



HAL
open science

Dynamique et toxicité de protéines avec une haute homologie de séquences : approche par dynamique moléculaire et calculs ab-initio.

Colette Foulie

► **To cite this version:**

Colette Foulie. Dynamique et toxicité de protéines avec une haute homologie de séquences : approche par dynamique moléculaire et calculs ab-initio.. Matière Condensée [cond-mat]. Université de Bourgogne, 2008. Français. NNT: . tel-00367817

HAL Id: tel-00367817

<https://theses.hal.science/tel-00367817v1>

Submitted on 12 Mar 2009

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

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

THÈSE

présentée devant l'

Université de Bourgogne

pour obtenir le grade de :

DOCTEUR DE L'UNIVