également des motifs **GXXPG** et qui sont des sites potentiels d'interaction avec l'EBP.

Le **Tableau 41** regroupe les protéines de la matrice extracellulaire, contenant au moins 3 motifs **GXXPG**, présentes dans les bases de données protéiques de la NCBI (National Center for Biotechnology Information) et de la Swiss-Prot (Bairoch et Apweiler 1997).

Nom	Nombre de séquence GXXPG	Banque de données	
		NCBI	Swiss-Prot
Tropoélastine	**32**	**NP_000492**	**P15502**
Tenascine X	**11**	**NP_061978**	**P22105**
Tenascine N	**8**	**XP_040527**	**Q9UQP3**
Fibronectine	**7**	**NP_997641**	**P02751**
Fibrilline 2	**7**	**NP_001990**	**P35556**
Collagène VIIα1	**7**	**NP_000085**	**Q02388**
Galectine 3	**5**	**NP_002297**	**P17931**
Fibrilline 3	**4**	**NP_115823**	**Q75N90**
Collagène IIIα1	**3**	**NP_000081**	**P02461**
Collagène IVα5	**3**	**NP_00486**	**P29400**
Collagène Vα2	**3**	**NP_00384**	**P05997**
Collagène VIIIα2	**3**	**NP_005193**	**P25067**
Laminine α3	**3**	**NP_000218**	**Q16787**
Fibrilline 1	**3**	**NP_000129**	**Q75N87**
Tenascine C	**3**	**NP_002151**	**-**
Aggrecane 1	**3**	**NP_001126**	**-**
Netrine 2	**3**	**NP_006172**	**O00634**
EMILIN 1	**3**	**NP_008977**	**Q9Y6C2**
EMILIN 2	**3**	**NP_114437**	**Q9BXX0**

Tableau 41 : Protéines de la matrice extracellulaire contenant au moins 3 motif GXXPG.

Plusieurs pathologies associées à un dysfonctionnement de la régulation d'enzymes protéolytiques pourrait être dues et / ou amplifiées par la dégradation de protéines de la MEC. Les peptides, issus de la dégradation contenant un motif **GXXPG**, pourraient être capables de

se lier à l'EBP, induisant à leur tour la production d'enzymes protéolytiques, contribuant ainsi à l'entretien, voire l'augmentation de la dégradation des protéines de la MEC et de l'inflammation.
La tenascine C, par exemple, contenant 3 séquences **GXXP** est connue pour induire une augmentation de l'expression d'une élastase, la MMP-9 (Kalembeyi et al. 2003; Ilunga et al. 2004), et sa présence est associée à l'augmentation de la production d'une autre élastase, la MMP-2 dans la sténose aortique calcifiée (Jian et al. 2001).

Le syndrome de Marfan est une maladie héréditaire, dont les symptômes principaux sont un dysfonctionnement des tissus conjonctifs cardiovasculaires, oculaires et squelettiques, qui se caractérise par une importante fragmentation des fibres élastiques (Fleischer et al. 1997; Pyeritz 2000). Il a été démontré que ce sont des mutations du gène de la fibrilline 1 (FNB1), un des composants essentiels pour la formation des fibres élastiques, qui sont responsables de ce syndrome (Dietz et al. 1991). Le mécanisme exact de cette maladie reste indéterminé, mais l'une des hypothèses avancées est que les FNB1 mutées seraient plus facilement dégradées par les enzymes protéolytiques de la matrice extracellulaire et que leurs fragments seraient eux-mêmes à l'origine d'une augmentation de la dégradation protéolytique. Cette hypothèse est notamment étayée par la détection de taux élevés de métalloprotéinases matricielles chez les malades atteints du syndrome de Marfan (Segura et al. 1998; Sachdev et al. 2002).

Les effets biologiques de peptides provenant de la fibrilline 1 et contenant une séquence **GXXPG** ont été étudiés *in vitro* sur la production de MMP-1 (Institut de Génétique médicale, Hôpital Universitaire Charité, Université d'Humboldt, Berlin, Allemagne).

B.1. Etudes expérimentales

B.1.a.Peptides étudiés

Un peptide recombinant de 276 résidus, appelé rFib47^{wt} (wt pour "wild type" : type sauvage), a été construit par synthèse biochimique, correspondant aux exons 47 à 53 de la fibrilline 1 (acides aminés $\mathbf{I_{1.929}}$ à $\mathbf{E_{2.205}}$).

La séquence **GXXPG** est située dans un domaine protéique particulier appelé EGF (Epidermal Growth Factor), entre les acides aminés 2.194 à 2.199 (**Figure 132**).

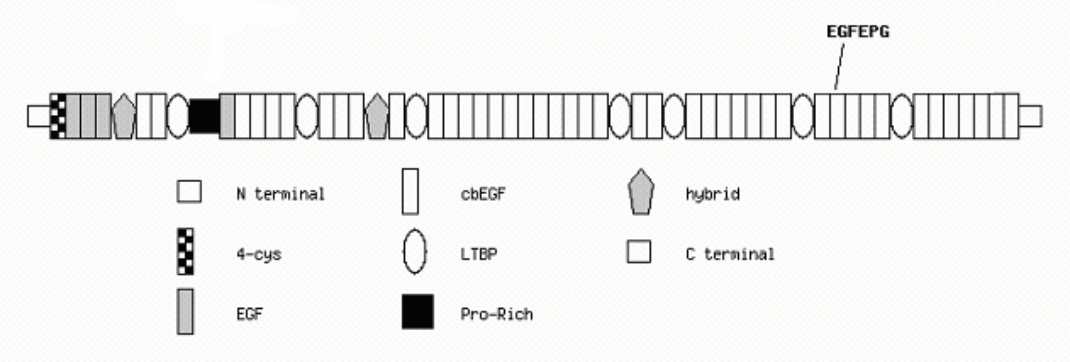


Figure 132 : Domaines constituant le peptide de fibrilline 1.

Un deuxième peptide recombinant, appelé rFib47^{mt}, (mt pour "mutated type" : type muté), a également été testé. Ce dernier est exactement le même que le premier, hormis que la proline 2.198 du motif **GXXPG** a été remplacée par une sérine (**EGFEPG → EGFESG**).

Les peptides **VGVAPG**, qui va servir de témoin positif, et **EGFEPG** ont été produits par synthèse chimique (Ansynth Service, Roosendaal, Pays-Bas).

B.1.b. Protocole d'étude

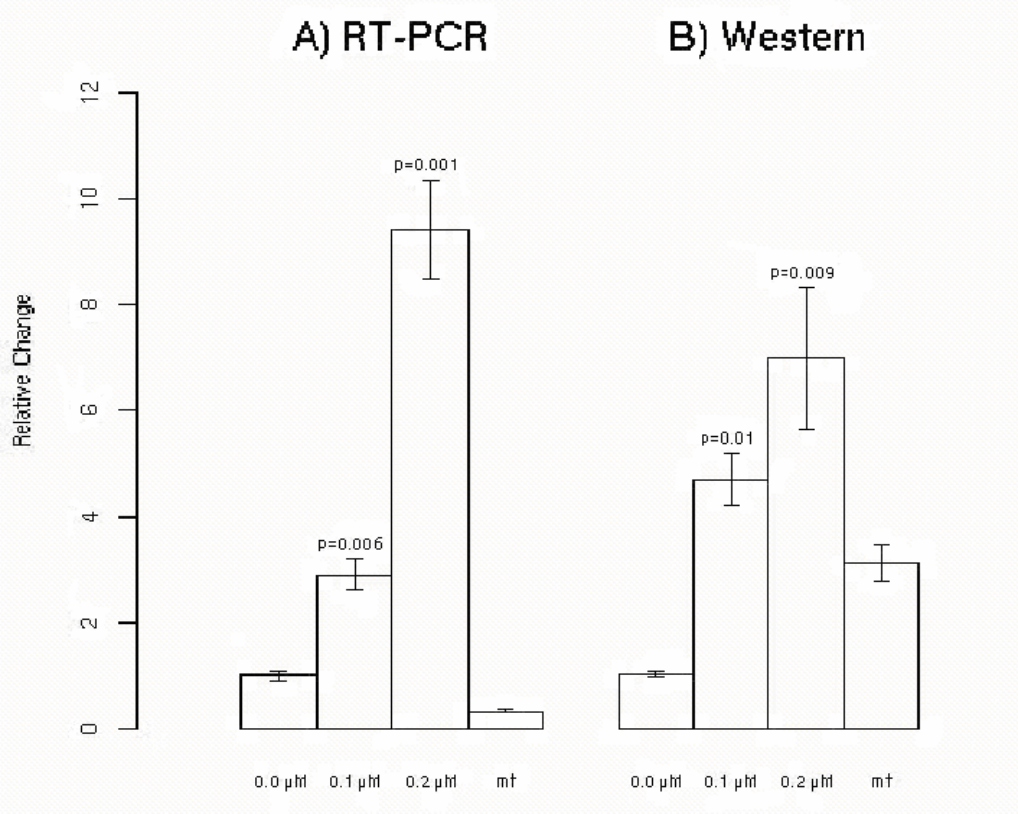
Les tests *in vitro* ont été réalisés sur des fibroblastes dermiques humains sur lesquels ont été incubés les peptides recombinants à différentes concentrations (0 µM ; 0,1 µM et 0,2 µM). Après 48 heures d'incubation, l'ARN total a été isolé pour permettre l'utilisation de la méthode RT-PCR (Reverse Transcription-Polymerase Chain Reaction) pour détecter et quantifier la production d'ARN messager de la MMP-1, c'est-à-dire l'expression du gène de la MMP-1 (**Figure 133A**). Les différentes stimulations ont été répétées chacune 6 fois. Pour mettre en évidence la production de MMP-1, des expériences de "Western blot" ont, de plus, été effectuées séparément avec les peptides recombinants rFib47^{wt} et rFib47^{mt} (**Figure 133B**) et les hexapeptide synthétiques **VGVAPG** et **EGFEPG** (**Figure 134**).

B.1.c. Résultats :

La **Figure 133** donne les résultats de la RT-PCR (A) et du Western blot (B).

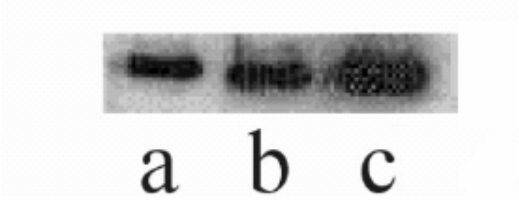
On peut voir que l'incubation, à une concentration de 0,2 µM de peptide recombinant contenant la séquence **GFEPG** (rFib47^{wt}), augmente d'un facteur de 9,4 l'expression du gène de la MMP-1, d'une manière dose-dépendante. Alors qu'à la même concentration, le peptide recombinant ne possédant pas de motif **GXXPG** (rFib47^{mt}) n'induit aucune expression du gène de la MMP-1 (**Figure 133A**). De même, le peptide recombinant avec la séquence

GFEPG montre un effet dose-dépendant sur la production de MMP-1, tandis que l'absence de motif **GXXPG** provoque une diminution significative de la production de MMP-1 (**Figure 133 B**).



<u>Figure 133</u> : Influence des peptides recombinants sur A) l'expression du gène de la MMP-1 et B) la production de MMP-1. Les stimulations à différentes concentrations sont celles du peptide rFib47^{wt} ; mt indique la stimulation effectuée à 0,2 µM avec le peptide recombinant muté rFib47^{mt}.

L'analyse des résultats des expériences de Western blot obtenus par stimulation avec les hexapeptides synthétiques **VGVAPG** et **EGFEPG** confirme la stimulation de la production de MMP-1 puisque chacun de ces peptides montre clairement une augmentation de la production de MMP-1 (**Figure 139**).



<u>Figure 134</u> : Western blot de la MMP-1 obtenus après stimulation a) sans peptide, b) avec VGVAPG et c) avec EGFEPG.

En conclusion, le peptide recombinant dérivant de la fibrilline 1 contenant un motif **GXXPG**, qui est le motif d'interaction entre l'EBP et ses ligands, est capable d'induire une augmentation de l'expression et de la production de la MMP-1. Dans le même peptide dans lequel un seul résidu a été muté de manière à ne plus avoir de motif **GXXPG**, mais une séquence **GFESG**, les effets précédents sur la régulation de la MMP-1 ne sont plus observables. Sous forme d'hexapeptide, **GEFPG** augmente également la production de MMP-1. Ce type de comportement a déjà été précédemment observé avec des peptides dérivant de l'élastine (Brassart et al. 2001), mais cette étude permet de mettre en évidence qu'une mutation au niveau du motif **GXXPG**, induit la perte de l'activité biologique.

Pour comprendre la conséquence de la mutation, nous avons étudié par simulation de dynamique moléculaire la structure des hexapeptides **EGFEPG** et **EGFESG**.

B.2. Etudes par simulations de dynamique moléculaire

B.2.a. Protocole des simulations de dynamique moléculaire

Nous avons eu recours au modèle de solvant implicite SASA dans les mêmes conditions que celles utilisées pour le calcul des cartes adiabatiques des peptides d'élastine.
Deux modèles de départ ont été construits pour chacun des deux peptides :

- un modèle "tout étendu", appelé T_1 (pour Trajectoire 1), avec les angles dièdres ϕ et ψ à 180°, excepté pour la proline pour laquelle ϕ est égale à - 76,1°.

- un modèle construit par homologies, appelé T_2, obtenu à partir de la structure RMN du domaine cbEGF (calcium binding Epidermal Growth Factor) de la fibrilline-1 (**Figure 135**), dont la référence PDB est **1emn** (Downing et al. 1996). La séquence **EGFEPG** est en structure non ordonnée et non pas en coude β de type VIII.

Figure 135 : Domaine cbEGF de la fibrilline 1 ; représentation *ruban* pour l'ensemble de la protéine et représentation *bâtons* pour le motif EGFEPG.

Les structures de départ ont été minimalisées avec 500 pas de l'algorithme SD et 500 pas de l'algorithme GC. La phase de chauffage a duré 50 ps et celle d'équilibrage 80 ps, selon le protocole précédemment décrit (Cf. Matériel et méthodes).

La phase de production d'une durée de 45 ns, avec un pas d'intégration de 1 fs, a eut sa température de 300 K régulée par un couplage avec un bain externe, la constante de couplage étant de 5 ps (Berendsen et al. 1984).

L'analyse des trajectoires a été faite sur 45.000 conformations (1 par ps), en terme de structures secondaires avec le programme STRIDE (Frishman et Argos 1995), tandis que la présence d'hélice polyproline II et des différents types de coudes β a été étudiée à l'aide des deux programmes précédemment utilisés (Cf. Matériel et méthodes).

L'énergie libre de formation du coude β de type VIII a été calculée pour chacune des trajectoires selon la formule de Gibbs (Cf. Matériel et méthodes).

B2.b. Résultats des simulations de dynamique moléculaire

L'analyse en terme d'hélice polyproline II sur la séquence FE**P** ou FE**S** a permis de montrer que seulement environ 1,0 % des conformations avaient une telle structure et ce pour l'un ou l'autre des peptides.
Outre ce faible pourcentage, aucune différence ne peut être faite suivant l'activité biologique, ce qui laisse à penser que ce n'est pas dans l'hélice polyproline II que réside l'origine de l'activité biologique des peptides contenant la séquence **GXXP**.

Pour les deux trajectoires de chacun des peptides, nous avons reporté **Tableau 42**, les pourcentages des différents types de coudes β présents sur chaque tétrapeptide possible (quadruplet de résidus), c'est-à-dire pour chacune des séquences **EGFE**, **GFEX**, **FEXG**.

La **Figure 136** donne la présence de ces différents types de coudes β en fonction du temps pour les deux simulations de DM, correspondant respectivement aux deux modèles de départ tout étendu (DM T1) ou construit par homologie (DM T2) .

a)

Type de coude β (%)	EGFE**P**G					
	Etendu (T_1)			Homologie (T_2)		
	EGFE	GFE**P**	FE**P**G	EGFE	GFE**P**	FE**P**G
I	2,7	–	–	2,9	–	–
I'	–	–	–	–	–	–
II	–	–	–	–	–	–
II'	**73,8**	–	–	**78,5**	–	–
VIII	1,9	**35,1**	–	1,1	**38,8**	–
Total	**78,4**	**35,1**	–	**82,5**	**38,8**	–

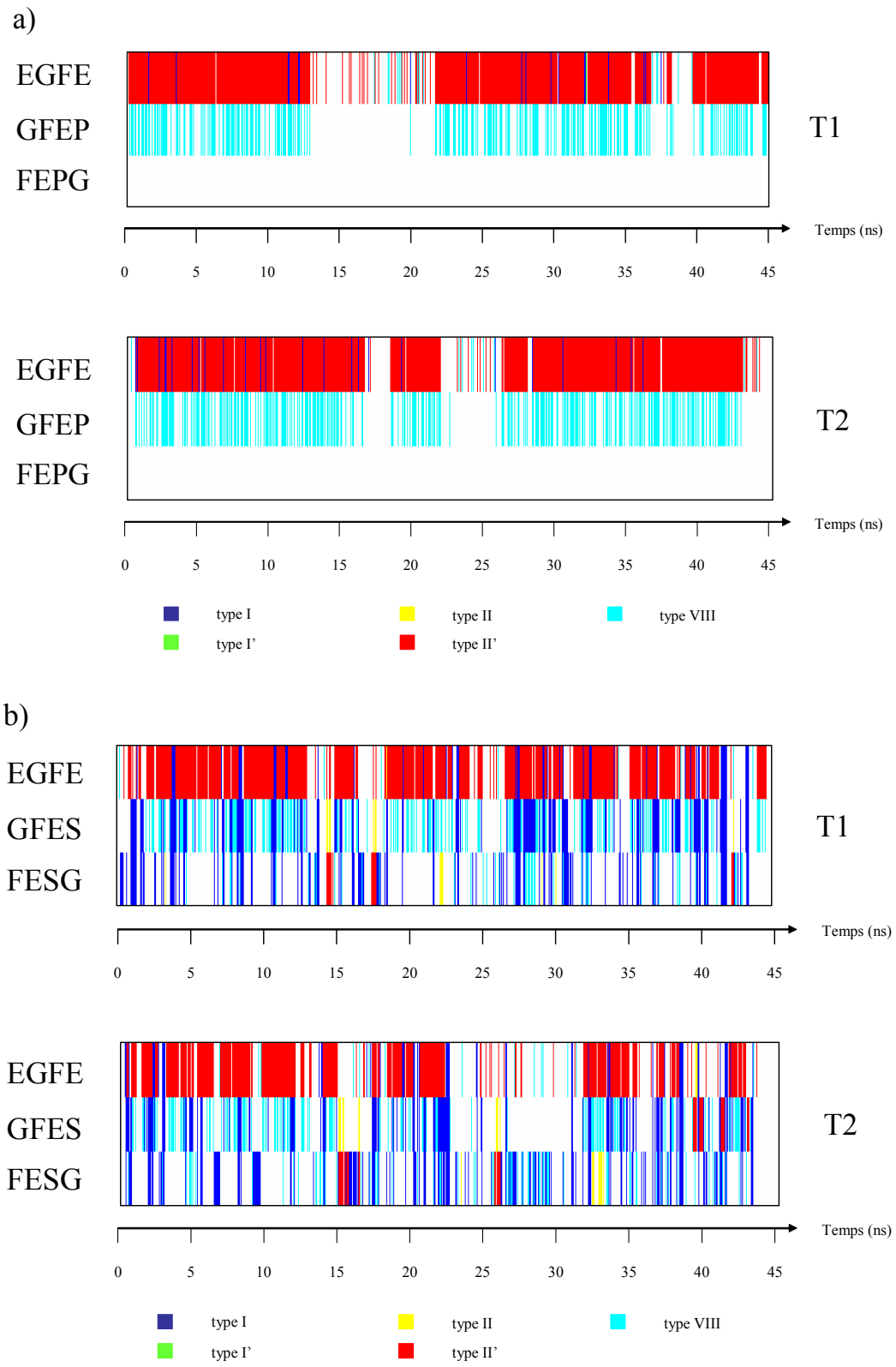
b)

Type de coude β (%)	EGFE**S**G					
	Etendu (T_1)			Homologie (T_2)		
	EGFE	GFE**S**	FE**S**G	EGFE	GFE**S**	FE**S**G
I	7,9	**25,8**	**20,8**	6,6	**19,4**	**23,3**
I'	–	–	–	–	–	–
II	–	0,9	0,9	–	0,9	1,3
II'	**64,7**	–	2,3	**45,7**	2,1	2,3
VIII	1,4	**22,7**	6,5	2,3	**15,3**	8,1
Total	**74,0**	**49,4**	**30,5**	**54,6**	**37,7**	**35,3**

Tableau 42 : Types de coudes β observés pour les peptides a) EGFEPG et b) EGFESG.

Il est intéressant de constater que le peptide **EGFEPG** se replie de manière quasi exclusive en deux coudes β consécutifs : de type II' sur **EGFE** (T_1 : 73,8 % et T_2 : 78,5 %) et de type VIII sur **GFEP** (T_1 : 35,1 % et T_2 : 38,8 %). Ces deux types de coude β ont des valeurs d'angles dièdres compatibles lorsqu'ils sont ainsi placés sur cinq résidus consécutifs.

Le coude β de type II' sur **EGFE** est stabilisé par une liaison hydrogène entre des atomes du squelette peptidique, entre le **O=C-E_1** et le **H-N-E_4**, dans 44,2 % des conformations pour T_1 et 32,4 % pour T_2. Par contre, aucun coude β n'est observable sur la séquence **FEPG**.



<u>Figure 136</u> : Types de coudes β présents dans les DM T1 et T2 sur a) EGFE, GFEP et FEPG et b) EGFE, GFES, FESG.

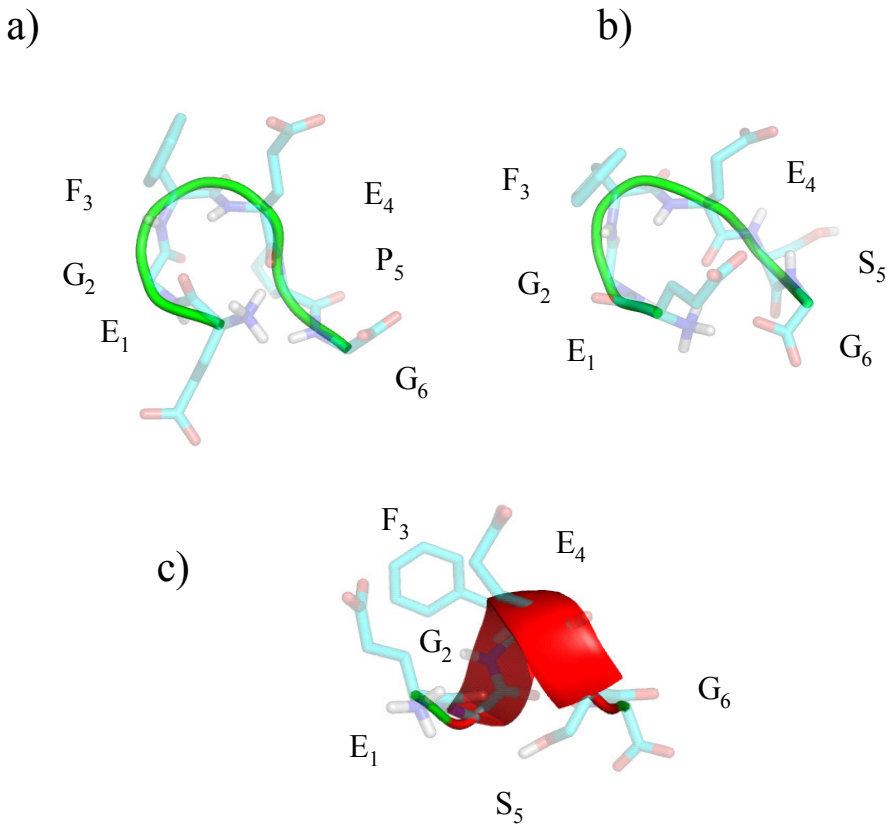
Le peptide **EGFE<u>S</u>G** présente un comportement structural différent du peptide **EGFE<u>P</u>G**. Un coude β de type II' est toujours observable sur les résidus **EGFE** (T_1 : 64,7 % et T_2 : 45,7 %) qui est encore stabilisé par la même liaison hydrogène, présente dans 27,9 % des conformations pour T_1 et 19,4 % pour T_2.

Mais désormais, la séquence **GFES** peut se replier en coude β de type VIII (T_1 : 22,7 % et T_2 : 15,3 %), mais également, de manière plus importante, en type I (T_1 : 25,8 % et T_2 : 19,4 %) et par contre, très rarement en type II (T_1 : 0,9 % et T_2 : 0,9 %).

Il est assez surprenant de remarquer qu'aucune liaison hydrogène n'est détectée ici, alors que la sérine, contrairement à la proline, pourrait en créer une avec la glycine.

Un coude β est désormais présent sur la séquence **FE<u>S</u>G**, majoritairement de type I (T_1 : 20,8 % et T_2 : 23,3 %), plus faiblement de type VIII (T_1 : 6,5 % et T_2 : 8,1 %) et transitoirement de type II' (T_1 : 2,3 % et T_2 : 2,3 %) ou de type II (T_1 : 0,9 % et T_2 : 1,3 %).

L'enchaînement de 3 coudes β sur le peptide **EGFE<u>S</u>G** entraîne un repliement en hélice α (T_1 : 2,2 % et T_2 : 1,9 %) ou en hélice $_{3\text{-}10}$ (T_1 : 3,9 % et T_2 : 3,4 %).



<u>Figure 137</u> : Structures des peptides a) EGFEPG (EGFE: coude $\beta_{II'}$; GFEP: coude β_{VIII}), b) EGFESG (EGFE : coude $\beta_{II'}$; GFES : coude β_{VIII} ; FESG : β_I) et c) EGFESG (EGFE : coude $\beta_{II'}$; GFES : coude β_I; FESG : β_I).

La présence de la sérine confère ainsi une plus grande liberté de conformation au peptide **EGFESG**. La proline ne restreignant pas le résidu précédent, c'est-à-dire l'aspartate (**E**), dans une zone favorable au coude β de type VIII, celle correspondant à la zone des brins β, les pourcentages coudes β de type VIII diminuent logiquement. Le coude β peut être désormais de type I ou VIII. Il est intéressant de remarquer que seul le résidu **E**, permet de différencier le type I et le type VIII, car les valeurs d'angles dièdres sont les mêmes pour le résidu **F**. En type I, **E** a pour valeur d'angles dièdres $\phi = -90°$ et $\psi = 0°$, alors qu'en type VIII, **E** prend les valeurs $\phi = -120°$ et $\psi = +120°$. Il semble donc que l'absence de la proline permet la "libération" de l'angle ψ du résidu la précédant. Un coude β, n'est possible sur le motif **XXPX** que si le résidu précédant la proline est une glycine (voir § **IV.B.1.a.β.**). Cependant, la mutation de la proline en sérine, permet la formation de ce type de coude β sur les quatre derniers résidus, ce qui n'est le cas d'aucun des peptides contenant le motif **GXXP** ayant une activité biologique, précédemment étudié.

L'énergie libre ΔG pour le coude β de type VIII est tracée en fonction du temps pour chacune des trajectoires de simulations de dynamique moléculaire (**Figure 138**).

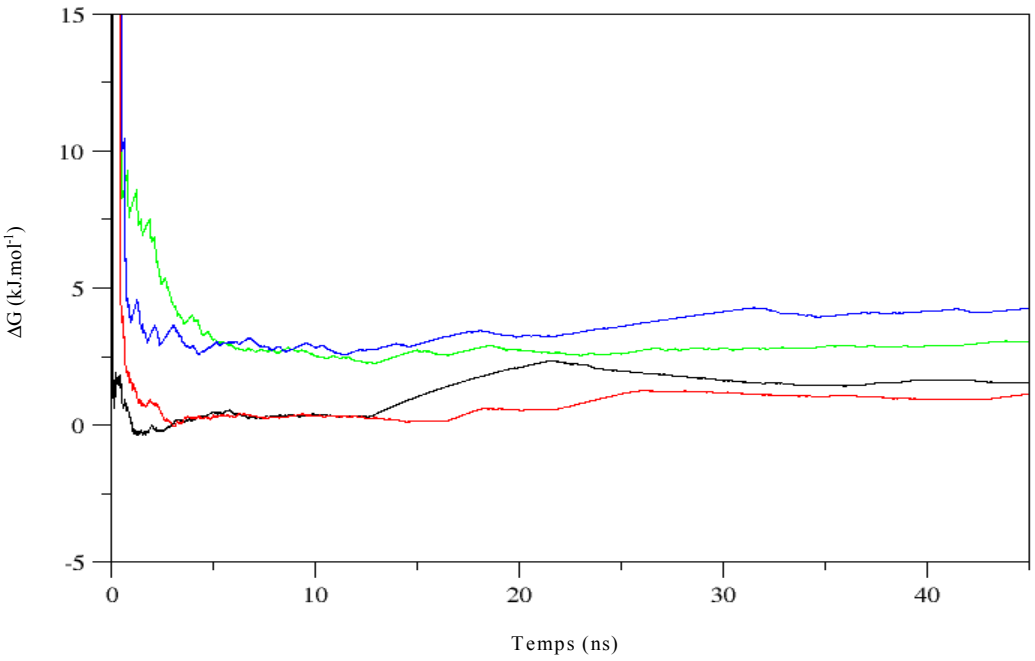


Figure 138 : ΔG en fonction du temps : en noir EGFEPG (T_1), en rouge EGFEPG (T_2), en vert EGFESG (T_1) et en bleu EGEFSG (T_2).

Les ΔG convergent au bout de 30 ns ce qui signifie qu'un équilibre, entre les conformations en coudes β de type VIII et celles qui ne le sont pas, est atteint. Les ΔG moyens à convergence sont reportés en **Tableau 43**.

Le peptide **EGFE<u>P</u>G** se replie de manière plus favorable en coude β de type VIII puisque il a un ΔG plus faible que celui du peptide **EGFE<u>S</u>G.**

	EGFE**P**G		EGF**S**G	
	Etendu (T_1)	Homologie (T_2)	Etendu (T_1)	Homologie (T_2)
Valeur de ΔG à convergence ($kJ.mol^{-1}$)	1,5	1,1	3,0	4,2

<u>Tableau 43</u> : ΔG à convergence pour les quatre simulations de DM.

B.3. Conclusions

Le peptide recombinant natif dérivant de la fibrilline 1 contenant une séquence **GXXPG**, c'est-à-dire ici **GFEPG**, possède expérimentalement la capacité d'induire une surexpression et une surproduction de MMP-1. Mais, la mutation de la séquence **GXX<u>P</u>G** en **GXX<u>S</u>G** inhibe totalement l'influence du même peptide sur la MMP-1. Nous avons pu voir que sous forme d'hexapeptide le peptide **EGFE<u>P</u>G** continuait d'avoir un effet activateur sur la production de MMP-1.

En étudiant sa structure par simulation de dynamique moléculaire, nous avons pu montrer qu'il possédait comme les EDPs un coude β de type VIII sur **GFEP** précédé par un autre coude β, ici de type II', sur la séquence **EGFE**.

A l'inverse, le peptide muté **EGFE<u>S</u>G** entraîne un comportement structural différent. Comme le peptide natif, il possède un coude β de type II' sur **EGFE**, mais en plus, il se replie sur **GFES** en coude β de type VIII et de type I, tandis qu'un nouveau coude β plutôt de type I apparaît sur **FESG**, ce qui induit la présence d'hélice α ou d'hélice $\beta 3_{10}$. La présence de ces types d'hélice n'a jamais été observée, tant dans le peptide **EGFEPG** que dans les EDPs précédemment étudiés.

La différence d'activité entre le peptide natif et le peptide muté est donc certainement due à cette différence de structure : les peptides actifs sont caractérisés par un coude β sur **XGXX** et un coude β de type VIII sur **GXX<u>P</u>**, tandis que **EGFE<u>S</u>G** a la particularité d'avoir le motif

GXXS en 2 types possibles de coudes β mais également la présence d'un autre coude β sur la séquence **XXSG**.

La différence entre le coude β de type I et celui de type VIII d'un motif $\mathbf{GX_1X_2X_3}$ est l'orientation de la liaison peptidique entre les résidus $\mathbf{X_2}$ et $\mathbf{X_3}$, ce qui entraîne une différence de position du groupement carbonyle du résidus $\mathbf{X_2}$ qui ne pointe pas dans la même direction (voir ci-dessous **Figure 139**).

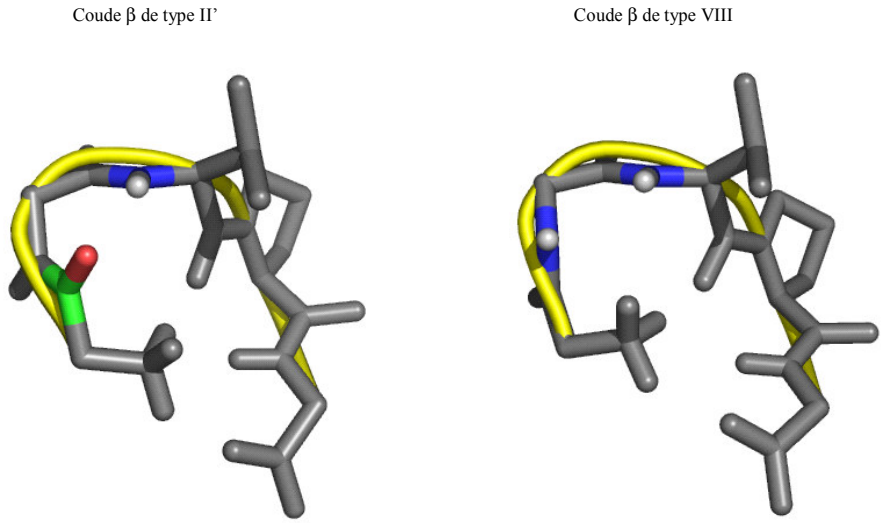


Figure 139 : Différence de positionnement de la liaison peptidique entre E et G du motif GFES suivant le type de coude β présent.

Le fait que **EGFESG** puisse adopter plusieurs types de coudes provoque une plus grande entropie et donc diminue les chances d'interagir avec son récepteur. La glycine peut être à l'origine également du manque d'activité du peptide, sa présence étant indispensable pour l'activité biologique des EDPs (Brassart et al. 2001). Dans les peptides actifs, la proline impose une certaine rigidité à la séquence **XXPG**, lui interdisant un repliement en coude β, mais permettant à la glycine d'être bien exposée. Dans la séquence **FESG**, cette glycine peut être désormais impliquée dans un coude β, mais cela l'empêche alors d'être accessible pour le récepteur. De la même manière, le remplacement d'une proline en sérine interdit, peut-être à cause du changement de nature de la chaîne latérale, l'interaction avec une poche du site de fixation du récepteur.

C. CONCLUSIONS GENERALES sur les LIGANDS de L'EBP

Les différentes méthodes théoriques (dynamique moléculaire, Monte-Carlo, carte adiabatique) utilisées ont montré que les peptides contenant le motif **GXXP** se repliaient de manière préférentielle en coude β de type VIII.

En comparant le comportement structural des peptides biologiquement actifs et inactifs, il a été possible de corréler leur activité biologique avec leur propension à se replier en coude β de type VIII, laissant penser que cette conformation est nécessaire pour se lier à l'EBP.

Deux sortes de peptides inactifs, c'est-à-dire ne se liant pas à l'EBP, ont été étudiées, les causes de leur inactivité semblant être différentes.

Le peptide **EGFESG**, provenant de la mutation (**P => S**) du peptide actif de la fibrilline-1 **EGFEPG**, à cause du remplacement de la proline, gagne de la flexibilité ce qui diminue ses chances d'être dans la bonne conformation pour interagir avec l'EBP. Le peptide **EGFESG** peut également être en coude β, non seulement de type VIII, mais également de type I.

L'autre peptide inactif étudié, le peptide **GGVPG**, peut adopter de manière faible deux types de coudes : le type VIII, mais également, de manière préférentielle le type II'. La différence structurale principale entre ces deux types de coudes β, provenant de la différence des valeurs d'angles dièdres, est l'orientation de la liaison peptidique entre les résidus **G** et **X** du motif **GXXP**. En effet, le groupement carbonyle de $\mathbf{G_2}$ pointe dans la même direction que le groupement amine de $\mathbf{V_3}$ dans le coude β de type II' alors que dans le type VIII, c'est le groupement amine de $\mathbf{G_1}$ qui pointe dans la même direction que le groupement amine de $\mathbf{V_3}$ (**Figure 140**). De plus, une Glycine au sein du motif **GXXP**, tout comme le remplacement de la proline par une sérine dans le peptide de fibrilline-1, augmente la flexibilité et l'entropie du système et donc diminue ses chances d'adopter la conformation adéquate. La proline du motif **GXXP** est certainement le résidu dont la présence est la plus importante car elle force le résidu précédent à posséder les angles dièdres du coude β de type VIII. Elle permet ainsi la bonne orientation des groupements amides et carboxyliques de ce dernier tout en lui imposant une certaine rigidité. De plus, elle-même étant le seul résidu cyclique, elle ne peut explorer qu'un petit nombre de conformations, renforçant également la rigidité locale du peptide, ce qui permet de diminuer son entropie. On peut d'ailleurs remarquer que de nombreux peptides d'élastine pouvant se lier à l'EBP possèdent une séquence **XPG** (Bisaccia et al. 1999).

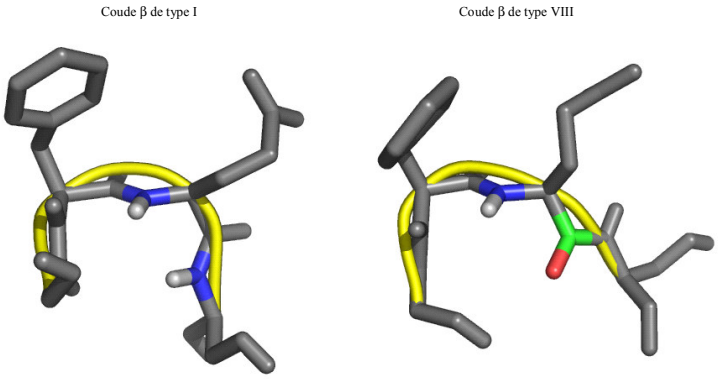


Figure 140 : Différence de position de la liaison peptidique située entre G et X du motif GXXP pour des coudes β de types II' et VIII.

En définitive, les peptides d'élastine possédant un motif **GX_1X_2P** doivent pour se lier à l'EBP, posséder plusieurs propriétés. Ils doivent pouvoir se replier en coude β de type VIII, ce qui permet d'avoir deux groupements amides pointant dans la même direction, ceux des résidus centraux **X_1** et **X_2**, et qui permet également au groupement carbonyle de **X_2** de pointer lui aussi dans la même direction (**Figure 145**). Ces groupements sont certainement impliqués dans des liaisons hydrogène avec l'EBP. Les peptides actifs doivent se replier de manière quasi exclusive en coude β de type VIII car, les autres types ne permettent pas les mêmes orientations des liaisons peptidiques et surtout l'entropie augmentant, les chances d'interaction diminuent. De même le résidu glycine, étant un résidu très flexible, de part sa présence en position centrale du motif **GX_1X_2P** augmente également l'entropie du peptide et son absence en position centrale est donc préférable. Enfin, une glycine après le motif **GX_1X_2P** semble être indispensable pour la fixation à l'EBP, mais ne présente pas de rôle structural sur ce motif. Sa présence est requise car elle favorise certainement la fixation du peptide à l'EBP, en interagissant avec une poche pour laquelle elle possède des affinités physico-chimiques. Il semble donc que pour cela, elle doit la plus exposée possible au solvant.

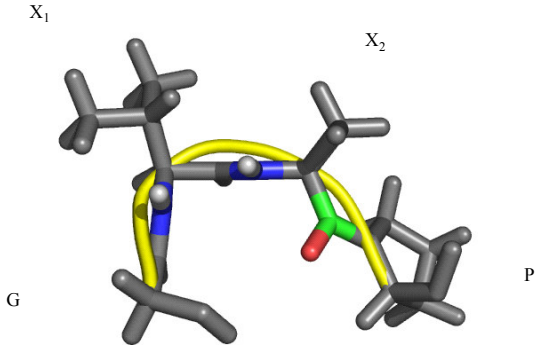


Figure 141 : Coude β de type VIII.

Partie V

INTERACTIONS LIGANDS - RECEPTEUR

Il a été démontré qu'un même anticorps reconnaissait de manière spécifique la séquence permettant la fixation de élastine à l'EBP, et celle de l'élastase pancréatique porcine (Hinek et Rabinovich 1994). Cette séquence provenant de l'EBP sous forme peptidique, appelée peptide Sgal, possède toujours la propriété de se fixer à l'élastine.

Nous avons donc tout d'abord étudié la fixation de certains peptides dérivant de l'élastine à l'élastase pancréatique porcine (PPE pour Porcine Pancreatic Elastase), puis l'interaction entre certains peptides dérivant de l'élastine et le peptide Sgal.

A. INTERACTION PPE / EDP

A.1. Etude par amarrage moléculaire :

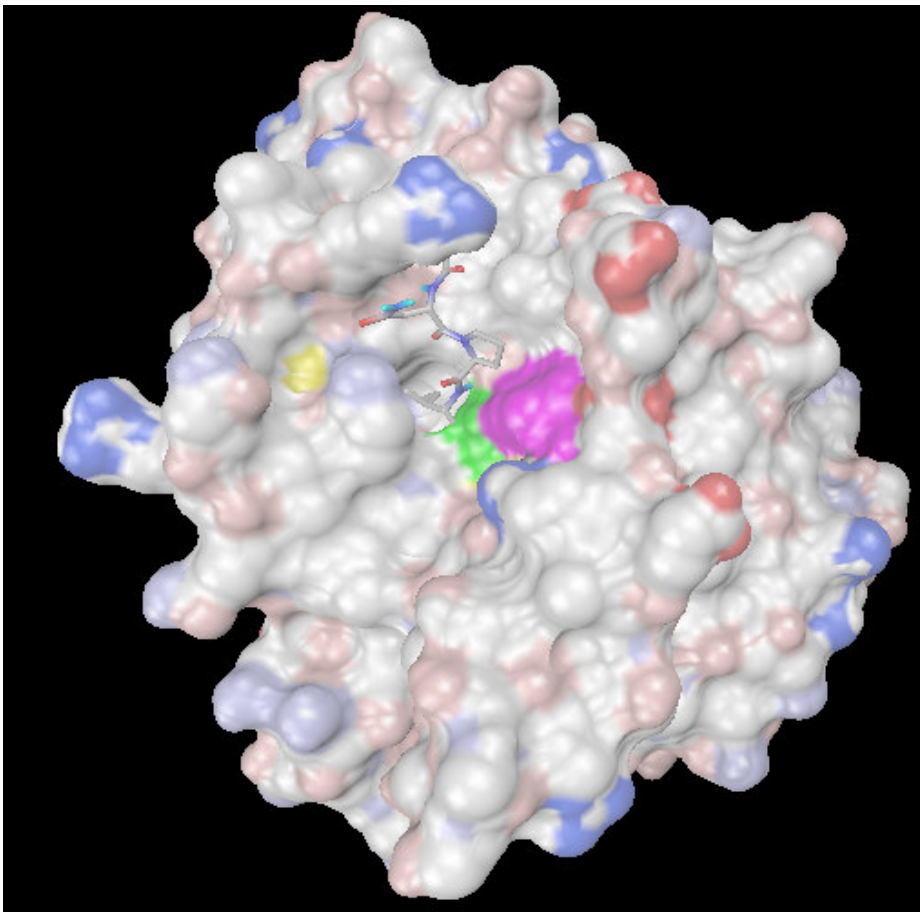
A.1.a Vérification du protocole d'amarrage moléculaire utilisé :

Pour effectuer les calculs d'amarrage moléculaire, nous avons utilisé la version 3.0 du programme AutoDock (Morris et al. 1998), avec la méthode d'échantillonnage des conformations la plus récemment implémentée, basée sur l'algorithme génétique de Lamarck. Nous avons tout d'abord plongé entièrement la PPE dans une grille suffisamment grande pour que les ligands testés, quelle que soit leur conformation, ne puissent en sortir, tout en ayant la maille la plus faible pour que le calcul soit le plus précis possible. Ensuite, les endroits les plus favorables, c'est-à-dire ceux où les ligands possédaient les plus faibles énergies d'interaction, ont été à l'origine d'une nouvelle grille de taille plus restreinte avec un maillage plus petit, pour avoir un calcul le plus précis possible.

Pour vérifier si le protocole que nous souhaitons employer était correct, nous avons essayé de retrouver l'emplacement où l'inhibiteur se fixe à la PPE. Nous avons choisi la structure 3D de la PPE présente dans la Protein Data Bank sous la référence **1GVK**, cette structure ayant été résolue avec une précision de 0,95 Å (Katona et al. 2002).

Le site actif des enzymes de la famille des sérine-protéinases est composé d'une triade catalytique, formée par les résidus sérine-histidine-aspartate, qui permet la rupture de la liaison peptidique de certaines protéines, dont l'élastine (Blow et Steitz 1970). L'inhibiteur co-

cristallisé avec la PPE est un ligand tripeptidique N-acétylé : N-acétyl-N-P-I (**Figure 142**). Son activité inhibitrice est due à formation au niveau du site actif d'une liaison covalente avec la $\mathbf{S_{195}}$ responsable de la coupure de la liaison peptidique.

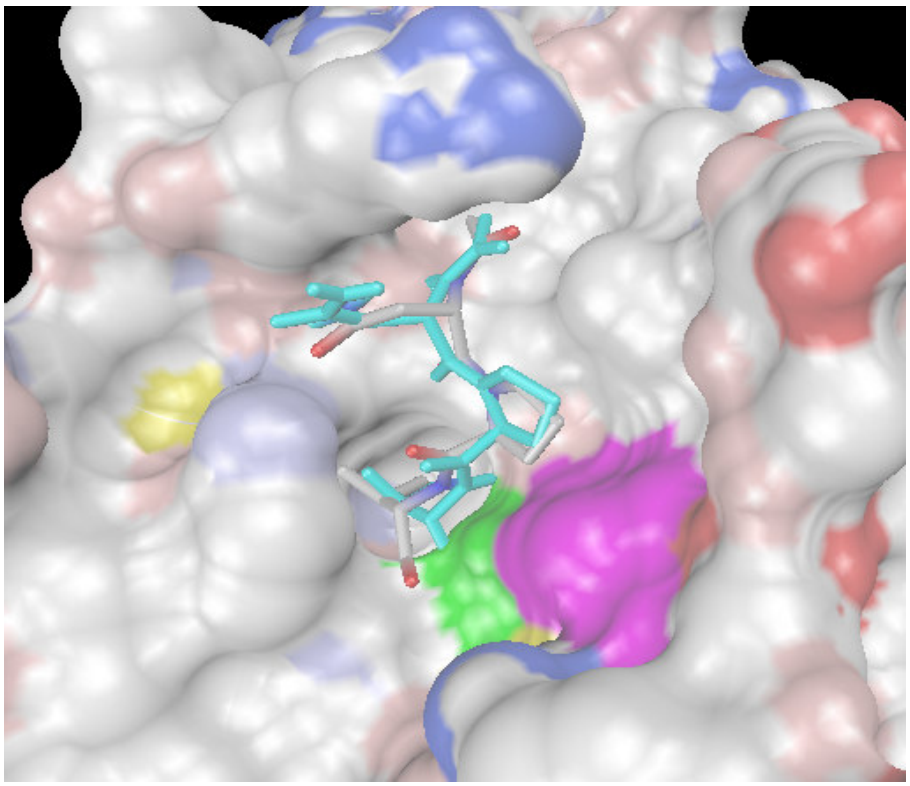


<u>Figure 142</u> : Représentation de la PPE (1GVK) avec sa surface accessible au solvant.
Les couleurs sont en fonction de la polarité des atomes (rouge : +, bleu : -).
L'histidine 57 est coloriée en violet et la sérine 195 en vert.
L'inhibiteur est représenté en *bâtons* et les couleurs fonction du type d'atome.

Les charges partielles de chacun des atomes du ligand ont été calculées avec le module `Ampac/Mopac` du programme INSIGHT II (Accelrys Inc) avec la méthode AM1. La PPE a été plongé dans une grille parallélépipédique de 126 × 110 × 126 points avec une maille de 0,45 Å. L'algorithme génétique de Lamarck a été configuré pour avoir au départ 50 individus et un maximum de 500.000 générations par cycle. Un cycle de 250 calculs d'amarrage moléculaire a permis de localiser les meilleurs ligands dans un voisinage très proche du site de protéolyse de la PPE. Une autre grille cubique de 126 × 126 × 126 points, a été utilisée

dans cette zone restreinte du système avec un maillage de 0,2 Å. Un nouveau cycle de 250 tirages a été effectué.

Le modèle finalement obtenu et présentant la plus basse énergie d'interaction, est représenté en **Figure 143**. Il est parfaitement placé et a une conformation très proche de celle obtenue expérimentalement, avec une RMSD inférieure à 1,5 Å. La seule différence marquée est la rotation de l'angle χ_1 du résidu $\mathbf{N_1}$. La chaîne latérale de I_3 est, quant à elle, bien placée : dans la même poche que celle du ligand provenant de la structure cristallisée, tout comme le groupement méthyle de la fonction N-acétyle porté par le résidu $\mathbf{N_1}$. Les distances entre les atomes de l'inhibiteur et ceux de la PPE dans le modèle obtenu par les calculs d'amarrage sont très voisines de celles obtenues expérimentalement (**Tableau 44**). La petite différence toutefois obtenue est vraisemblablement due au fait que l'inhibiteur se lie à la PPE grâce à une liaison covalente entre l'atome de carbone de la fonction carboxylique de I_3 de l'inhibiteur et l'atome d'oxygène de la fonction alcool de la $\mathbf{S_{195}}$ de la PPE.



<u>Figure 143</u> : Site actif de la PPE (1GVK) représentée avec sa surface accessible au solvant. Les inhibiteurs sont représentés en *bâtons*, avec en couleur cyan, l'inhibiteur cristallisé avec la PPE et en couleur dépendant de la nature des atomes, l'inhibiteur obtenu après les calculs d'amarrage moléculaire.

Le programme AutoDock est, bien sûr, incapable de créer des liaisons covalentes, il en résulte un encombrement stérique à ce niveau qui tend à éloigner le résidu **I_3** de la PPE.

Interaction			Distance (Å)	
Inhibiteur		PPE	Expérimentale (1GVK)	Théorique (Amarrage moléculaire)
I_3–N	->	O–S_{214}	**3,04**	**3,54**
N_1–O	->	N–V_{216}	**2,97**	**2,91**
I_3–O	->	N–G_{193}	**2,85**	**2,82**
I_3–0	->	N–G_{195}	**2,72**	**2,70**

Tableau 44 : Distance entre certains atomes de l'inhibiteur et de la PPE pour la structure expérimentale et pour le modèle théorique obtenu après les calculs d'amarrage moléculaire.

Ces résultats montrent que le protocole et le programme utilisés sont bien adaptés à l'étude de la PPE et des ses ligands peptidiques.

A.1.b. Amarrage moléculaire des EDP sur la PPE :

Les charges partielles des EDP d'activité connue **VGVAPG**, **PGAIPG**, **GLVPG**, **GVAPG**, **GGVPG**) ont été calculées avec la méthode AM1 du module Ampac/Mopac d'INSIGHT II. Pour chaque peptide, 4 cycles de 250 calculs d'amarrage moléculaire chacun, ont alors été effectués.

Toutes les meilleures solutions proposées par le programme AutoDock sont localisées au même endroit et dans pratiquement la même configuration (**Figure 144**). L'EDP en effet adopte une structure contenant un coude β de type VIII déformé sur la séquence **GX_1X_2P**. Ce coude β est déformé dans le sens où le critère de distance entre le C_α de **G** et le C_α de **P** est légèrement supérieur à 7,0 Å (**Tableau 45**).

En fait, le coude β de type déformé de **GX_1X_2P** entoure le résidu **Q_8**, tout en formant 3 liaisons hydrogène impliquant différents atomes du squelette peptidique :

- entre le **Q_8-NH** de la PPE et le **X_2-C=O** de l'EDP
- entre le **S_6-C=O** de la PPE et le **X_2-NH** de l'EDP
- entre le **S_6-C=O** de la PPE et le **X_1-NH** de l'EDP

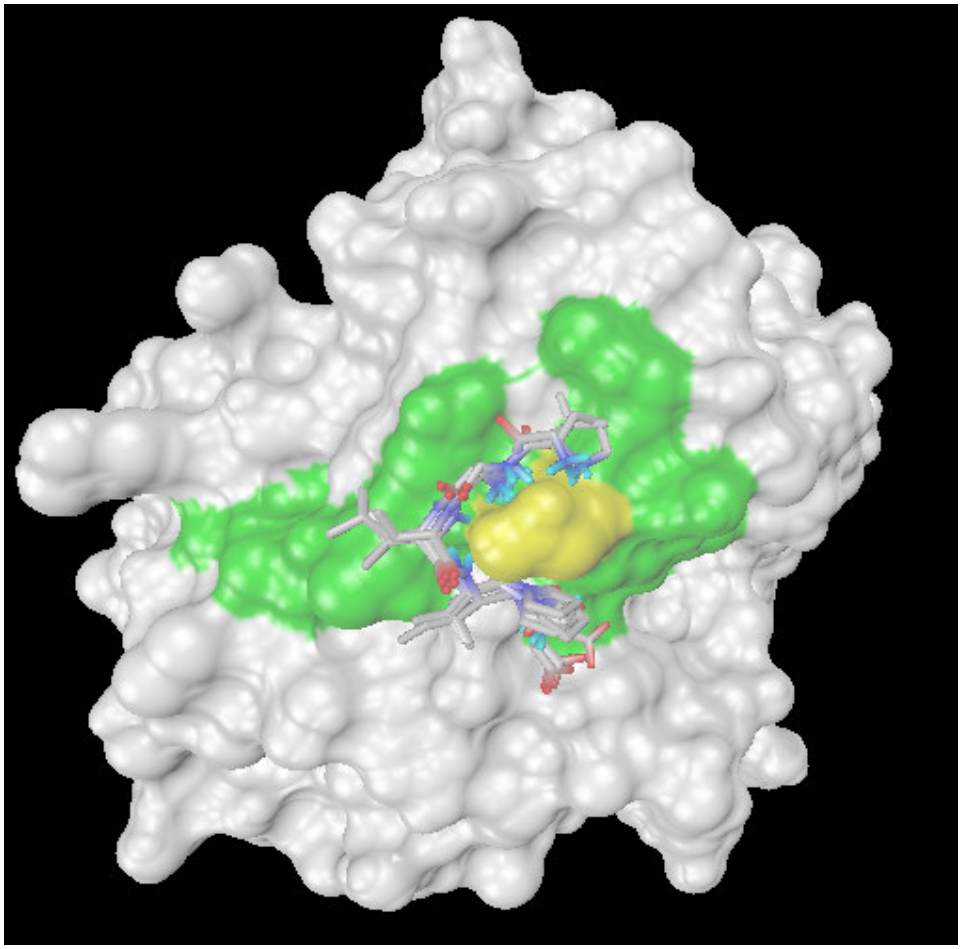


Figure 144 : Localisation du meilleur candidat proposé par AutoDock pour chaque EDP. Les EDP sont représentés en "bâtons" et leurs atomes sont colorés en fonction de leur type; la PPE est représentée avec sa surface accessible au solvant, et en vert les 15 premiers résidus correspondant au peptide Sgal et en jaune, le résidu Q_8.

séquence (X)GX$_1$X$_2$PG	Résidus X$_1$		Résidus X$_2$		Distance Cα G - Cα P (Å)
	ϕ (°)	ψ(°)	ϕ(°)	ψ(°)	
VGVAPG	– 112,3	– 38,6	– 117,3	127,6	7,2
PGAIPG	– 115,6	– 45,9	– 138,6	114,5	7,8
GVAPG	– 98,5	– 52,8	– 128,1	119,1	7,3
GLVPG	– 103,5	– 64,5	– 133,5	142,8	7,5
GGVPG	– 93,0	– 65,0	– 124,5	132,5	7,3

Tableau 45 : Caractéristiques des meilleures solutions proposées par AutoDock pour chacun des peptides

Il est intéressant de constater que ces 3 liaisons hydrogène impliquent les atomes des liaisons peptidiques et que leur direction est caractéristique d'un coude β de type VIII. Le **NH** des résidus $\mathbf{X_2}$ et $\mathbf{X_1}$, ainsi que le **C=O** du résidu $\mathbf{X_2}$ pointent tous les trois dans le même sens comme cela avait été observé lors de l'étude structurale des ligands peptidiques de l'EBP.

Le peptide **GGVPG**, bien qu'inactif biologiquement, semble pouvoir se fixer à la PPE comme les autres EDPs actifs. Il peut donc adopter une conformation favorable ayant les caractéristiques structurales requises pour se lier à la PPE. On peut donc penser que c'est plutôt sa fixation qui serait défavorisée à cause de l'augmentation de son entropie. En effet, la glycine étant extrêmement flexible, elle favorise de multiples conformations et diminue d'autant les chances qu'à le peptide de prendre la structure adéquate interagir avec la PPE, mais par là même également avec l'EBP.

A.2. Dynamique moléculaire de la PPE avec le peptide VGVAPG :

Les simulations de dynamique moléculaire de la PPE ont été réalisées avec la version c27b3 du programme CHARMM (Brooks et al. 1983) avec le jeu de paramètres "tout atome" PARAM22 (MacKerell et al. 1998). Le protocole de simulation (interactions non liantes, conditions périodiques aux limites, minimalisation, chauffage et équilibrage) est le même que celui précédemment utilisé. La seule différence est le recours à l'algorithme SHAKE pour fixer la distance des liaisons entre les atomes lourds et les atomes d'hydrogène, ce qui nous perme d'utiliser un pas d'intégration de 0,2 fs pour toutes les phases de la simulation (chauffage, équilibrage et production).

Afin de tester et valider notre protocole de simulation, mais aussi d'évaluer l'incidence de la présence du ligand sur la structure de la PPE, nous avons déjà réalisé une simulation de dynamique moléculaire de la PPE seule. La simulation du complexe PPE-**VGVAPG** a ensuite été effectuée avec la structure de plus basse énergie obtenue pour ce complexe avec les calculs d'amarrage moléculaire.

A.2.a. Dynamique moléculaire de la PPE seule

La PPE a été plongée dans une boîte parallélépipédique de dimension 70 Å x 65 Å x 62 Å contenant 9.451 molécules d'eau TIP3P (Jorgensen et al. 1983). 1.285 molécules d'eau à moins de 2,8 Å d'un atome lourd de la PPE ont été ôtées, donnant ainsi un système constitué d'un total de 27.940 atomes. La phase de production a une durée de 2 ns.

Durant l'analyse, nous avons vérifié que la PPE ne perdait pas sa structure en étudiant l'évolution temporelle des structures secondaires présentes initialement à l'aide du programme STRIDE (**Figure 145).**

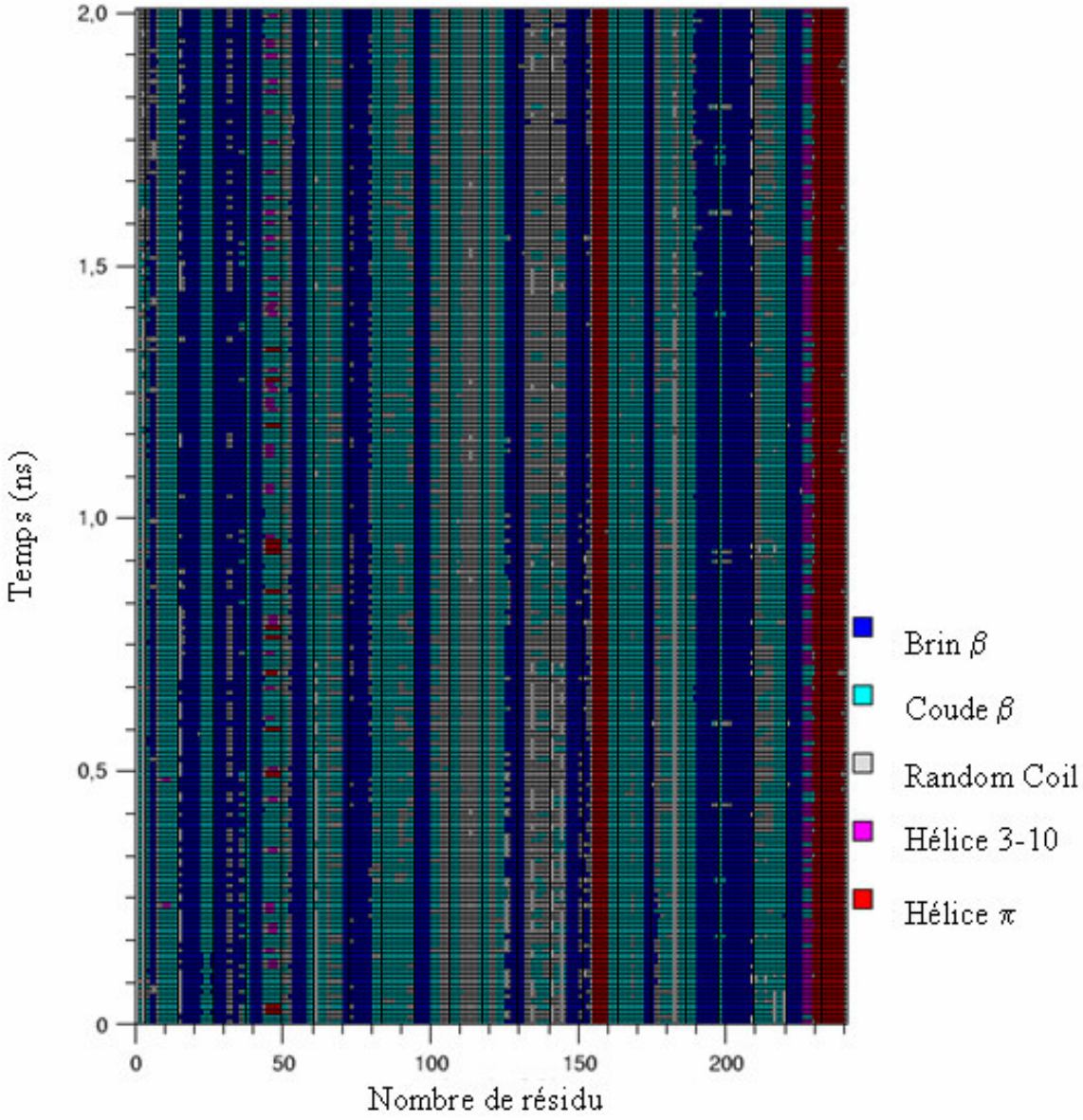


Figure 145 : Suivi des structures secondaires en fonction du temps au cours de la simulation de DM de la PPE seule.

On peut s'apercevoir que les structures secondaires de la PPE sont bien conservées au cours de la trajectoire. Du côté N-terminal, correspondant à la partie ayant de fortes homologies avec le peptide Sgal, les 4 premiers résidus perdent leur conformation en coude β à partir de 1,5 ns de trajectoire.

Le suivi de la position des types de coudes de la séquence correspondant au peptide Sgal, c'est-à-dire les 15 premiers résidus du côté N-terminal confirme cette observation (**Figure 146**).

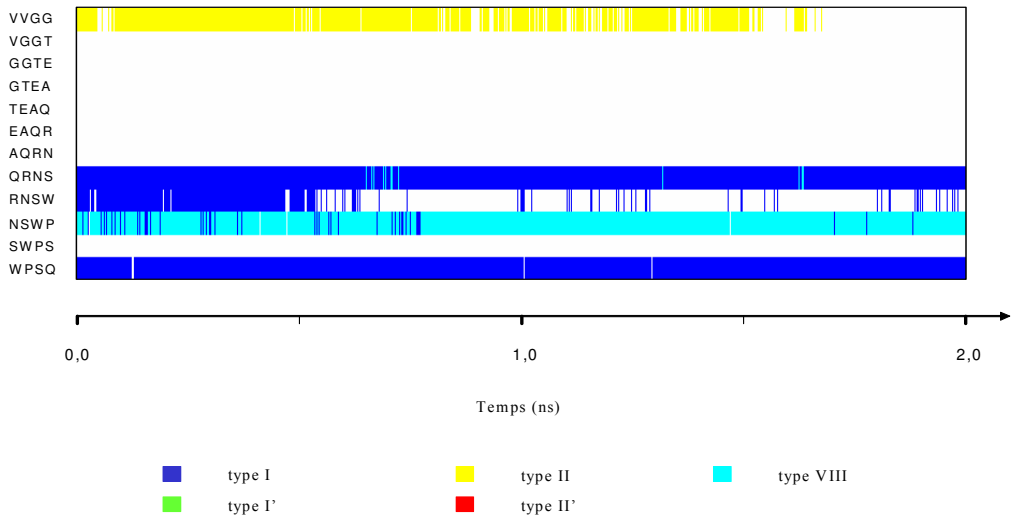


Figure 146 : Présence des différents types de coudes β au cours de la DM de la PPE seule.

En effet, il apparaît clairement, qu'après 1,5 ns de trajectoire, le coude β de type II sur **V_1VGG_4** disparaît, tout comme les coude β de type I sur **RNSW** qui n'est présent que transitoirement après 0,5 ns. A l'inverse, 3 coudes β restent très stables au cours du temps : sur les séquences **Q_8RNS_{11}**, **$N_{10}SWP_{13}$** et **$W_{12}PSQ_{15}$** (**Tableau 46**).

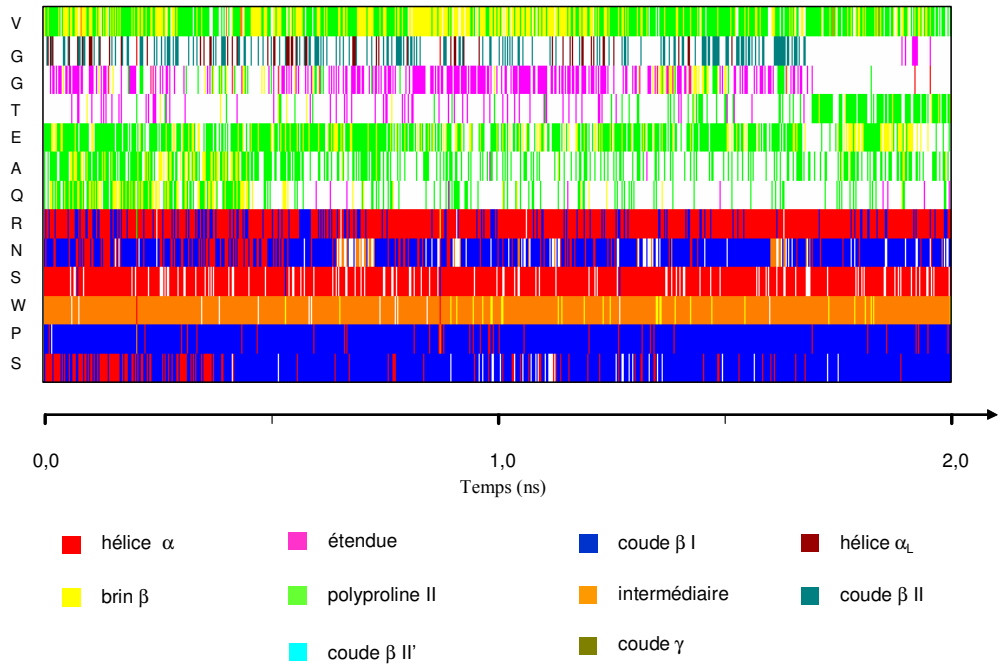
Séquences	Type de coude β canoniques ou déformés (% de conformations)		Total
	I	VIII	
QRNS	98,6	1,4	100,0
NSWP	5,6	93,8	99,4
WPSQ	99,5	-	99,5

Tableau 46 : Pourcentages de coudes β sur les séquences QRNS, NSWP et WPSQ au cours de la DM de la PPE seule.

Il est intéressant également de noter que le coude β sur **QRNS** de type I était initialement de type II, dans la structure expérimentale de la PPE. La séquence équivalente à **QRNS** dans le peptide Sgal est **QDEA** qui est le coude β le plus observé durant les dynamiques moléculaires, et qui est également de type I. Le comportement du motif **QRNS** de la PPE est

donc le même que son équivalent, **QDEA**, dans le peptide Sgal, ce qui laisse à penser qu'un coude β de type I à cet endroit dans la séquence est extrêmement favorisé et pourrait donc être important pour l'activité : la fixation des EDP.

L'analyse des différentes zones de la carte de Ramachandran visitées par les 15 différents résidus composant la séquence N-terminale, montre qu'à partir de **Q_8** inclus, chaque résidu est très contraint dans une seule zone propre (**Figure 147**). A l'inverse, de manière prévisible, les résidus les précédant étant en bout de chaîne peptidique, sont beaucoup plus flexibles et explorent différentes régions de la carte de Ramachandran, expliquant ainsi la perte du coude β de type II sur les résidus **VVGG**.

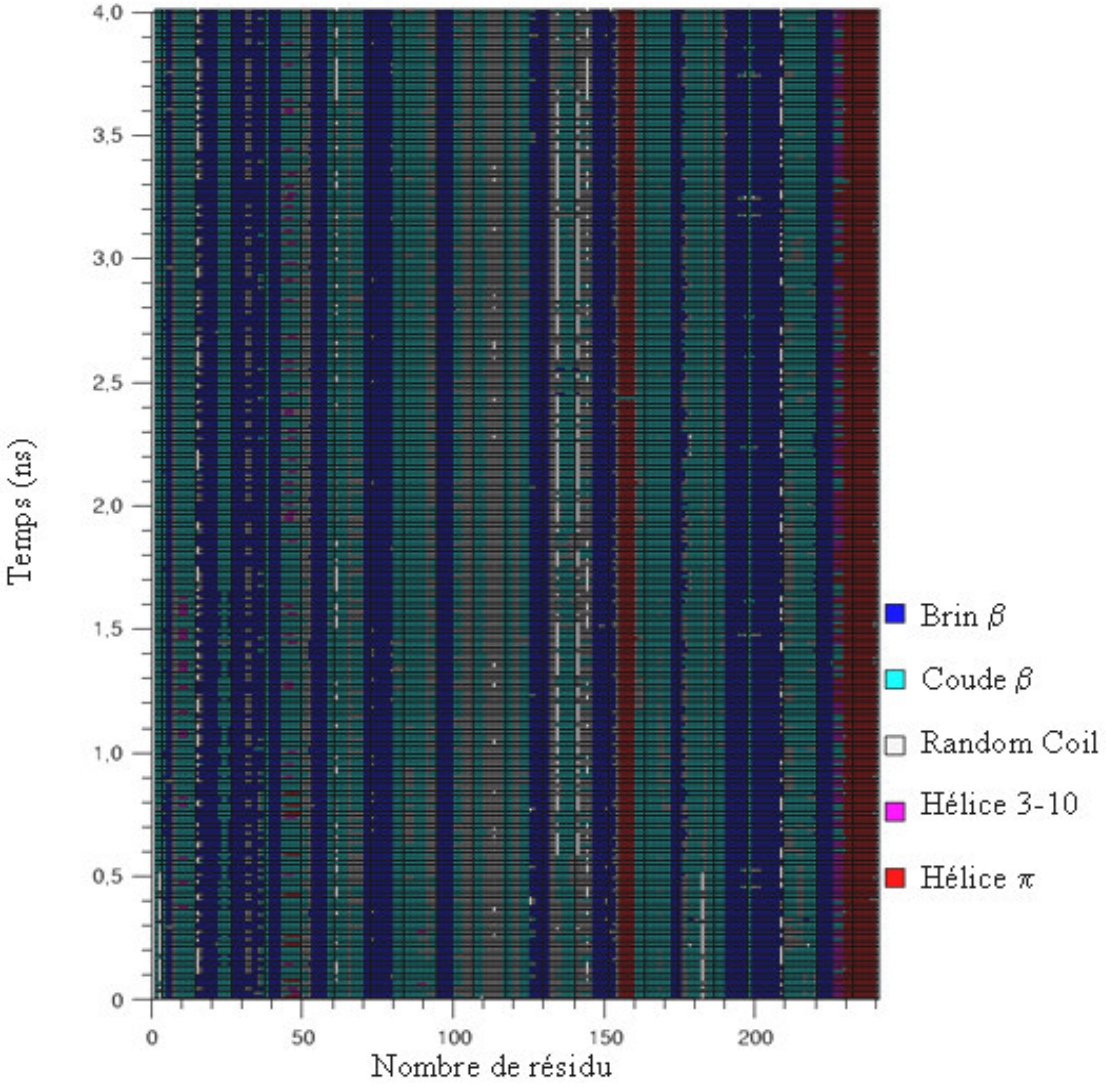


<u>Figure 147</u> : Zones explorées par les 15 premiers résidus au cours de la trajectoire de la dynamique moléculaire de PPE seule.

A.2.b. Dynamique moléculaire de la PPE en interaction avec VGVAPG :

La PPE en interaction avec la meilleure position du peptide **VGVAPG** obtenue par amarrage moléculaire a été plongée dans une boîte d'eau parallélépipédique de dimension 70 Å × 65 Å × 62 Å. Après le retrait de 1.359 molécules d'eau situées à moins de 2,8 Å de la PPE ou du peptide **VGVAPG**, le système final est constitué de 27.839 atomes. La phase de production a duré 4 ns.

Comme le montre la **Figure 148**, les hélices α, les brins β ou les coudes β de la PPE sont conservés durant toute la trajectoire, aucune modification d'importance des structures secondaires n'étant visible.



<u>Figure 148</u> : Suivi des structures secondaires en fonction du temps au cours de la simulation de la PPE en interaction avec le peptide VGVAPG.

La présence des coudes β au cours de la trajectoire des 15 premiers résidus du côté N-terminal correspondant à la séquence du peptide Sgal est reportée en **Figure 149**.

On peut remarquer que contrairement à la dynamique de la PPE seule, le coude β de type II sur $\mathbf{V_1VGG_4}$ reste présent tout au long de la trajectoire. De même, le coude β de type I sur **RNSW** est visible pendant plus longtemps, mais finit par disparaître après 1,7 ns de simulation. Les 3 coudes β sur $\mathbf{Q_8RNS_{11}}$, $\mathbf{N_{10}SWP_{13}}$ et $\mathbf{W_{12}PSQ_{15}}$ reste en revanche toujours

aussi stables que durant la trajectoire de le PPE seule (**Tableau 47**). Ce résultat confirme l'importance du coude β de type I sur **Q_8RNS$_{11}$** puisque même lorsque **VGVAPG** se place autour de la chaîne latérale de la **Q_8**, ce motif structural reste inchangé.

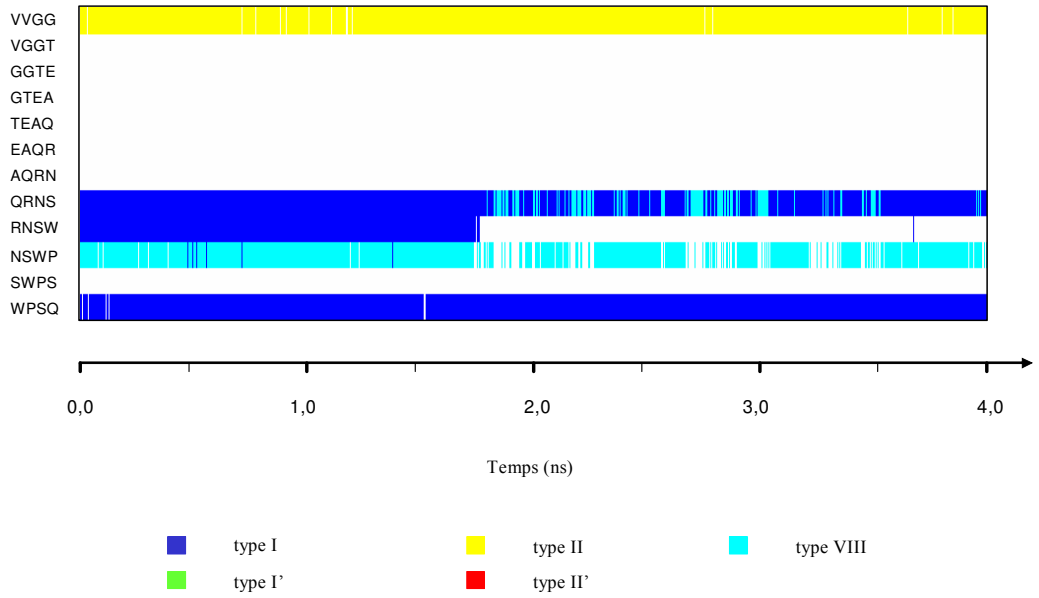


Figure 149 : Présence des différents types de coudes β au cours de la trajectoire de la PPE en interaction avec le peptide VGVAPG.

Séquences	Types de coudes β canoniques ou déformés (%)			total
	I	II	VIII	
VVGG	-	98,4	-	98,4
QRNS	85,8	-	14,2	100,0
NSWP	3,8	-	83,0	83,8
WPSQ	99,4		-	99,4

Tableau 47 : Pourcentages de coudes β sur les séquences VVGG, QRNS, NSWP et WPSQ au cours de la trajectoire de la PPE en interaction avec le peptide VGVAPG.

La présence du peptide **VGVAPG** contraint le résidu **E_6** et surtout les résidus **A_7** et **Q_8** à rester dans la zone de la polyproline II, tandis que les autres résidus ne présentent pratiquement pas de différence quant aux zones qu'ils occupent par rapport à ce que l'on observe pour la protéine seule (**Figure 150**).

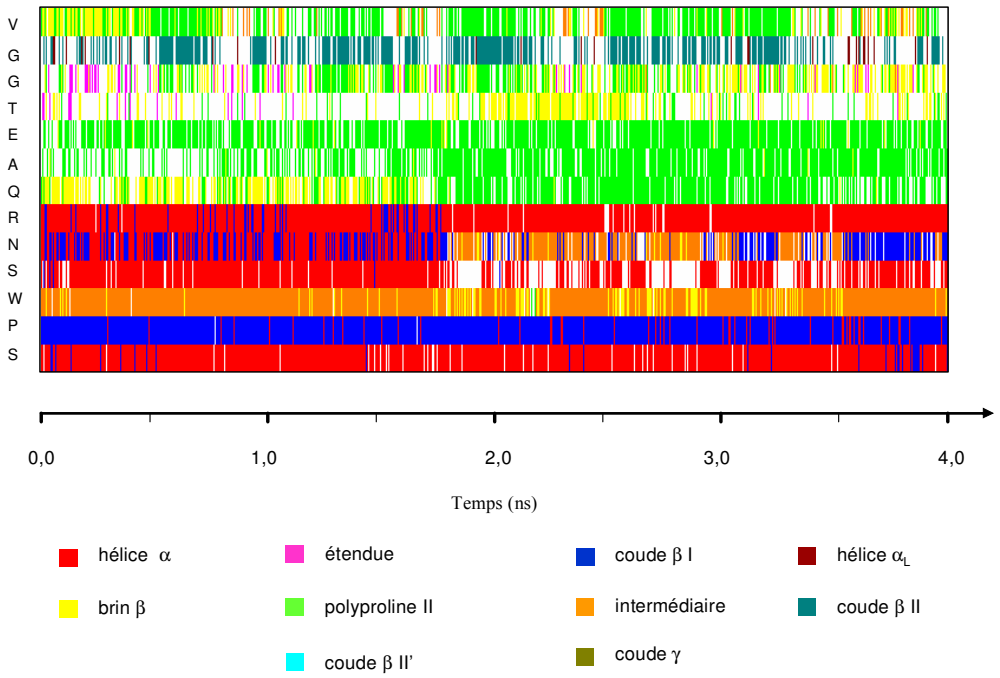


Figure 150 : Zones explorées par les 15 premiers résidus de la PPE au cours de la DM de la PPE en interaction avec le peptide VGVAPG.

Lorsque l'on s'intéresse maintenant au comportement du peptide **VGVAPG** en interaction avec la PPE, on s'aperçoit que sa structure reste très stable au cours du temps.

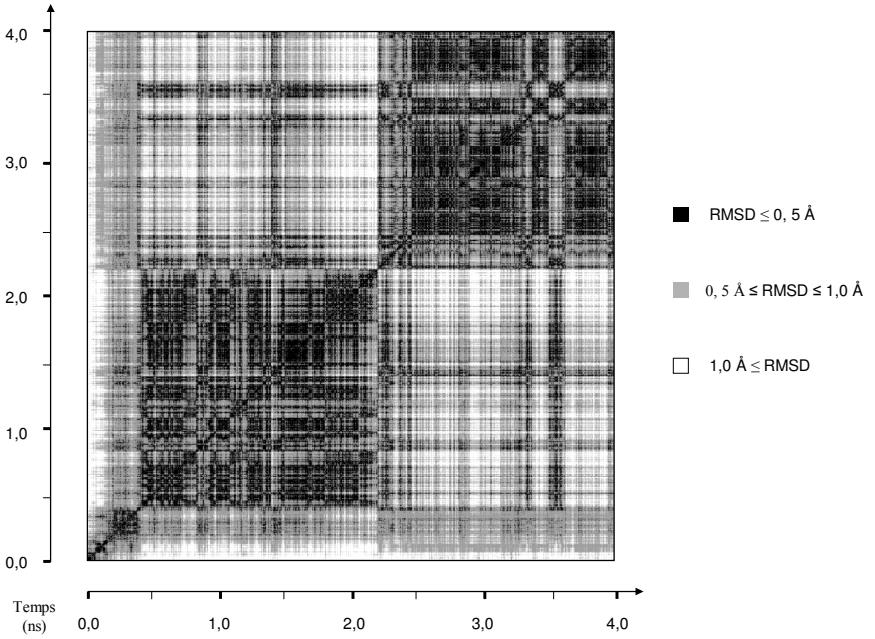
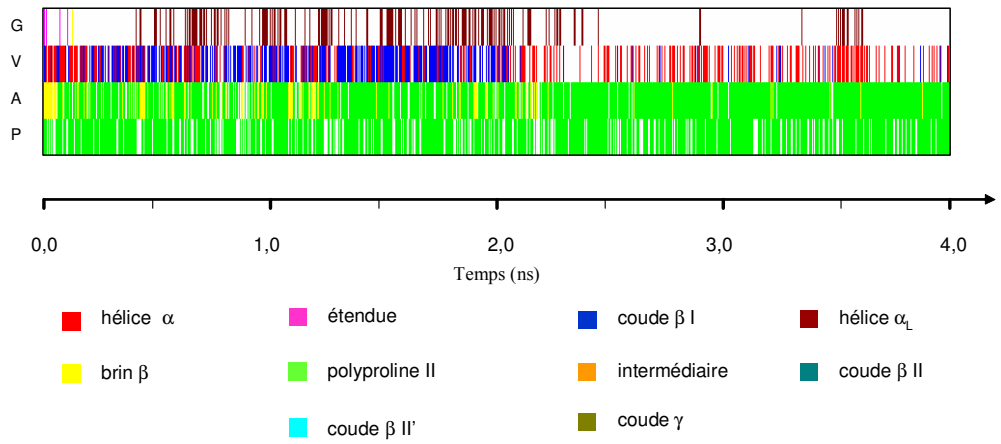


Figure 151 : Carte des RMSD croisés durant la DM de la PPE en interaction avec. VGVAPG

La carte de RMSD croisés, calculée sur le squelette peptidique du peptide seul, montre que les différentes conformations sont très proches les une des autres, c'est-à-dire avec une valeur de RMSD moyenne inférieure à 1,0 Å (voir ci-dessus en **Figure 151**). Cependant, un premier amas de conformations est tout de même visible jusqu'à 2,2 ns de simulation, un deuxième amas est présent de 2,2 ns à la fin de la trajectoire. Les deux amas de **VGVAPG** présentent un coude β déformé, c'est-à-dire une distance C_α-C_α entre $\mathbf{P_5}$ et $\mathbf{G_2}$ supérieure à 7,0 Å. La seule différence existant entre les deux amas provient de l'orientation de $\mathbf{V_3}$.

Bien que les valeurs des angles dièdres des résidus $\mathbf{V_3}$ et $\mathbf{A_4}$ soient très stables et possèdent les valeurs adéquates pour avoir un coude β de type VIII (**Figure 152**), seules 4,8 % des conformations du peptide **VGVAPG** sont effectivement en coude β de type VIII. On peut s'apercevoir également qu'au début de la trajectoire, le résidu $\mathbf{V_3}$ est dans la zone du coude β de type I (ϕ = - 90°; ψ = 0°) puis à partir 2,2 ns, il passe dans la zone la plus favorable à un coude β de type VIII : la zone de l'hélice α (ϕ = - 60°; ψ = - 30°). Ce réarrangement est également visible sur la carte des RMSD croisés précédente : un nouvel amas de conformations proches apparaît à cette même période.

Aucun coude β n'est en revanche observable sur les motifs **VGVA** ou **VAPG**.



<u>Figure 152</u> : Zones explorées par les résidus au cours de la trajectoire de la PPE en interaction avec VGVAPG

Le peptide **VGVAPG** continue en fait d'avoir une structure proche d'un coude β de type VIII, seule la distance C_α ($\mathbf{G_2}$) - C_α ($\mathbf{P_5}$) supérieure à 7,0 Å (**Figure 153**), empêche une telle classification. On peut cependant aussi observer qu'après 2,2 ns, le peptide a subi un

réarrangement de sa structure car d'une distance de 8 Å en moyenne on passe ensuite à une distance d'environ 7,2 Å.

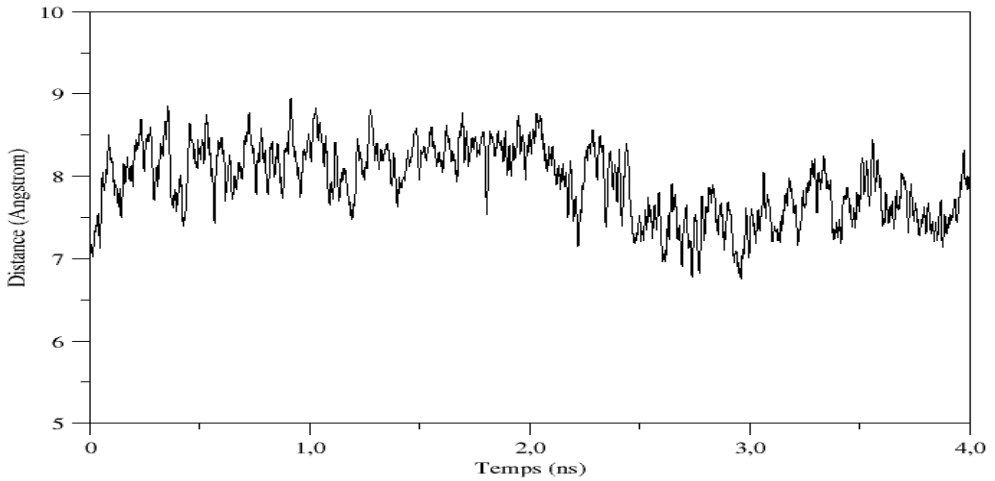


Figure 153 : Distance C_a (G_2) - C_a (P_5) du peptide VGVAPG en fonction du temps, au cours de sa dynamique en interaction avec la PPE.

Trois liaisons hydrogène étaient présentes entre la PPE et le peptide **VGVAPG** dans le modèle obtenu par amarrage moléculaire. Seules deux persistent tout au long de la trajectoire : entre **G_2-NH** du peptide **VGVAPG** et **E_6-C=O** de la PPE, et entre **V_4-NH** du peptide VGVAPG et **E_6-C=O** de la PPE. Celles-ci sont présentes dans respectivement 50,6 % et 95,8 % des conformations analysées (**Figure 154**). La liaison hydrogène, entre **G_2-NH** du peptide **VGVAPG** et **E_6-C=O** de la PPE, devient en outre plus stable après 2,2 ns de trajectoire.

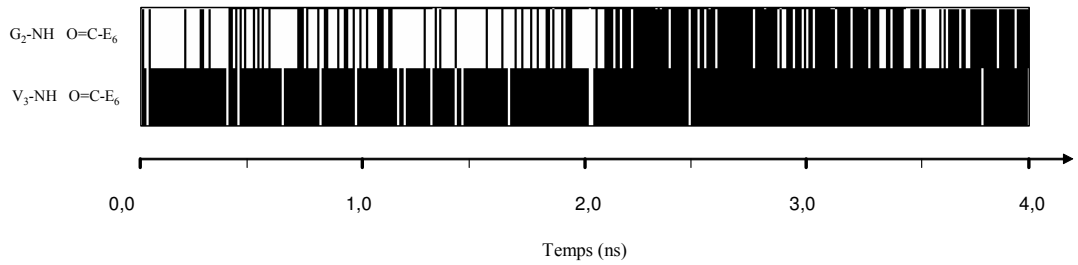


Figure 154 : Liaison hydrogène entre VGVAPG et la PPE au cours de la DM.

A.3. Conclusions

Grâce aux calculs d'amarrage moléculaire sur la PPE de certains des peptides dérivant de l'élastine contenant un motif **GXXP**, nous avons pu montrer que la partie de 15 résidus située du côté N-terminal de la PPE et correspondant au peptide Sgal avait une affinité particulière pour ces derniers.

La structure des EDPs en interaction avec la PPE est caractérisée par un coude β déformé de type VIII sur **GXXP**, car la distance Cα–Cα entre **G** et **P** est légèrement supérieure à 7,0 Å.

Pendant 4 ns de dynamique moléculaire, la PPE et le peptide **VGVAPG** restent fixés l'un à l'autre, sans changement de conformation important, démontrant qu'ils ont bien une affinité structurale réciproque. Nous avons pu également observer un réarrangement structural du peptide **VGVAPG**, le conduisant à adopter une conformation encore plus proche de celle du coude β de type VIII sur **GVAP**.

Même si elle méritait d'être prolongée, la simulation de dynamique moléculaire a montré que nous avons très certainement obtenu par amarrage moléculaire l'endroit et la structure nécessaire au peptide **VGVAPG** pour interagir avec la PPE. Des preuves indirectes vont dans ce sens. Tout d'abord, l'étude des peptides dérivant de l'élastine a montré qu'ils adoptaient préférentiellement un coude β de type VIII sur **GXXP**, c'est-à-dire pratiquement le même motif structural que celui obtenu par les calculs d'amarrage moléculaire de ces peptides sur la PPE. De plus, l'étude du peptide Sgal a montré qu'un coude β de type I sur **QDEA** était le motif le plus stable au cours du temps. C'est ce même coude β qui permet au résidu $\mathbf{Q_8}$ d'être suffisamment exposé pour permettre au peptide **VGVAPG** de se lier autour.

B. INTERACTION PEPTIDE Sgal / EDP :

Il a été démontré que le peptide Sgal pouvait se fixer à l'élastine (Hinek et al. 1993). De même il a été démontré que les EDP contenant, notamment ceux contenant un motif **GXXP**, pouvaient se lier à l'EBP et ainsi induire de nombreuses activités biologiques (Senior et al. 1984; Grosso et Scott 1993b; Bisaccia et al. 1994; Castiglione-Morelli et al. 1997; Brassart et al. 2001). Toutefois, rien à l'heure actuelle ne prouve que le peptide Sgal et un EDP puissent interagir ensemble. Nous avons donc étudié expérimentalement par spectroscopie de DC et de fluorescence l'interaction entre le peptide Sgal et les EDP. *In silico*, par amarrage moléculaire, nous avons essayé de voir quelle est la structure nécessaire au peptide Sgal et au peptide **VGVAPG** pour que l'interaction ait lieu. Enfin, nous avons vérifié la stabilité des modèles proposés par simulation de dynamique moléculaire.

B.1. Résultats expérimentaux :

B.1a. Dichroïsme circulaire électronique

Le peptide Sgal se fixe à l'élastine entière *via* les répétitions **VGVAPG** de son exon 24. L'interaction entre le peptide Sgal et un peptide de 18 résidus composé de 3 répétitions **VGVAPG** (**Figure 155**), puis entre le peptide Sgal et l'hexapeptide **VGVAPG** (**Figure 156**) ont été étudié par dichroïsme circulaire électronique dans l'eau à 25 °C. Nous avons reporté **Figures 155-156** les spectres de DCE des peptides en interaction et les spectres de la somme théorique des contributions individuelles de chacun des peptides.

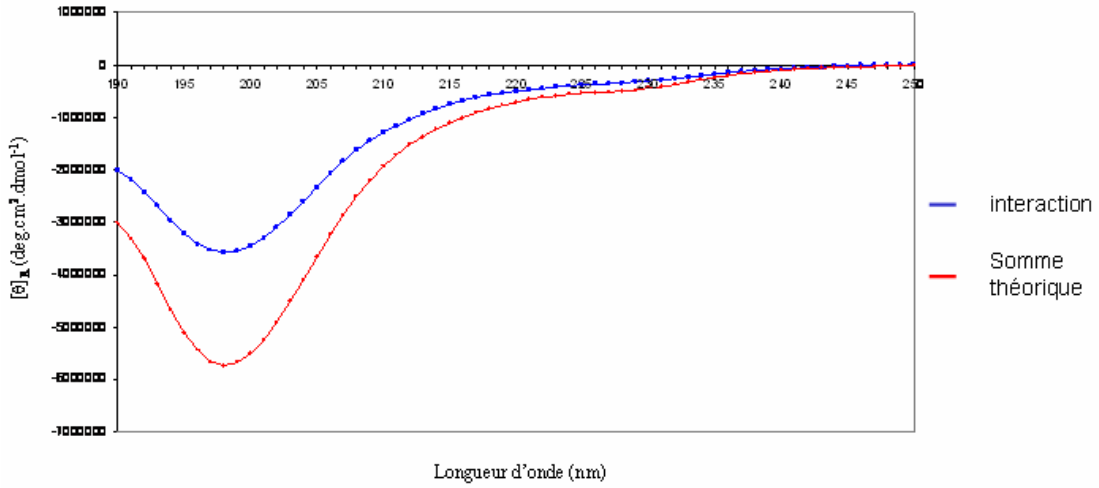


Figure 155 : Spectre DC de l'interaction entre le peptide Sgal et $(VGVAPG)_3$.

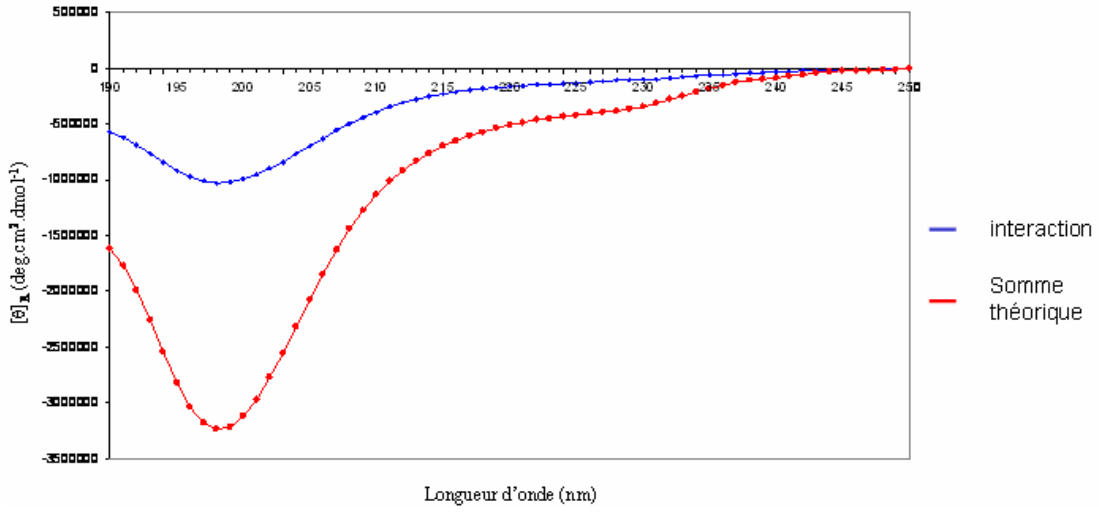


Figure 156 : Spectre DCE de l'interaction entre le peptide Sgal et VGVAPG.

La différence significative qui existe entre le spectre de la somme théorique et le spectre de l'interaction montre que le peptide Sgal peut se lier aussi bien au peptide (**VGVAPG**)$_3$ qu'à l'hexapeptide **VGVAPG**. En effet, cette différence dans l'allure des spectres de DCE indique qu'au moins l'un des peptides (peptide Sgal ou EDP) a vu sa structure modifiée par la présence du second peptide, indiquant ainsi qu'une interaction a eut lieu entre les 2 peptides (peptide Sgal et EDP). Malheureusement, le DCE ne permet pas d'en apprendre d'avantage sur la nature de cette l'interaction, comme par exemple : quels sont les résidus mis en jeu, quelles sont les conformations des peptides en présence et encore moins l'affinité qu'il peut exister entre le peptide Sgal et les EDP.

B.1.b. Fluorescence :

Pour étudier l'interaction entre deux molécules par spectroscopie de fluorescence, il est nécessaire qu'au moins l'une des deux soit fluorescente. Le peptide Sgal ne possède pas de résidus ayant cette propriété. C'est pourquoi, nous avons laissé le groupement Fmoc qui est utilisé dans la synthèse peptidique du peptide **VGVAPG** car il est constitué de noyaux aromatiques lui conférant des propriétés de fluorescence. Les mesures de fluorescence ont été effectuées dans l'eau à 25 °C.

Le graphique de Scatchard obtenu pour l'interaction entre le peptide Sgal et **VGVAPG** est reporté **Figure 157**. Grâce au coefficient directeur de la droite, nous pouvons en déduire la constante de dissociation du complexe peptide Sgal - peptide **VGVAPG**.

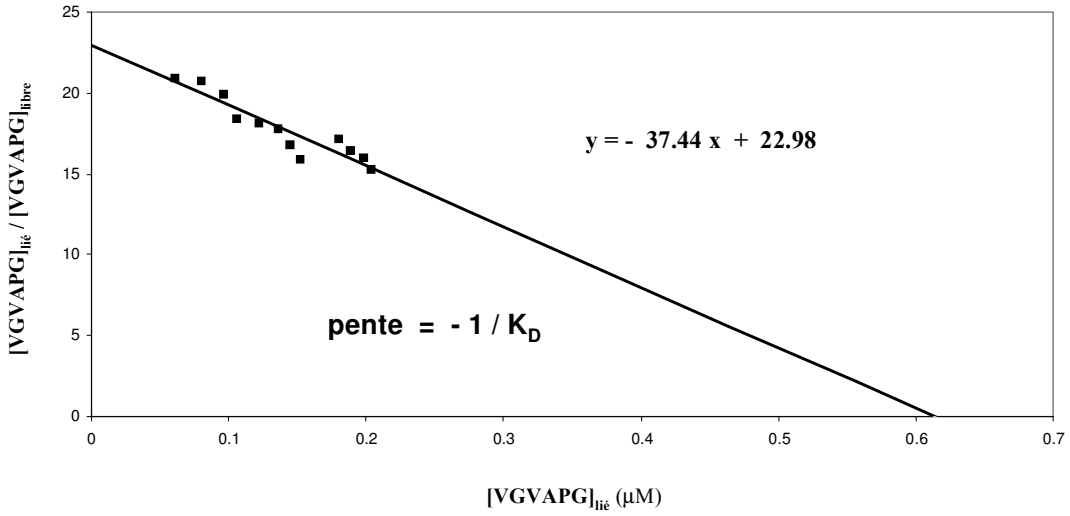


Figure 157 : Graphique de Scatchard pour le complexe VGVAPG - peptide Sgal.

Nous avons procédé de la même manière pour l'interaction entre le peptide Sgal et un autre EDP, le peptide **YGVG**, qui ne possède pas de motif **GXXP**, mais qui a toutefois une activité biologique *via* sa fixation à l'EBP. L'avantage de ce peptide est qu'il possède naturellement par l'intermédiaire du résidu **Y** un groupement fluorescent. Le graphique de Scatchard obtenu pour l'interaction entre le peptide Sgal et le peptide **YGVG** est représenté en **Figure 158**.

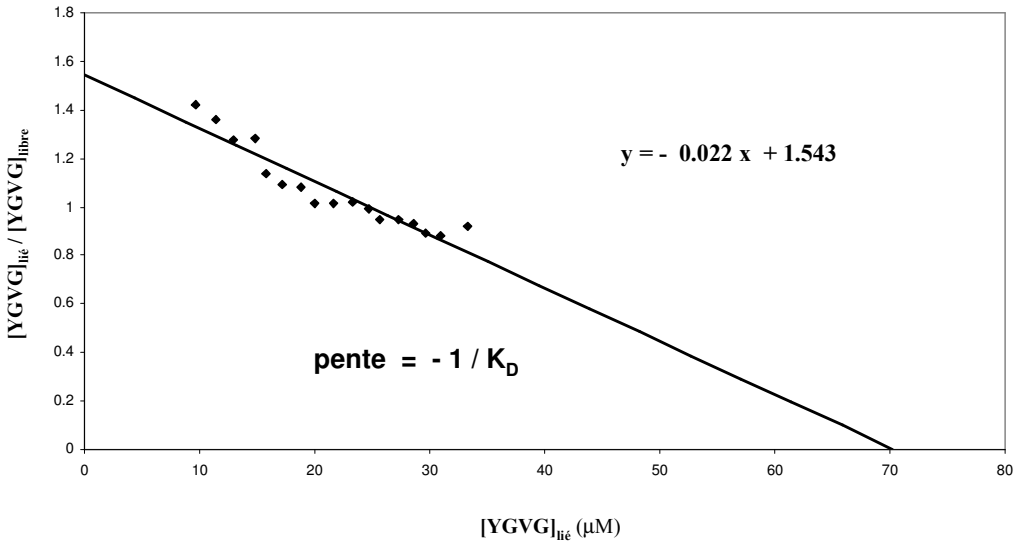


Figure 158 : Graphique de Scatchard pour le complexe YGVG - peptide Sgal.

Grâce aux graphiques de Scatchard, on obtient les constantes de dissociation suivantes :

- K_D = 27,6 nM pour le complexe **VGVAPG** - peptide Sgal, proche de celle de 8 nM calculée pour le complexe tropoélastine - EBP (Wrenn et al. 1988). La différence observée provient de ce que les 2 peptides sont beaucoup plus flexibles que les 2 protéines, cela diminue donc leurs chances d'être dans la conformation adéquate pour l'interaction. Cette interaction de part la valeur de son K_D est très spécifique.

- K_D = 45,4 µM pour le complexe **YGVG** - peptide Sgal. L'interaction entre ces 2 peptides est donc moins forte qu'entre **VGVAPG** et le peptide Sgal, cette interaction est plutôt de type non spécifique.

B.1.c. Conclusions :

Nous avons pu montrer que le peptide Sgal peut se fixer avec des peptides dérivant de l'élastine. L'interaction est aspécifique dans le cas du peptide **YGVG** et spécifique dans le cas du peptide **VGVAPG** contenant un motif **GXXP**, avec une constante de dissociation K_D de 27,6 nM.

B.2. Résultats théoriques :

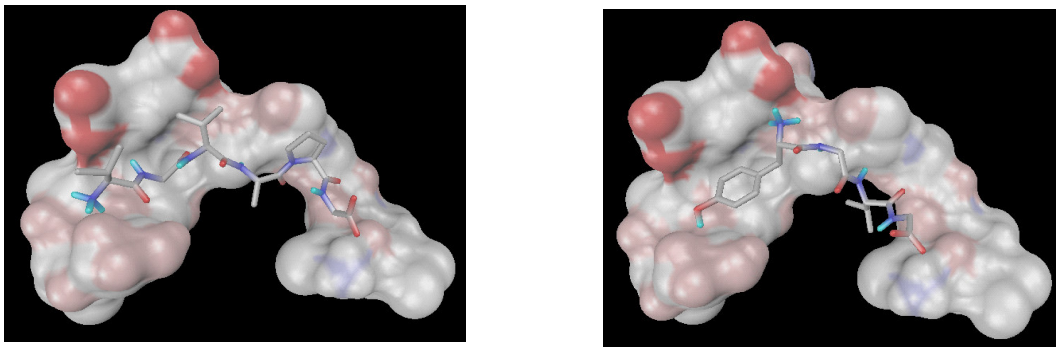
B.2.a. Amarrage moléculaire :

Les résultats expérimentaux montrant que les peptides **VGVAPG** et **YGVG** pouvaient se fixer sur le peptide Sgal, nous avons voulu étudier, par amarrage moléculaire, où et en adoptant quelle conformation ces peptides pouvaient se fixer au peptide Sgal. La structure du peptide Sgal utilisé a été celle du modèle construit par homologie à partir de la PPE.

Les calculs d'amarrage moléculaire de **VGVAPG** ou **YGVG** sur le peptide Sgal ont été effectués dans les mêmes conditions et avec le même protocole que celui utilisé pour les calculs d'amarrage moléculaire sur la PPE.

Les résultats obtenus sont malheureusement décevants car la structure et la position des peptides proposés par AutoDock ne correspondent pas à celles obtenues lors des calculs d'amarrage moléculaire sur la PPE. Plus grave encore, les meilleurs candidats proposés se situent à un endroit correspondant à une partie non exposée dans la PPE et donc très certainement à une partie non exposée de l'EBP. En effet, les peptides amarrés présentent tous un pont salin avec l'une des extrémités N- ou C-terminale du peptide Sgal qui sont normalement enfouies au sein dans la PPE et vraisemblablement également au de l'EBP (**Figure 159**).

a) b)



VGVAPG YGVG

<u>Figure 159</u> : Meilleurs candidats proposés par AutoDock pour l'interaction entre le peptide Sgal et a) VGVAPG ou b) YGVG. Le peptide Sgal est représenté avec sa surface accessible au solvant, les peptides VGVAPG et YGVG sont représentés en "bâtons".

Le programme AutoDock a été développé pour des calculs d'amarrage moléculaire sur des protéines et non pas des peptides. Il est donc possible que le traitement des interactions

non liantes de type électrostatique ne soit donc pas adapté à notre système et favorise les ponts salins entre la molécule cible et les ligands.

On peut constater que les ligands se placent dans une position où ils sont le moins exposés, ce qui nous laisse penser que le terme de désolvatation n'est probablement pas adapté, lui non plus, à des études où les molécules cibles sont des peptides.

Nous avions initialement prévu, pour compenser la rigidité qu'impose le programme AutoDock à la molécule cible, d'effectuer d'autres calculs d'amarrage moléculaire en utilisant différentes structures du peptide Sgal provenant des calculs de Monte-Carlo ou de dynamique moléculaire. Cependant, nous y avons renoncé puisque AutoDock ne semble pas adapté pour traiter des interactions de type peptide-peptide.

B.2.b. Simulations de dynamique moléculaire :

α. En partant des modèles proposés par l'amarrage moléculaire sur Sgal :

Nous avons effectué une simulation de dynamique moléculaire de l'interaction entre le peptide Sgal et chacun des peptides **VGVAPG** et **YGVG**, avec les structures correspondant à la meilleure solution proposée par AutoDdock.

Les simulations ont été réalisées avec le même protocole et dans les mêmes conditions que celles du peptide Sgal seul. Les phases de production ont été de 20 ns.

(i) Peptide Sgal / VGVAPG :

La simulation de dynamique moléculaire de l'interaction peptide Sgal et **VGVAPG** a confirmé nos doutes sur la validité des solutions d'amarrage moléculaire proposé par AutoDock. En effet, en moins de 4 ns, les deux peptides n'étaient plus du tout en interaction l'un avec l'autre.

(ii) Peptide Sgal / YGVG :

A l'inverse de la simulation précédente, les deux peptides sont restés en interaction plus longtemps. La distance minimale entre les atomes lourds de chacun des peptides reste inférieure à 4 Å et ce pendant environ 8 ns. Ceci est dû au groupement aromatique du résidus **Y** du peptide **YGVG**. En effet, durant cette période, malgré certains changements de conformation du peptide **YGVG**, le résidu **Y** est à l'origine d'une interaction de type hydrophobe avec plusieurs résidus du peptide Sgal, mais sans réelle spécificité vis à vis d'un en particulier. Il semble donc que l'on retrouve l'interaction non spécifique entre le peptide

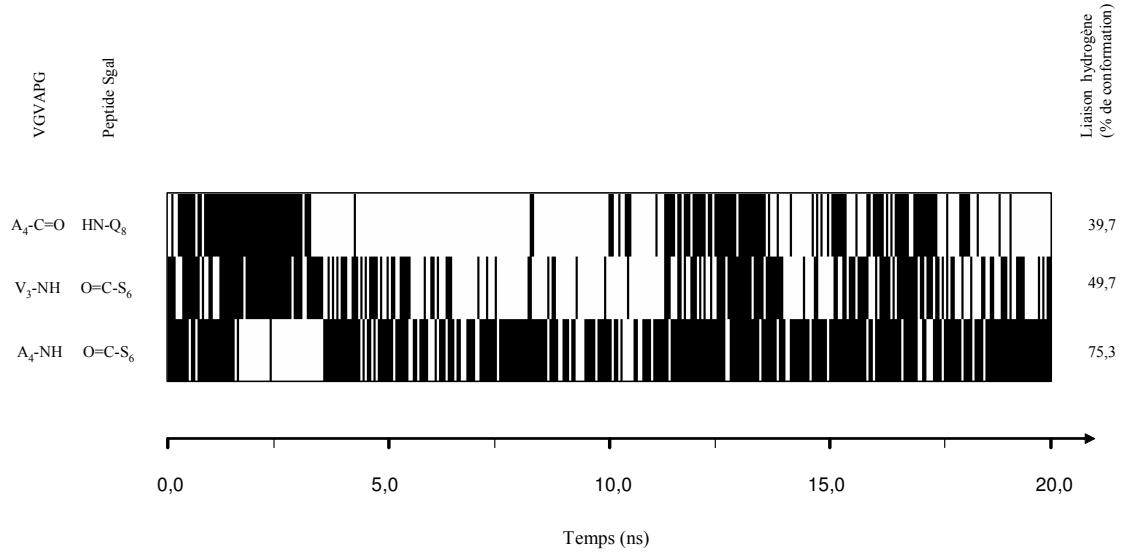
Sgal et **YGVG** qui avait été démontrée expérimentalement. Mais à cause des doutes que nous avons sur la validité du modèle de départ, il semble, tout de même, hasardeux de tirer de quelconques conclusions à partir de cette simulation de dynamique moléculaire.

β. En partant des modèles proposés par l'amarrage moléculaire sur la PPE :

Nous avons effectué une simulation de dynamique moléculaire de peptide Sgal en présence du peptide **VGVAPG**, dans une configuration ne provenant **pas** des calculs d'amarrage de **VGVAPG** sur le **peptide Sgal**, mais de **VGVAPG** sur la **PPE**.

Nous avons placé le peptide **VGVAPG**, dans la meilleure configuration proposée par les calculs d'amarrage moléculaire ayant pour cible la PPE, sur la structure construite par homologie du peptide Sgal. Le protocole de simulation de dynamique moléculaire est le même que celui précédemment utilisé. La phase de production a duré 20 ns.

Les liaisons hydrogène présentes dans le modèle de départ entre le peptide Sgal et **VGVAPG** sont régulièrement observables au cours de la trajectoire (**Figure 160)**. La présence de ces liaisons hydrogène prouve que les deux peptides restent bien en interaction durant toute la simulation de dynamique moléculaire.



<u>Figure 160</u> : Présence des liaisons hydrogène entre VGVAPG et le peptide Sgal au cours de la simulation de 20 ns dans l'eau.

Comme pour les simulations de dynamique moléculaire du peptide Sgal seul, aucune structure régulière n'est détectable hormis des coudes β (**Tableau 49 et Figure161**).

	Type de coude β (%)					
	canoniques et déformés					Total
	I	I'	II	II'	VIII	
VVGS	–	–	12,9	–	0,2	13,1
VGSP	–	–	–	1,3	8,7	10,0
GSPS	–	–	–	–	–	–
SPSA	–	–	–	–	–	–
PSAQ	0,2	–	–	–	0,2	0,4
SAQD	29,9	–	–	–	3,2	53,1
AQDE	14,7	–	–	–	9,2	23,9
QDEA	6,0	–	–	–	3,4	9,4
DEAS	9,7	–	–	–	2,8	12,
EASP	–	–	–	–	2,6	2,6
ASPL	–	–	–	–	–	–
SPLS	49,3	–	–	–	8,0	57,3

Tableau 49 : Pourcentages de coudes β du peptide Sgal en interaction avec VGVAPG au cours de la simulation de 20 ns dans l'eau.

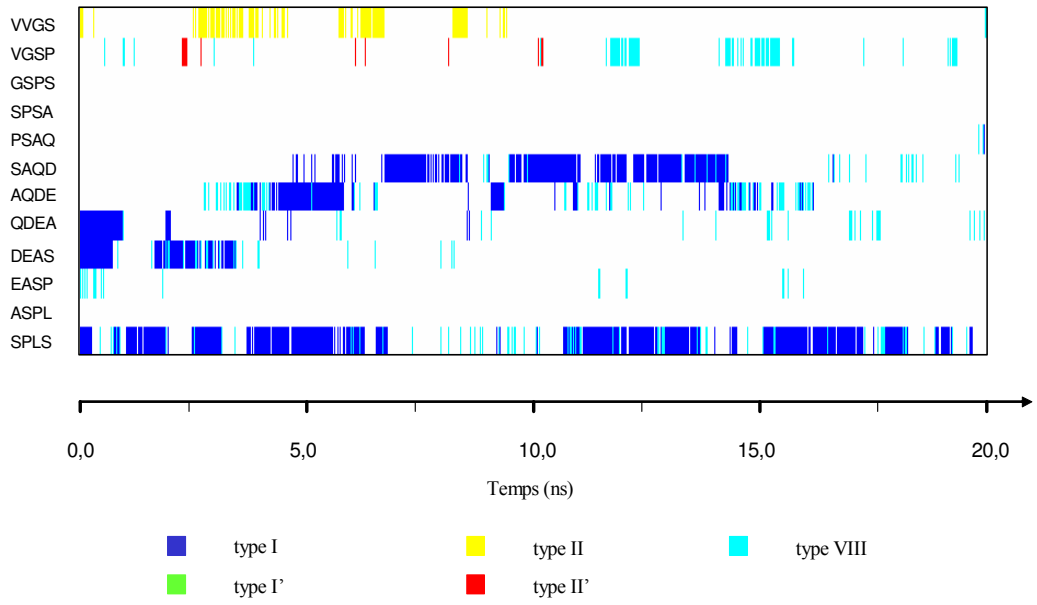
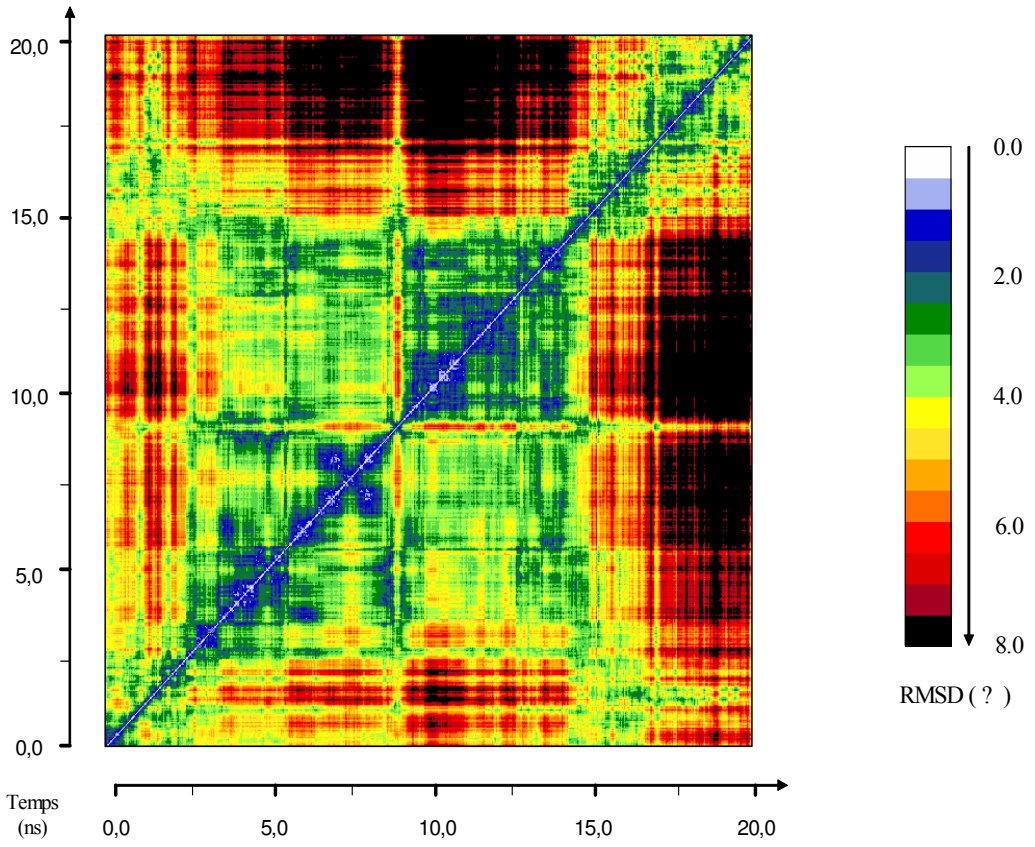


Figure 161 : **Présence des différents types de coudes β du peptide Sgal en interaction avec le peptide VGVAPG au cours de la simulation de 20 ns dans l'eau.**

Il est assez étonnant de constater que la présence du peptide **VGVAPG** ne stabilise pas les coudes β initialement présents sur la conformation de départ du peptide Sgal. Mais, il est à noter que notamment le motif **QDEA**, même si il ne satisfait plus au critère exact d'un coude β, est tout de même dans une conformation proche. En effet, seule la distance Cα-Cα entre $\mathbf{Q_8}$ et $\mathbf{A_{11}}$ est supérieure à 7, 0 Å et ne permet plus de classer le motif **QDEA** comme étant un coude β.

La carte des RMSD croisés de la structure du peptide Sgal confirme également que ce peptide présente toujours une certaine flexibilité même en présence de son ligand, le peptide **VGVAPG**. On peut distinguer 3 principaux amas de conformations proches. Les différences structurales entre les conformations représentatives de chacun de ces amas proviennent essentiellement des mouvements des résidus en bout de chaîne peptidique. Mais l'affinité et la stabilité des structures que prennent localement ces deux peptides, sont tout de même suffisantes pour conserver une interaction entre les deux.



<u>Figure 162</u> : Carte des RMSD croisés du peptide Sgal durant sa DM en interaction avec VGVAPG.

A l'inverse, le peptide **VGVAPG** reste pratiquement dans la même structure au cours de la trajectoire de dynamique moléculaire. Une structure proche d'un coude β de type VIII sur le motif **GVAP** le caractérise, ainsi qu'un coude β de type I ou II' sur la séquence **VGVA**, tandis qu'aucun coude β n'est observable sur **VAPG** (**Tableau 50 et Figure 163**).

Séquences	Types de coudes β canoniques ou déformés (%)			total
	I	II'	VIII	
VGVA	**7,8**	**18,8**	**-**	**26,6**
GVAP	**-**	**-**	**28,6**	**28,6**
VAPG	**-**	**-**	**-**	**-**

Tableau 50 : Pourcentages des coudes β observés dans la structure du peptide VGVAPG en interaction avec le peptide Sgal au cours de la DM de 20 ns dans l'eau.

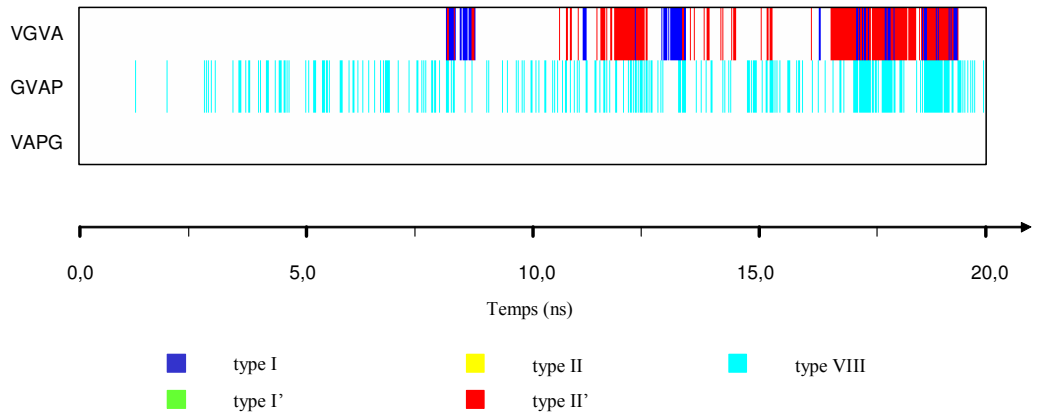
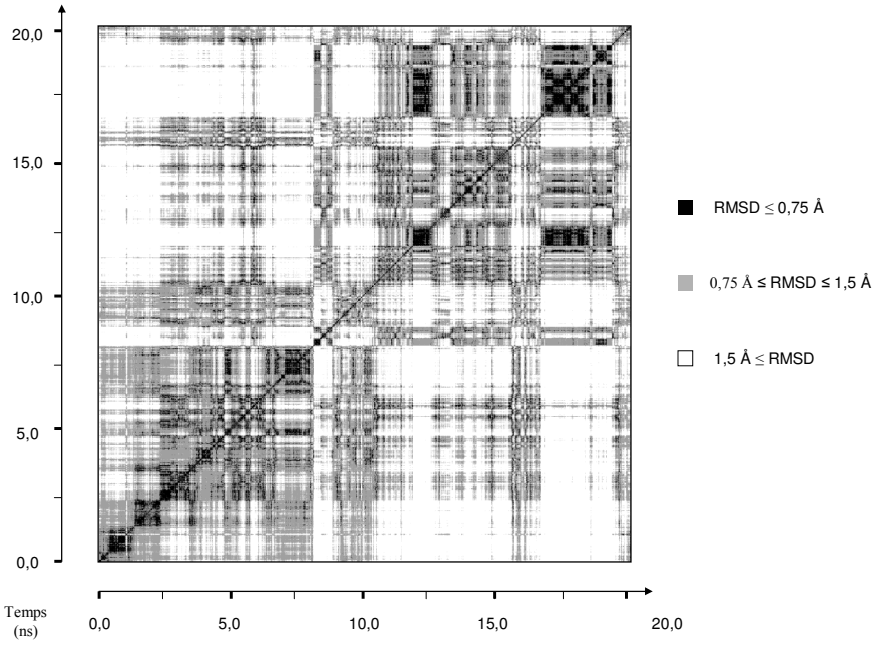


Figure 163 : Présence des différents types de coudes β du peptide VGVAPG en interaction avec le peptide Sgal au cours de la DM de 20 ns dans l'eau.

La carte de RMSD croisés montre que le peptide **VGVAPG** présente 2 conformations majeures très proches l'une de l'autre.

La première correspondant à l'amas observable au début de la trajectoire, durant les 10 premières ns, est composée d'un coude β de type VIII sur **GVAP**.

La deuxième, présente dans la deuxième partie de la simulation, à partir, de 10 ns jusqu'à la fin de la trajectoire, possède en plus du coude β de type VIII sur **GVAP**, un coude β oscillant entre le type I et, de manière prépondérante, le type II' sur **VGVA**.



<u>Figure 164</u> : Carte des RMSD croisés du peptide VGVAPG durant sa DM en interaction avec le peptide Sgal.

Il existe donc bien une interaction forte entre le peptide Sgal et le peptide **VGVAPG**, puisque après 20 ns, ces deux molécules sont toujours liées entre elles par 3 liaisons hydrogène.

Des deux peptides, ce n'est pas celui qui est considérée comme molécule réceptrice, c'est-à-dire le peptide Sgal, qui est la structure la plus stable, mais le peptide **VGVAPG** qui conserve tout au long de la trajectoire un coude β de type VIII sur **GVAP**.

C. CONCLUSIONS

Nous avons pu montrer par amarrage moléculaire que les EDPs possédaient bien une grande affinité pour les 15 premiers résidus de la PPE qui correspondent au peptide Sgal. Cette séquence est bien une séquence de liaison avec les EDPs et donc l'élastine.

Les EDPs se lient tous dans une conformation similaire caractérisée par un coude β déformé de type VIII sur le motif **GXXP**. Ce coude β entoure un résidu particulier de la PPE, le $\mathbf{Q_8}$, qui est également présent dans la séquence du peptide Sgal. Le pentapeptide **GGVPG**, qui est biologiquement inactif, peut lui aussi adopter cette conformation. On peut donc penser que l'absence d'interaction avec l'EBP est due à l'entropie dont est à l'origine le résidu $\mathbf{G_2}$, ce qui diminue ainsi les chances d'adopter la conformation adéquate pour se fixer à l'EBP.

Les résultats expérimentaux portant sur l'étude de l'interaction directe entre le peptide Sgal et les EDPs montrent, tout d'abord, que cette interaction existe bel et bien, mais également qu'elle est très spécifique dans le cas du peptide **VGVAPG** contenant un motif **GXXP**, avec une constante de dissociation de 27,6 nM pour le peptide **VGVAPG**.
En revanche, l'interaction entre le peptide Sgal et le peptide **YGVG** est aspécifique, la constante de dissociation étant de 46,4 μM.

Malheureusement, les résultats d'amarrage moléculaire entre le peptide Sgal et les EDP, ainsi que les simulations de dynamique moléculaire réalisées à partir des meilleures solutions obtenues, ont donné d'improbables résultats. Ceux-ci sont très certainement faussés par le fait que le programme AutoDock utilisé n'est pas bien adapté aux calculs d'amarrage moléculaire entre peptides.

En revanche, la dynamique moléculaire ayant comme modèle de départ une conformation provenant des calculs d'amarrage moléculaire du peptide **VGVAPG** sur la PPE, montre que durant 20 ns, 3 liaisons hydrogène stabilisent le complexe peptide Sgal - peptide **VGVAPG**. En outre, ce dernier présente tout au long de la trajectoire une conformation caractérisée par un coude β de type VIII sur le motif **GVAP**.

Il est donc envisagé de caractériser expérimentalement le complexe peptide Sgal - peptide **VGVAPG** par spectroscopie de RMN comme modèle restreint de l'interaction existant entre l'EBP et les EDPs ou l'élastine.

Partie VI

CONCLUSIONS GENERALES ET PERSPECTIVES

L'EBP (pour Elastin Binding Protein) est la sous-unité du récepteur de l'élastine qui permet la fixation des ligands sur celui-ci. Malheureusement aucune donnée expérimentale sur sa structure n'est pour l'instant connue. La construction d'un modèle par homologie n'étant pas réalisable, notre étude s'est concentrée sur une séquence de 15 résidus de l'EBP qui a été montré comme étant responsable de l'interaction avec l'élastine. Ces 15 résidus sous forme peptidique, appelé peptide Sgal, conserve ses propriétés de fixation vis-à-vis de l'élastine.

Nous avons donc dans un premier temps étudié expérimentalement la structure que pouvait prendre le peptide Sgal.

Les spectroscopies RMN et de DCE ont montré que, dans l'eau, le peptide Sgal présentait les caractéristiques d'une hélice polyproline II en équilibre avec une conformation non ordonnée.

La modélisation moléculaire prend donc ici toute son importance pour suppléer les manques de données structurales. Ainsi, l'étude par modélisation moléculaire du peptide Sgal a permis de mettre en évidence la présence récurrente d'un coude β de type I sur la séquence **QDEA** et ce, quelle que soit la méthode utilisée (simulations de Monte-Carlo ou de dynamique moléculaire). Nous avons en outre constaté que les résidus du peptide Sgal possédaient majoritairement, au cours des 70 ns de simulation de dynamique moléculaire réalisées, des valeurs d'angles dièdres correspondant à celles de la polyproline II, ce qui est en bon accord avec les données expérimentales obtenues.

Nous nous sommes ensuite intéressés aux ligands de l'EBP les plus connus : les peptides dérivant de l'élastine (EDPs) et contenant un motif **GXXP**. A partir de 128 simulations de dynamique de tous les tétra, penta ou hexapeptides contenant un motif **GXXP** de l'élastine humaine, nous avons pu étayer l'hypothèse que ce motif se repliait de manière préférentielle en coude β de type VIII.

Nous avons pu également constater que la présence d'une glycine en position centrale du motif **GXXP**, diminuait les chances de repliement en coude β de type VIII. La présence de la glycine dans cette position fait également apparaître un coude β de type II', jamais observé chez les autres EDPs.
Un résidu quelconque placé avant le motif **GXXP** augmente le taux de présence de coude β de type VIII sur ce motif.
Cependant, un résidu situé après ce motif ne semble avoir aucune conséquence sur le repliement en coude β de type VIII sur **GXXP**, laissant à penser que la glycine dont la présence est indispensable à l'activité biologique doit avoir un rôle dans la fixation de l'EDP à l'EBP.

La présence de coude β de type VIII sur le motif **GXXP** a également pu être corrélée avec l'activité biologique connue de certains de ces peptides. Quatre peptides actifs (**VGVAPG**, **PGAIPG**, **GVAPG** et **GLVPG**) et le seul peptide inactif connu, **GGVPG**, ont été plus précisément étudiés à l'aide de différentes méthodes de modélisation moléculaire (Monte-Carlo, carte adiabatique et dynamique moléculaire). Nous avons ainsi pu renforcer l'hypothèse selon laquelle la présence d'un coude β de type VIII sur la séquence **GXXP** était nécessaire à ces peptides pour se fixer à l'EBP, car tous les peptides actifs ne se replient que dans cette structure.
A l'inverse, le peptide inactif **GGVPG** non seulement présente peu de coudes β de type VIII sur **GGVP**, mais également est le seul à se replier en coude β de type II' sur **GGVP**. C'est d'ailleurs ce type de coude que les calculs de Monte-Carlo et de carte adiabatique proposent comme étant la structure la plus favorable pour le peptide **GGVPG**.

L'étude de deux peptides dérivant de la fibrilline-1, l'un actif (**EGFE<u>P</u>G**) et l'autre non (mutation P => S <=> **EGFE<u>S</u>G**) montre que le peptide inactif même s'il peut adopter un coude β de type VIII sur **GFES**, peut également présenter un coude β de type I.
Le peptide actif dérivant de la fibrilline-1, lui, tout comme les peptides dérivant de l'élastine, ne se replie qu'en coude β de type VIII.
Outre le fait qu'ils ont de plus faibles taux de coudes β de type VIII, on se rend compte que le point commun que possèdent les deux peptides inactifs est qu'ils sont plus flexibles que les peptides actifs, ayant donc une plus grande entropie diminuant ainsi les chances de se lier à l'EBP.

Après avoir étudié séparément la structure que peut prendre la séquence réceptrice de l'EBP, le peptide Sgal, et celle que peuvent adopter certains des ligands de l'EBP, les EDPs, nous nous sommes ensuite intéressés à leur interaction.

Par l'intermédiaire de la PPE, nous avons voulu vérifier que les EDP possédaient bien une affinité particulière pour la séquence de 15 résidus de la PPE ayant de grandes homologies de séquence avec le peptide Sgal.
Les calculs d'amarrage moléculaire, sur la PPE, des EDPs d'activité connue ont montré qu'effectivement l'enchaînement de 15 résidus correspondant au peptide Sgal est un site de fixation pour ces EDPs. Les cinq peptides interagissent avec la PPE par l'intermédiaire de 3 liaisons hydrogènes avec la PPE et sont dans une structure proche du coude β de type VIII. Une dynamique moléculaire de la PPE en interaction a permis de montrer que cette interaction était stable sur 4 ns.

Nous avons ensuite démontré expérimentalement, par spectroscopie de DCE, pour la première fois que le peptide Sgal est capable de se lier directement à des EDPs, sous la forme d'un peptide de 18 résidus constitué de 3 répétitions de la séquence **VGVAPG**, mais également sous la forme d'un simple hexapeptide : le peptide **VGVAPG**.
Pour quantifier l'affinité qui existe entre le peptide Sgal et les EDP, par spectroscopie de fluorescence, nous pu montrer que cette interaction était spécifique avec une constante de dissociation de 27,6 nM pour un peptide contenant un motif **GXXP**, le peptide **VGVAPG**. Par contre, l'interaction est aspécifique pour un EDP ne contenant pas de motif **GXXP**, le peptide **YGVG**, avec une constante de dissociation de 45,4 μM.

Enfin, un complexe peptide Sgal - peptide **VGVAPG** a été construit, basé sur les résultats d'amarrage moléculaire sur la PPE. Après 20 ns de dynamique moléculaire, les deux peptides forment toujours un complexe, stabilisé par trois liaisons hydrogènes qui sont les mêmes que celles qui caractérisaient la fixation de l'EDP sur la PPE.

Il donc très probable que nous avons réussi à obtenir les bonnes configurations favorisant l'interaction entre le peptide Sgal et le peptide **VGVAPG** et par là même l'interaction entre l'EBP et la tropoélastine ou ses peptides dérivés.

Il est désormais envisageable d'étudier la possibilité de synthétiser des analogues chimiques possédant les mêmes caractéristiques structurales qu'un coude β de type VIII. Le but étant que l'analogue se fixe sur la partie correspondant au peptide Sgal de l'EBP pour interdire aux EDPs de se lier au récepteur de l'élastine, et ainsi empêcher son implication dans l'invasion tumorale, l'athérosclérose et vieillissement.

De même, les élastases de la famille de la PPE sont impliquées dans de nombreux processus pathologiques, dont les processus inflammatoires, cancéreux et le vieillissement. Ils sont, depuis des années, une cible de choix pour de nombreuses études qui cherchent à inhiber leurs activités protéolytiques vis-à-vis de la matrice extracellulaire.

Peut-être, au vu des résultats que nous avons obtenus, serait-il possible d'essayer, non plus d'inhiber le site de protéolyse de cette enzyme, mais plutôt le site de liaison au moyen d'un ligand mimant une structure en coude β de type VIII.

Partie VII

RÉFÉRENCES BIBLIOGRAPHIQUES

Abagyan, R. and Totrov, M. (1994). *Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins*. J Mol Biol **235**: 983-1002.

Abe, H., Braun, W., Noguti, T. and Go, N. (1984). *Rapid calculation of first and second derivatives of conformationnal energy with respect to dihedral angles for proteins. General recurrent equations*. Comput Chem **8**: 239-247.

Alix, A. J. P. (1997). *Molecular modeling of globular proteins : Strategy 1D -> 3D : secondary structures and epitopes.* NATO ASI Series E Applied Sciences. **342:** 121-150.

Alix, A. J. P. (1999). *Predictive estimation of protein linear epitopes by using the program PEOPLE*. Vaccine **18**(3-4): 311-314.

Alix, A. J. P. (2001). *A turning point in the knowledge of the structure-function-activity relations of elastin*. J Soc Biol **195**(2): 181-193.

Alix, A. J. P., Berjot, M. et Marx, J. (1985). Determination of the secondary structure of proteins. Dans A. J. P. Alix , L. Bernard & M. Manfait (eds), *Spectroscopy of biological molecules,* Wiley, Chichester, pp. 149-154.

Alix, A. J. P., Bochichio, B., Dauchez, M., Debelle, L., Floquet, N., Fuchs, P., Héry-Huynh, S., Kaplan, H., Moroy, G. and Tamburro, A. M. (2003). *ELASTIN in silico*. Elastin 2002. Edited by Tamburro, A. M. and Pepe, A.

Alix A. J. P., Pedanou, G. and Berjot, M. (1988). *Fast determination of the quantitative secondary structure of proteins by using some parameters of the Raman Amide I band*. J Mol Struct (G.B.), **174:** 159-164.

Argos, P., Hanei, M. and Garavito, R. M. (1978). *The Chou-Fasman secondary structure prediction method with an extended data base*. FEBS Lett **93**(1): 19-24.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997). *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs*. Nucleic Acids Res **25**(17): 3389-402.

Bairoch, A. and Apweiler, R. (1997). *The SWISS-PROT protein sequence database: its relevance to human molecular medical research.* J Mol Med **75**: 312-216.

Bashford, D. and Case, D. A. (2000). *Generalized born models of macromolecular solvation effects*. Annu Rev Phys Chem **51**: 129-152.

Bashir, M., Indik, Z., Yeh, H., Abrams, W., Ornstein-Goldstein, N., Rosenbloom, J. C., Fazio, M. J., Uitto, J., Mecham, R. and Rosembloom, J. (1990). *Elastin gene structure and mRNA alternative splicing.* Elastin: chemical and biological aspects. Edited by Tamburro, A. M. and Davidson, J. M.

Berendsen, H. J. C., Postma, J. P. M., van Gusteren, W. F., DiNola, A. and Haak, J. R. (1984). *Molecular dynamics with coupling to an external bath*. J Chem Phys **81**: 3684-3690.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. and Bourne, P. E. (2000). *The Protein Data Bank*. Nucleic Acids Res **28**(1): 235-42.

Berjot, M., Marx, J. and Alix, A. J. P. (1987). *Determination of the secondary structure of proteins from the Raman Amide I band : the Reference Intensity Profiles Method.* J. Raman Spectrosc. **18:** 289-300.

Bisaccia, F., Castiglione-Morelli, M. A., Spisani, S., Ostuni, A., Serafini-Fracassini, A., Bavoso, A. and Tamburro, A. M. (1998). *The amino acid sequence coded by the rarely expressed exon 26A of human elastin contains a stable beta-turn with chemotactic activity for monocytes*. Biochemistry **37**(31): 11128-35.

Bisaccia, F., Castiglione-Morelli, M. A., Spisani, S., Serafini-Fracassini, A. and Tamburro, A. M. (2000). *Solution structure of the amino acid sequence coded by the rarely expressed exon 26A of human elastin: the N-terminal region*. J Pept Res **56**(4): 201-209.

Bisaccia, F., Morelli, M. A., De Biasi, M., Traniello, S., Spisani, S. and Tamburro, A. M. (1994). *Migration of monocytes in the presence of elastolytic fragments of elastin and in synthetic derivates. Structure-activity relationships*. Int J Pept Protein Res **44**(4): 332-41.

Bisaccia, F., Ostuni, A. and Tamburro, A. M. (1999). *Biological activities of elastin peptides*. Current Topics in Peptide and Protein Research **3**.

Blow, D. M. and Steitz, T. A. (1970). *X-ray diffraction studies of enzymes*. Annu Rev Biochem **39**: 63-100.

Bochicchio, B., Ait-Ali, A., Tamburro, A. M. and Alix, A. J. P. (2004a). *Spectroscopic evidence revealing polyproline II structure in hydrophobic, putatively elastomeric sequences encoded by specific exons of human tropoelastin*. Biopolymers **73**(4): 484-493.

Bochicchio, B., Floquet, N., Pepe, A., Alix, A. J. P. and Tamburro, A. M. (2004b). *Dissection of human tropoelastin: solution structure, dynamics and self-assembly of the exon 5 peptide*. Chemistry **10**(13): 3166-3176.

Bochicchio, B., Pepe, A. and Tamburro, A. M. (2001). *On (GGLGY) synthetic repeating sequences of lamprin and analogous sequences*. Matrix Biol **20**(4): 243-250.

Brassart, B., Fuchs, P., Huet, E., Alix, A. J. P., Wallach, J., Tamburro, A. M., Delacoux, F., Haye, B., Emonard, H., Hornebeck, W. and Debelle, L. (2001). *Conformational dependence of collagenase (matrix metalloproteinase-1) up-regulation by elastin peptides in cultured fibroblasts*. J Biol Chem **276**(7): 5222-5227.

Brassart, B., Randoux, A., Hornebeck, W. and Emonard, H. (1998). *Regulation of matrix metalloproteinase-2 (gelatinase A, MMP-2), membrane-type matrix metalloproteinase-1 (MT1-MMP) and tissue inhibitor of metalloproteinases-2 (TIMP-2) expression by elastin-derived peptides in human HT-1080 fibrosarcoma cell line*. Clin Exp Metastasis **16**(6): 489-500.

Bressan, G. M. and Prockop, D. J. (1977). *Synthesis of elastin in aortas from chick embryos. Conversion of newly secreted elastin to cross-linked elastin without apparent proteolysis of the molecule*. Biochemistry **16**(7): 1406-12.

Brooks, B. R., 3rd, Karplus, M. and Pettitt, B. M. (1988). *Protein. A Theoritical Perspective of Dynamics, Structure, and Thermodynamics Advances in Chemical Physics*. Wiley-Interscience. New-York. **71**.

Brooks, B. R., Bruccoleri, B. D., Olafson, B. D., States, D. J., Swaminathan, S. and Karplus, M. (1983). *CHARMM: A program for macromolecular energy, minimization and dynamics calculations*. J Comp Chem **4**(2): 187-217.

Brooks, B. R., Pettitt, B. M. and Karplus, M. (1985). *Structural and energetic effects of truncating long ranged interactions in ionic and polar fluids*. J Chem Phys **83**: 5897-5908.

Brown, P. L., Mecham, L., Tisdale, C. and Mecham, R. P. (1992). *The cysteine residues in the carboxy terminal domain of tropoelastin form an intrachain disulfide bond that stabilizes a loop structure and positively charged pocket*. Biochem Biophys Res Commun **186**(1): 549-55.

Brown-Augsburger, P., Broekelmann, T., Mecham, L., Mercer, R., Gibson, M. A., Cleary, E. G., Abrams, W. R., Rosenbloom, J. and Mecham, R. P. (1994). *Microfibril-associated glycoprotein binds to the carboxyl-terminal domain of tropoelastin and is a substrate for transglutaminase*. J Biol Chem **269**(45): 28443-9.

Brünger, A. T. and Karplus, M. (1988). *Polar hydrogen positions in proteins: empirical energy placement and neutron diffraction comparision*. Proteins **4**(2): 148-56.

Busetta, B. (1987). *The use of folding patterns in the prediction of protein topologies*. Biochim Biophys Acta **916**(1): 54-65.

Carpenter, G. A. and Grossberg, S. (1987). *ART2: self-organization of stable category recognition codes for analog input patterns*. Appl Opt **26**: 4919-4930.

Castiglione Morelli, M. A., Bisaccia, F., Spisani, S., De Biasi, M., Traniello, S. and Tamburro, A. M. (1997). *Structure-activity relationships for some elastin-derived peptide chemoattractants*. J Pept Res **49**(6): 492-9.

Castiglione-Morelli, A., Scopa, A., Tamburro, A. M. and Guantieri, V. (1990). *Spectroscopic studies on elastin-like synthetic polypeptides*. Int J Biol Macromol **12**(6): 363-8.

Chang, D. K., Venkatachalam, C. M., Prasad, K. U. and Urry, D. W. (1989). *Nuclear Overhauser effect and computational characterization of the beta-spiral of the polypentapeptide of elastin*. J Biomol Struct Dyn **6**(5): 851-8.

Chen, I. J., Yin, D. and MacKerell, A. D. (2002). *Combined Ab initio/Empirical Approach for Optimization of Lennard-Jones Parameters for Polar-Neutral Compounds*. J Comp Chem **23**: 199-213.

Chou, K. C. (2000). *Prediction of tight turns and their types in proteins*. Anal Biochem **286**(1): 1-16.

Clarke, A. W. and Weiss, A. S. (2004). *Microfibril-associated glycoprotein-1 binding to tropoelastin: multiple binding sites and the role of divalent cations*. Eur J Biochem **271**(14): 3085-3090.

Colonna-Cesari, F. and Sander, C. (1990). *Excluded volume approximation to protein-solvent interaction*. Biophys J **57**(5): 1103-1107.

Cregge, R. J., Durham, S. L., Farr, R. A., Gallion, S. L., Hare, C. M., Hoffman, R. V., Janusz, M. J., Kim, H. O., Koehl, J. R., Mehdi, S., Metz, W. A., Peet, N. P., Pelton, J. T., Schreuder, H. A., Sunder, S. and Tardif, C. (1998). *Inhibition of human neutrophil elastase. 4. Design, synthesis, X-ray crystallographic analysis, and structure-activity relationships for a series of P2-modified, orally active peptidyl pentafluoroethyl ketones*. J Med Chem **41**(14): 2461-80.

D'Agrosa, R. M. and Callahan, J. W. (1988). *In vitro activation of neuraminidase in the beta-galactosidase-neuraminidase-protective protein complex by cathepsin C*. Biochem Biophys Res Commun **157**(2): 770-5.

D'Agrosa, R. M., Hubbes, M., Zhang, S., Shankaran, R. and Callahan, J. W. (1992). *Characteristics of the beta-galactosidase-carboxypeptidase complex in GM1-gangliosidosis and beta-galactosialidosis fibroblasts*. Biochem J **285 (Pt 3)**: 833-8.

Darden, T. A., York, D. M. and Pedersen, L. G. (1993). *Particule mesh Ewald: an N log(N) method for Ewald sums in large systems*. J Chem Phys: 10089-10092.

Davis, G. E., Bayless, K. J., Davis, M. J. and Meininger, G. A. (2000). *Regulation of tissue injury responses by the exposure of matricryptic sites within extracellular matrix molecules*. Am J Pathol **156**(5): 1489-1498.

Debelle, L. and Alix, A. J. P. (1995). *Optical spectroscopic determination of bovine tropoelastin molecular model.* J Mol Struct **348**: 321-324.

Debelle, L. and Alix, A. J. P. (1999). *The structures of elastins and their function*. Biochimie **81**(10): 981-994.

Debelle, L., Alix, A. J. P., Jacob, M. P., Huvenne, J. P., Berjot, M., Sombret, B. and Legrand, P. (1995). *Bovine elastin and kappa-elastin secondary structure determination by optical spectroscopies*. J Biol Chem **270**(44): 26099-103.

Debelle, L., Alix, A. J. P., Wei, S. M., Jacob, M. P., Huvenne, J. P., Berjot, M. and Legrand, P. (1998). *The secondary structure and architecture of human elastin*. Eur J Biochem **258**(2): 533-9.

Debelle, L. and Tamburro, A. M. (1999). *Elastin: molecular description and function*. Int J Biochem Cell Biol **31**(2): 261-272.

Debelle, L., Wei, S. M., Jacob, M. P., Hornebeck, W. and Alix, A. J. P. (1992). *Predictions of the secondary structure and antigenicity of human and bovine tropoelastins*. Eur Biophys J **21**(5): 321-9.

Derreumaux, P. (1997). *A diffusion process-controlled Monte Carlo method for finding the global energy minimum of the polypeptide chain. I. Formulation and test on a hexadecapeptide*. J Chem Phys **106**(12): 5260-5270.

Derreumaux, P. (2001). *Evidence that the 127-164 region of prion proteins has two equi-energetic conformations with beta or alpha features*. Biophys J **81**(3): 1657-1665.

Dewar, M. J. S., Zoebisch, E. G., Healy, E. F. and Stewart, J. J. P. (1985). *AM1: A New General Purpose Quantum Mechanical Model*. J Am Chem Soc **107**: 3902-3909.

Dietz, H. C., Cutting, G. R., Pyeritz, R. E., Maslen, C. L., Sakai, L. Y., Corson, G. M., Puffenberger, E. G., Hamosh, A., Nanthakumar, E. J., Curristin, S. M. and et al. (1991). *Marfan syndrome caused by a recurrent de novo missense mutation in the fibrillin gene*. Nature **352**(6333): 337-9.

Dominy, B. and Brooks, B. R., 3rd (1999). *Development of a Generalized Born Model Parameterization for Proteins and Nucleic Acids*. J Phys Chem **103**: 3765-3773.

Downing, A. K., Knott, V., Werner, J. M., Cardy, C. M., Campbell, I. D. and Handford, P. A. (1996). *Solution structure of a pair of calcium-binding epidermal growth factor-like domains: implications for the Marfan syndrome and other genetic disorders*. Cell **85**(4): 597-605.

Duca, L., Debelle, L., Debret, R., Antonicelli, F., Hornebeck, W. and Haye, B. (2002). *The elastin peptides-mediated induction of pro-collagenase-1 production by human fibroblasts involves activation of MEK/ERK pathway via PKA- and PI(3)K-dependent signaling*. FEBS Lett **524**(1-3): 193-198.

Duca, L., Floquet, N., Alix, A. J. P., Haye, B. and Debelle, L. (2004). *Elastin as a matrikine*. Crit Rev Oncol Hematol **49**(3): 235-244.

Efremov, R. G. and Alix, A. J. P. (1993). *Environmental characteristics of residues in proteins: three-dimensional molecular hydrophobicity potential approach.* J Biomol Struct Dyn **11**(3): 483-507.

Efremov, R. G., Nolde, D. E., Vergoten, G. and Arseniev, A. S. (1999). *A Solvent Model for Simulations of Peptides in bilayers. I. Membrane-Promoting α-helix Formation.* Biophys J **76**: 2448-2459.

Eisenberg, D. and McLachlan, A. D. (1986). *Solvation energy in protein folding and binding*. Nature **319**(6050): 199-203.

Emini, E. A., Schleif, W. A., Colonno, R. J. and Wimmer, E. (1985). *Antigenic conservation and divergence between the viral-specific proteins of poliovirus type 1 and various picornaviruses*. Virology **140**(1): 13-20.

Essmann, U., Perera, L., Bertkowitz, M. L., Darden, T. A., Lee, H. and Pedersen, L. G. (1995). *A smooth particle mesh ewald method*. J Chem Phys: 8577-8593.

Ewald, P. (1921). *Die Berechnung optischer und elektrostatischer Gitterpotentiale*. Ann Phys **64**: 253-287.

Faury, G., Garnier, S., Weiss, A. S., Wallach, J., Fulop, T., Jr., Jacob, M. P., Mecham, R. P., Robert, L. and Verdetti, J. (1998). *Action of tropoelastin and synthetic elastin sequences on vascular tone and on free Ca2+ level in human vascular endothelial cells*. Circ Res **82**(3): 328-36.

Fazio, M. J., Mattei, M. G., Passage, E., Chu, M. L., Black, D., Solomon, E., Davidson, J. M. and Uitto, J. (1991). *Human elastin gene: new evidence for localization to the long arm of chromosome 7.* Am J Hum Genet **48**(4): 696-703.

Feig, M. and Brooks, B. R., 3rd (2004). *Recent advances in the development and application of implicit solvent models in biomolecule simulations*. Curr Opin Struct Biol **14**(2): 217-224.

Feig, M., Wompil, I. and Brooks, C. L., 3rd (2004). *Implicit solvation based on generalized Born theory in different dielectric environments*. J Chem Phys **120**(2): 903-911.

Ferrara, P. and Caflisch, A. (2000). *Folding simulations of a three-stranded antiparallel beta -sheet peptide*. Proc Natl Acad Sci U S A **97**(20): 10780-5.

Ferrara, P., Apostolakis, J. and Caflisch, A. (2002). *Evaluation of a fast implicit solvent model for molecular dynamics simulations*. Proteins **46**(1): 24-33.

Fischer, S., Dunbrack, R. L. J. and Karplus, M. (1994). *Cis-Trans Imide Isomerization of the Proline Dipeptide*. J Am Chem Soc **116**: 11931-11937.

Fleischer, K. J., Nousari, H. C., Anhalt, G. J., Stone, C. D. and Laschinger, J. C. (1997). *Immunohistochemical abnormalities of fibrillin in cardiovascular tissues in Marfan's syndrome*. Ann Thorac Surg **63**(4): 1012-7.

Fletcher, R. and Reeves, C. M. (1964). *Function minimization by conjugate gradients*. Comput J **7**(2): 149-154.

Floquet, N., Héry-Huynh, S., Dauchez, M., Derreumaux, P., Tamburro, A. M. and Alix, A. J. P. (2004). *Structural characterization of VGVAPG, an elastin-derived peptide*. Biopolymers **76**(3): 266-280.

Floquet, N., Pepe, A., Dauchez, M., Bochicchio, B., Tamburro, A .M. and Alix A. J. P. (2005). *Structure and modeling studies of the carboxy-terminus region of tropoelastin.* Matrix Biology, **24:** 271-282

Floquet, N. (2004). *Etudes structurales de peptides d'élastine et de collagène.* Thèse de doctorat, URCA.

Frishman, D. and Argos, P. (1995). *Knowledge-based secondary structure assignment*. Proteins **23**: 566-579.

Fuchs, P., Debelle, L. and Alix, A. J. P. (2001). *Structural study of some specific elastin hexapeptides activating MMP1*. J Mol Struc **565-566**: 335-339.

Fuchs, P., Bochicchio, B., Pepe, A., Bonvin, A., Alix, A. J. P. and Tamburro, A. M. (2003). *Spectroscopic and theoretical characterisation of a type VIII β-turn : Study of the tetrapeptide GDNP.* Elastin 2002**.** Eds A. M. Tamburro and A. Pepe, Editric Ermes, Potenza, Italia, pp 81-88.

Fuchs, P. F. J. and Alix, A. J. P. (2005). *High accuracy prediction of beta-turns and their types using propensities and multiple alignments*. Proteins: Structure, Function, Bioinformatics **59**(4): 828-839.

Fuchs, P. F. J., Bonvin, A. M. J. J., Bochicchio, B., Pepe, A., Alix, A. J. P. and Tamburro, A. M. (2005). *Kinetics and thermodynamics of type VIII β-turn formation: a CD, NMR and microsecond explicit molecular dynamics study of GDNP tetrapeptide*. **soumis.**

Fulop, T., Jr., Jacob, M. P., Khalil, A., Wallach, J. and Robert, L. (1998). *Biological effects of elastin peptides*. Pathol Biol (Paris) **46**(7): 497-506.

Galjart, N. J., Gillemans, N., Harris, A., van der Horst, G. T., Verheijen, F. W., Galjaard, H. and d'Azzo, A. (1988). *Expression of cDNA encoding the human "protective protein" associated with lysosomal beta-galactosidase and neuraminidase: homology to yeast proteases*. Cell **54**(6): 755-64.

Garnier, J., Osguthorpe, D. J. and Robson, B. (1978). *Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins*. J Mol Biol **120**(1): 97-120.

Gasteiger, J. and Marsili, M. (1980). *Iterative Partial Equalization of Orbital Electronegativity - A Rapid Access to Atomic Charges.* Tetrahedron **36**: 3219-3228.

Gibson, K. D. and Scheraga, H. A. (1967). *Minimization of Polypeptide energy, II. Preliminary Structures of Oxytocin, Vasopressin, and an Octapeptide from Ribonuclease*. Proc Natl Acad Sci **58**(4): 1317-1323.

Goodsell, D. S. and Olson, A. J. (1990). *Automated Docking of Substrates to Proteins by Simulated Annealing.* Proteins **8**: 195-202.

Gouliaev, D. I. (1998). *Etudes, modélisations et simulations moléculaires de protéines globulaires et / ou trans-membranaires (structures, hydrophobie).* Thèse de doctorat, Département de physique, Université de Reims Champagne Ardenne.

Grosso, L. E. and Scott, M. (1993a). *Peptide sequences selected by BA4, a tropoelastin-specific monoclonal antibody, are ligands for the 67-kilodalton bovine elastin receptor*. Biochemistry **32**(48): 13369-74.

Grosso, L. E. and Scott, M. (1993b). *PGAIPG, a repeated hexapeptide of bovine and human tropoelastin, is chemotactic for neutrophils and Lewis lung carcinoma cells*. Arch Biochem Biophys **305**(2): 401-4.

Grosso, L. E. and Scott, M. (1993c). *PGAIPG, a repeated hexapeptide of bovine tropoelastin, is a ligand for the 67-kDa bovine elastin receptor*. Matrix **13**(2): 157-64.

Guantieri, V., Grando, S., Pandolfo, L. and Tamburro, A. M. (1990). *Synthetic fragments and analogues of elastin. I. The synthesis*. Biopolymers **29**(4-5): 845-54.

Hauck, M., Seres, I., Kiss, I., Saulnier, J., Mohacsi, A., Wallach, J. and Fulop, T., Jr. (1995). *Effects of synthesized elastin peptides on human leukocytes*. Biochem Mol Biol Int **37**(1): 45-55.

Hinek, A. (1996). *Biological roles of the non-integrin elastin/laminin receptor*. Biol Chem **377**(7-8): 471-80.

Hinek, A., Keeley, F. W. and Callahan, J. (1995). *Recycling of the 67-kDa elastin binding protein in arterial myocytes is imperative for secretion of tropoelastin*. Exp Cell Res **220**(2): 312-24.

Hinek, A. and Rabinovitch, M. (1994). *67-kD elastin-binding protein is a protective "companion" of extracellular insoluble elastin and intracellular tropoelastin*. J Cell Biol **126**(2): 563-74.

Hinek, A., Rabinovitch, M., Keeley, F., Okamura-Oho, Y. and Callahan, J. (1993). *The 67-kD elastin/laminin-binding protein is related to an enzymatically inactive, alternatively spliced form of beta-galactosidase*. J Clin Invest **91**(3): 1198-205.

Hinek, A., Wrenn, D. S., Mecham, R. P. and Barondes, S. H. (1988). *The elastin receptor: a galactoside-binding protein*. Science **239**(4847): 1539-41.

Hopp, T. P. and Woods, K. R. (1981). *Prediction of protein antigenic determinants from amino acid sequences*. Proc Natl Acad Sci U S A **78**(6): 3824-8.

Hoover, W. G. (1985). *Canonical dynamics: Equilibrium phase-space distributions.* Phys Rev **A31**(3): 1695-1697.

Humphrey, W., Dalke, A. and Schulten, K. (1996). *VMD: visual molecular dynamics*. J Mol Graph **14**(1): 33-8, 27-8.

Hutchinson, E. G. and Thornton, J. M. (1994). *A revised set of potentials for beta-turn formation in proteins*. Protein Sci **3**(12): 2207-16.

Huynh, T., Smith, J. C. and Sanson, A. (2002). *Protein unfolding transitions in an intrinsically unstable annexin domain: molecular dynamics simulation and comparison with nuclear magnetic resonance data*. Biophys J **83**(2): 681-98.

Ilunga, K., Nishiura, R., Inada, H., El-Karef, A., Imanaka-Yoshida, K., Sakakura, T. and Yoshida, T. (2004). *Co-stimulation of human breast cancer cells with transforming growth factor-beta and tenascin-C enhances matrix metalloproteinase-9 expression and cancer cell invasion*. Int J Exp Pathol **85**(6): 373-9.

Indik, Z., Yeh, H., Ornstein-Goldstein, N., Sheppard, P., Anderson, N., Rosenbloom, J. C., Peltonen, L. and Rosenbloom, J. (1987). *Alternative splicing of human elastin mRNA indicated by sequence analysis of cloned genomic and complementary DNA*. Proc Natl Acad Sci U S A **84**(16): 5680-4.

Janin, J. (1979). *Surface and inside volumes in globular proteins*. Nature **277**(5696): 491-2.

Jacob, M. P. and Hornebeck, W. (1985). *Isolation and characterization of insoluble and kappa-elastins.* Front Matrix Biol **10**: 92-129.

Jian, B., Jones, P. L., Li, Q., Mohler, E. R., 3rd, Schoen, F. J. and Levy, R. J. (2001). *Matrix metalloproteinase-2 is associated with tenascin-C in calcific aortic stenosis*. Am J Pathol **159**(1): 321-7.

Johnson, W. C. (1999). *Analyzing protein circular dichroism spectra for accurate secondary structures*. Proteins **35**(3): 307-12.

Jones, G., Willett, P. and Glen, R. C. (1995). *Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation*. J Mol Biol **245**: 43-53.

Jones, D. T. (1999). *Protein secondary structure prediction based on position-specific scoring matrices*. J Mol Biol **292**(2): 195-202.

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. and Klein, M. L. (1983). *Comparison of simple potential functions for simulating liquid water*. J Chem Phys **79**: 926-935.

Kabsch, W. and Sander, C. (1983). *Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features*. Biopolymers **22**(12): 2577-2637.

Kalembeyi, I., Inada, H., Nishiura, R., Imanaka-Yoshida, K., Sakakura, T. and Yoshida, T. (2003). *Tenascin-C upregulates matrix metalloproteinase-9 in breast cancer cells: direct and synergistic effects with transforming growth factor beta1*. Int J Cancer **105**(1): 53-60.

Karpen, M. E., Tobias, D. J. and Brooks, C. L., 3rd (1993). *Statistical clustering techniques for the analysis of long molecular dynamics trajectories: analysis of 2.2-ns trajectories of YPGDV*. Biochemistry **32**(2): 412-20.

Karplus, M. (1959). *Contact electron-spin Coupling of Nuclear Magnetic Moments*. J Phys Chem **30**(11): 11-15.

Karplus, P. A. and Schulz, G. E. (1985). *Prediction of Chain Flexibility in Proteins : A Tool for the Selection of Peptide Antigens*. Naturwissenschaften **72**: 212-213.

Khaled, M. A. and Urry, D. W. (1976). *Nuclear Overhauser enhancement demonstration of the type II beta-turn in repeat peptides of tropoelastin*. Biochem Biophys Res Commun **70**(2): 485-91.

Kirkwood, F. G. and Westheimer, F. H. (1938). *The electrostatic influence of substituents on the dissociation constants of organic acids*. J Chem Phys **6**: 506-512.

Konig, W. and Geiger, R. (1970). *A new method for synthesis of peptides: activation of the carboxyl group with dicyclohexylcarbodiimide using 1-hydroxybenzotriazoles as additives*. Chem Ber **103**(3): 788-98.

Koradi, R., Billeter, M. and Wuthrich, K. (1996). *MOLMOL: a program for display and analysis of macromolecular structures*. J Mol Graph **14**(1): 51-5, 29-32.

Krimm, S. and Bandekar, J. (1986). *Vibrational spectroscopy and conformation of peptides, polypeptides, and proteins*. Adv Protein Chem **38**: 181-364.

Kuntz, I. D., Blaney, J. M., Oatley, S. J., Langridge, R. and Ferrin, T. E. (1982). *A geometric approach to macromolecule-ligand interactions.* J Mol Biol **161**(269-288).

Kyte, J. and Doolittle, R. F. (1982). *A simple method for displaying the hydropathic character of a protein*. J Mol Biol **157**(1): 105-32.

Lazaridis, T. and Karplus, M. (1999). *Effective Energy Function for Proteins in Solution.* Proteins **288**: 477-487.

Lee, M. S., Feig, M., Salsbury, F. R., Jr. and Brooks, C. L., 3rd (2003). *New analytic approximation to the standard molecular volume definition and its application to generalized Born calculations*. J Comput Chem **24**(11): 1348-1356.

Lelj, F., Tamburro, A. M., Villani, V., Grimaldi, P. and Guantieri, V. (1992). *Molecular dynamics study of the conformational behavior of a representative elastin building block: Boc-Gly-Val-Gly-Gly-Leu-OMe*. Biopolymers **32**(2): 161-72.

Li, B., Alonso, D. O., Bennion, B. J. and Daggett, V. (2001a). *Hydrophobic hydration is an important source of elasticity in elastin-based biopolymers*. J Am Chem Soc **123**(48): 11991-11998.

Li, B., Alonso, D. O. and Daggett, V. (2001b). *The molecular basis for the inverse temperature transition of elastin*. J Mol Biol **305**(3): 581-592.

Li, B., Alonso, D. O. and Daggett, V. (2002). *Stabilization of globular proteins via introduction of temperature-activated elastin-based switches*. Structure (Camb) **10**(7): 989-998.

Li, B. and Daggett, V. (2002). *Molecular basis for the extensibility of elastin*. J Muscle Res Cell Motil **23**(5-6): 561-573.

Li, B. and Daggett, V. (2003). *The molecular basis of the temperature- and pH-induced conformational transitions in elastin-based peptides*. Biopolymers **68**(1): 121-129.

Liotta, L. A., Rao, C. N. and Wewer, U. M. (1986). *Biochemical interactions of tumor cells with the basement membrane*. Annu Rev Biochem **55**: 1037-57.

Liptay, M. J., Parks, W. C., Mecham, R. P., Roby, J., Kaiser, L. R., Cooper, J. D. and Botney, M. D. (1993). *Neointimal macrophages colocalize with extracellular matrix gene expression in human atherosclerotic pulmonary arteries*. J Clin Invest **91**(2): 588-94.

Lograno, M. D., Bisaccia, F., Ostuni, A., Daniele, E. and Tamburro, A. M. (1998). *Identification of elastin peptides with vasorelaxant activity on rat thoracic aorta*. Int J Biochem Cell Biol **30**(4): 497-503.

Loncharich, R. J. and Brooks, B. R. (1989). *The effects of truncating long range forces on protein dynamics*. Proteins **6**(1): 32-45.

Long, M. M., Rapaka, R. S., Volpin, D., Pasquali-Ronchetti, I. and Urry, D. W. (1980). *Spectroscopic and electron micrographic studies on the repeat tetrapeptide of tropoelastin: (Val-Pro-Gly-Gly)n*. Arch Biochem Biophys **201**(2): 445-52.

MacArthur, M. W. and Thornton, J. M. (1991). *Influence of Proline Residues on Protein Conformation*. J Mol Biol **218**: 397-412.

MacKerrel, A. D. J., Bashford, D., Bellott, M., Dunbrack, R. L. J., Evanseck, J. D., Field, M. J., Fisher, S., Gao, J., Guo, S. H., Ha, S., Joseph-MacCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E. I., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D. and Karplus, M. (1998). *All-atom potential for molecular modeling and dynamics studies of proteins*. J Phys Chem B **102**: 3586-3616.

Maquart, F. X., Pasco, S., Ramont, L., Hornebeck, W. and Monboisse, J. C. (2004). *An introduction to matrikines: extracellular matrix-derived peptides which regulate cell activity. Implication in tumor invasion*. Crit Rev Oncol Hematol **49**(3): 199-202.

Maquart, F. X., Simeon, A., Pasco, S. and Monboisse, J. C. (1999). *Regulation of cell activity by the extracellular matrix: the concept of matrikines*. J Soc Biol **193**(4-5): 423-428.

Martino, M., Bavoso, A., Saviano, M., Di Blasio, B. and Tamburro, A. M. (1998). *Structure and dynamics of elastin building blocks. Boc-LG-OEt, Boc-VGG-OH*. J Biomol Struct Dyn **15**(5): 861-75.

Martino, M., Coviello, A. and Tamburro, A. M. (2000). *Synthesis and structural characterization of poly(LGGVG), an elastin-like polypeptide*. Int J Biol Macromol **27**(1): 59-64.

Martino, M., Perri, T. and Tamburro, A. M. (2002). *Elastin-based biopolymers: chemical synthesis and structural characterization of linear and cross-linked poly(OrnGlyGlyOrnGly)*. Biomacromolecules **3**(2): 297-304.

Martino, M. and Tamburro, A. M. (2001). *Chemical synthesis of cross-linked poly(KGGVG), an elastin-like biopolymer*. Biopolymers **59**(1): 29-37.

Mecham, R. P., Hinek, A., Entwistle, R., Wrenn, D. S., Griffin, G. L. and Senior, R. M. (1989). *Elastin binds to a multifunctional 67-kilodalton peripheral membrane protein*. Biochemistry **28**(9): 3716-22.

Mecham, R. P., Whitehouse, L., Hay, M., Hinek, A. and Sheetz, M. P. (1991). *Ligand affinity of the 67-kD elastin/laminin binding protein is modulated by the protein's lectin domain: visualization of elastin/laminin-receptor complexes with gold-tagged ligands*. J Cell Biol **113**(1): 187-94.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. and Teller, E. (1953). *Equation of State Calculations by Fast Computing Machines*. J Phys Chem **99**: 1087-1092.

Milner-White, E., Ross, B. M., Ismail, R., Belhadj-Mostefa, K. and Poet, R. (1988). *One type of gamma-turn, rather than the other gives rise to chain-reversal in proteins*. J Mol Biol **204**(3): 777-82.

Morreau, H., Galjart, N. J., Gillemans, N., Willemsen, R., van der Horst, G. T. and d'Azzo, A. (1989). *Alternative splicing of beta-galactosidase mRNA generates the classic lysosomal enzyme and a beta-galactosidase-related protein*. J Biol Chem **264**(34): 20655-63.

Morreau, H., Galjart, N. J., Willemsen, R., Gillemans, N., Zhou, X. Y. and d'Azzo, A. (1992). *Human lysosomal protective protein. Glycosylation, intracellular transport, and association with beta-galactosidase in the endoplasmic reticulum*. J Biol Chem **267**(25): 17949-56.

Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K. and Olson, A. J. (1998). *Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function*. J Comp Chem **19**: 1639-1662.

Muiznieks, L. D., Jensen, S. A. and Weiss, A. S. (2003). *Structural changes and facilitated association of tropoelastin*. Arch Biochem Biophys **410**(2): 317-323.

Némethy, G., Pottle, M. S. and Shegara, H. A. (1983). *Energy paramters in polypeptides, 9. Updating of geometrical parameters, nonbonded interactions and hydrogen bond interactions for naturally occuring amino acids.* J Phys Chem **87**: 1883-1887.

Norberg, J. and Nilsson, L. (2000). *On the Truncation of Long-Range Electrostatic Interactions in DNA*. Biophys J **79**: 1537-1553.

Nosé, S. (1984). *A unified formulation of the constant temperature molecular dynamics methods.* J Chem Phys **81**(1): 511-519.

Ostuni, A., Lograno, M. D., Gasbarro, A. R., Bisaccia, F. and Tamburro, A. M. (2002). *Novel properties of peptides derived from the sequence coded by exon 26A of human elastin*. Int J Biochem Cell Biol **34**(2): 130-135.

Pardi, A., Billeter, M. and Wuthrich, K. (1984). *Calibration of the angular dependence of the amide proton-C alpha proton coupling constants, 3JHN alpha, in a globular protein. Use of 3JHN alpha for identification of helical secondary structure*. J Mol Biol **180**(3): 741-51.

Parker, J. M., Guo, D. and Hodges, R. S. (1986). *New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites*. Biochemistry **25**(19): 5425-32.

Partridge, S. M., Davis, H. F. and Adair, G. S. (1955). *The chemistry of connective tissues. 2. Soluble proteins derived from partial hydrolysis of elastin*. Biochem J **61**(1): 11-21.

Pedanou, G. A. K. (1994). *Etudes quantitatives de la structure secondaire des protéines : méthode prédictives et spectroscopiques*. Thèse de doctorat, Département de physique, Université de Reims Champagne Ardenne.

Perczel, A., Hollosi, M., Sandor, P. and Fasman, G. D. (1993). *The evaluation of type I and type II beta-turn mixtures. Circular dichroism, NMR and molecular dynamics studies*. Int J Pept Protein Res **41**(3): 223-36.

Pitzer, R. M. (1983). *The barrier to internal rotation in ethan*. J Am Chem Soc **110**: 775-782.

Pollastri, G., Przybylski, D., Rost, B. and Baldi, P. (2002). *Improving the prediction of protein secondary structure in three and eight classes using recurrent neural networks and profiles*. Proteins **47**(2): 228-35.

Potier, M., Michaud, L., Tranchemontagne, J. and Thauvette, L. (1990). *Structure of the lysosomal neuraminidase-beta-galactosidase-carboxypeptidase multienzymic complex*. Biochem J **267**(1): 197-202.

Privitera, S., Prody, C. A., Callahan, J. W. and Hinek, A. (1998). *The 67-kDa enzymatically inactive alternatively spliced variant of beta-galactosidase is identical to the elastin/laminin-binding protein*. J Biol Chem **273**(11): 6319-26.

Provencher, S. W. and Glockner, J. (1981). *Estimation of globular protein secondary structure from circular dichroism*. Biochemistry **20**(1): 33-7.

Pyeritz, R. E. (2000). *The Marfan syndrome*. Annu Rev Med **51**: 481-510.

Qui, D., Shenkin, P. S., Hollinger, F. P. and Still, W. C. (1997). *The GB/SA Continuum Model for Solvation. A Fast Analytical Method for the Calculation of Approximate Born Radii*. J Phys Chem A **101**: 3005-3014.

Ramachandran, G. N., Ramakrishnan, C. and Sasisekharan, V. (1963). *Stereochemistry of polypeptide chain configurations*. J Mol Biol **7**: 95-99.

Raman, C. V. and Krishnan, K. S. (1928). *A new type of Secondary Radiation*. Nature **121**: 501-502.

Rapaka, R. S., Okamoto, K. and Urry, D. W. (1978). *Coacervation properties in sequential polypeptide models of elastin. Synthesis of H-(Ala-Pro-Gly-Gly)n-Val-OMe and H-(Ala-Pro-Gly-Val-Gly)n-Val-OMe*. Int J Pept Protein Res **12**(2): 81-92.

Rapaka, R. S. and Urry, D. W. (1978). *Coacervation of sequential polypeptide models of tropoelastin. Synthesis of H-(Val-Ala-Pro-Gly)n-Val-OMe and H-(Val-Pro-Gly-Gly)n-Val-OMe*. Int J Pept Protein Res **11**(2): 97-108.

Rarey, M., Wefing, S. and Lengauer, T. (1996). *Placement of medium-sized molecular fragments into active sites of proteins.* J Comput Aided Mol Des **10**(1): 41-54.

Ribeiro, S. M., Poczatek, M., Schultz-Cherry, S., Villain, M. and Murphy-Ullrich, J. E. (1999). *The activation sequence of thrombospondin-1 interacts with the latency-associated peptide to regulate activation of latent transforming growth factor-beta*. J Biol Chem **274**(19): 13586-13593.

Robert, L., Jacob, M. P. and Fulop, T. (1995). *Elastin in blood vessels*. Ciba Found Symp **192**: 286-99; discussion 299-303.

Rodgers, U. R. and Weiss, A. S. (2004). *Integrin alpha v beta 3 binds a unique non-RGD site near the C-terminus of human tropoelastin*. Biochimie **86**(3): 173-178.

Rose, G. D., Gierasch, L. M. and Smith, J. A. (1985). *Turns in peptides and proteins*. Adv Prot Chem **37**: 1-109.

Ross, R. and Bornstein, P. (1969). *The elastic fiber. I. The separation and partial characterization of its macromolecular components*. J Cell Biol **40**(2): 366-81.

Rudenko, G., Bonten, E., d'Azzo, A. and Hol, W. G. (1995). *Three-dimensional structure of the human 'protective protein': structure of the precursor form suggests a complex activation mechanism*. Structure **3**(11): 1249-59.

Ryckaert, J. P., Ciccotti, G. and Berendsen, H. J. C. (1977). *Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes*. J Comp Phys **23**: 327-341.

Sachdev, N. H., Di Girolamo, N., McCluskey, P. J., Jennings, A. V., McGuinness, R., Wakefield, D. and Coroneo, M. T. (2002). *Lens dislocation in Marfan syndrome: potential role of matrix metalloproteinases in fibrillin degradation*. Arch Ophthalmol **120**(6): 833-5.

Sayle, R. A. and Milner-White, E. J. (1995). *RASMOL: biomolecular graphics for all*. Trends Biochem Sci **20**(9): 374.

Schaumann, T., Braun, W. and Wuthrich, K. (1990). *The program FANTOM for energy refinement of polypeptides and proteins using a Newton - Raphson minimizer in torsion angle space.* Biopolymers **29**: 679-694.

Schiffer, C. A., Caldwell, J. W., Striud, R. M. and Kollman, P. A. (1992). *Inclusion of solvation free energy with molecular mechanics energy: alanyl dipeptide as a test case*. Protein Sci **1**: 396-400.

Segura, A. M., Luna, R. E., Horiba, K., Stetler-Stevenson, W. G., McAllister, H. A., Jr., Willerson, J. T. and Ferrans, V. J. (1998). *Immunohistochemistry of matrix metalloproteinases and their inhibitors in thoracic aortic aneurysms and aortic valves of patients with Marfan's syndrome*. Circulation **98**(19 Suppl): II331-7; discussion II337-8.

Senior, R. M., Griffin, G. L., Mecham, R. P., Wrenn, D. S., Prasad, K. U. and Urry, D. W. (1984). *Val-Gly-Val-Ala-Pro-Gly, a repeating peptide in elastin, is chemotactic for fibroblasts and monocytes*. J Cell Biol **99**(3): 870-4.

Senior, R. M., Hinek, A., Griffin, G. L., Pipoly, D. J., Crouch, E. C. and Mecham, R. P. (1989). *Neutrophils show chemotaxis to type IV collagen and its 7S domain and contain a*

67 kD type IV collagen binding protein with lectin properties. Am J Respir Cell Mol Biol **1**(6): 479-87.

Simonson, T. (2001). *Macromolcular electrostatics: continuum models and their growing pains*. Curr Opin Struct Biol **11**(2): 243-252.

Smekal, A. (1923). *Zur Quantentheorie der Dispersion*. Die Naturwissenschaften **11**: 873-875.

Solis, F. J. and R.J.-B., Wets. (1981). *Minimization by random search techniques.* Math Oper Res **6**: 19-30.

Simons, K. T., Bonneau, R., Ruczinski, I. and Baker, D. (1999). *Ab initio protein structure prediction of CASP III targets using ROSETTA.* Proteins **Suppl 3**: 171-6.

Sreerama, N. and Woody, R. W. (1993). *A self-consistent method for the analysis of protein secondary structure from circular dichroism*. Anal Biochem **209**(1): 32-44.

Sreerama, N. and Woody, R. W. (1999). *Molecular dynamics simulations of polypeptide conformations in water: A comparison of alpha, beta, and poly(pro)II conformations*. Proteins **36**(4): 400-6.

Sreerama, N. and Woody, R. W. (2000). *Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set*. Anal Biochem **287**(2): 252-60.

Steinbach, P. J. and Brooks, B. R. (1994). *New spherical-cutoff methods for long-range forces in macromolecular simulation.* J Comp Chem **15**: 667-683.

Stortz, C. A. (2004). *MM3 potential energy surfaces of alpha-3-linked L-fucobiose and fucotriose and their sulfated counterparts*. Carbohydr Res **339**(14): 2381-2390.

Tamburro, A. M., Bochicchio, B. and Pepe, A. (2003). *Dissection of human tropoelastin: exon-by-exon chemical synthesis and related conformational studies*. Biochemistry **42**(45): 13347-13362.

Tamburro, A. M. and Guantieri, V. (1991). *Classical and fractal description of elastin structure*. Ecological Physical Chemistry. Elsevier**:** 391-400.

Tamburro, A. M., Guantieri, V. and Gordini, D. D. (1992). *Synthesis and structural studies of a pentapeptide sequence of elastin. Poly (Val-Gly-Gly-Leu-Gly)*. J Biomol Struct Dyn **10**(3): 441-54.

Tamburro, A. M., Guantieri, V., Pandolfo, L. and Scopa, A. (1990). *Synthetic fragments and analogues of elastin. II. Conformational studies*. Biopolymers **29**(4-5): 855-70.

Tamburro, A. M., Guantieri, V., Scopa, A. and Drabble, J. M. (1991). *Polypeptide models of elastin: CD and NMR studies on synthetic poly(X-Gly-Gly)*. Chirality **3**(4): 318-23.

Tatham, A. S. and Shewry, P.R. (2000). *Elastomeric proteins: biological roles, structures and mechanisms.* Trends Biochem Sci **25**(11): 567-571.

Urry, D. W., Long, M. M., Cox, B. A., Ohnishi, T., Mitchell, L. W. and Jacobs, M. (1974a). *The synthetic polypentapeptide of elastin coacervates and forms filamentous aggregates*. Biochim Biophys Acta **371**(2): 597-602.

Urry, D. W., Long, M. M., Ohnishi, T. and Jacobs, M. (1974b). *Circular dichroism and absorption of the polytetrapeptide of elastin: a polymer model for the beta-turn*. Biochem Biophys Res Commun **61**(4): 1427-33.

Urry, D. W., Mitchell, L. W., Ohnishi, T. and Long, M. M. (1975a). *Proton and carbon magnetic resonance studies of the synthetic polypentapeptide of elastin*. J Mol Biol **96**(1): 101-17.

Urry, D. W., Onishi, T., Long, M. M. and Mitchell, L. W. (1975b). *Studies on the conformation and interactions of elastin: nuclear magnetic resonance of the polyhexapeptide*. Int J Pept Protein Res **7**(5): 367-78.

Urry, D. W. and Parker, T. M. (2002). *Mechanics of elastin: molecular mechanism of biological elasticity and its relationship to contraction*. J Muscle Res Cell Motil **23**(5-6): 543-559.

Verlet, L. (1967). *Computer "experiments" on classical fluids: I. Thermodynamical properties of Lennard-Jones molecules*. Phys Rev **159**: 98-103.

Vihinen, M., Torkkila, E. and Riikonen, P. (1994). *Accuracy of protein flexibility predictions*. Proteins **19**(2): 141-9.

Villani, V. and Tamburro, A. M. (1995). *Conformational modeling of elastin tetrapeptide Boc-Gly-Leu-Gly-Gly-NMe by molecular dynamics simulations with improvements to the thermalization procedure*. J Biomol Struct Dyn **12**(6): 1173-202.

Villani, V. and Tamburro, A. M. (1999). *Conformational chaos of an elastin-related peptide in aqueous solution*. Ann N Y Acad Sci **879**: 284-287.

von Freyberg, B. and Braun, W. (1993). *Efficient search for all low energy conformations of polypeptides by Monte Carlo methods*. J Comp Chem **12**: 1065-1076.

Vrhovski, B., Jensen, S. and Weiss, A. S. (1997). *Coacervation characteristics of recombinant human tropoelastin*. Eur J Biochem **250**(1): 92-8.

Vrhovski, B. and Weiss, A. S. (1998). *Biochemistry of tropoelastin*. Eur J Biochem **258**(1): 1-18.

Wagner, M., Adamczak, R., Porollo, A. and Meller, J. (2005). *Linear regression models for solvent accessibility prediction in proteins*. J Comput Biol **12**(3): 355-69.

Wallace, B. A. and Janes, R. W. (2001). *Synchrotron radiation circular dichroism spectroscopy of proteins: secondary structure, fold recognition and structural genomics*. Curr Opin Chem Biol **5**(5): 567-71.

Wedemeyer, W. J., Welker, E. and Scheraga, H. A. (2002). *Proline cis-trans isomerization and protein folding*. Biochemistry **41**(50): 14637-44.

Wei, S. M., Erdei, J., Fulop, T., Jr., Robert, L. and Jacob, M. P. (1993). *Elastin peptide concentration in human serum: variation with antibodies and elastin peptides used for the enzyme-linked immunosorbent assay*. J Immunol Methods **164**(2): 175-87.

Wei, S. M., Katona, E., Fachet, J., Fulop, T., Jr., Robert, L. and Jacob, M. P. (1998). *Epitope specificity of monoclonal and polyclonal antibodies to human elastin*. Int Arch Allergy Immunol **115**(1): 33-41.

Weiner, P. K. and Kollman, P. A. (1981). *AMBER: Assisted Model Building with Energy Refinement. A General Program for Modeling Molecules and Their Interactions.* J Comp Chem **2**: 287-303.

Weisshoff, H., Prasang, C., Henklein, P., Frommel, C., Zschunke, A. and Mugge, C. (1999). *Mimicry of beta II'-turns of proteins in cyclic pentapeptides with one and without D-amino acids*. Eur J Biochem **259**(3): 776-88.

Wodak, S. J. and Janin, J. (1980). *Analytical Approximation to the Solvent Accessible Surface Area of Proteins.* Proc Natl Acad Sci **77**: 1736-1740.

Wrenn, D. S., Hinek, A. and Mecham, R. P. (1988). *Kinetics of receptor-mediated binding of tropoelastin to ligament fibroblasts*. J Biol Chem **263**(5): 2280-4.

Xie, D. L., Hui, F., Meyers, R. and Homandberg, G. A. (1994). *Cartilage chondrolysis by fibronectin fragments is associated with release of several proteinases: stromelysin plays a major role in chondrolysis*. Arch Biochem Biophys **311**(2): 205-12.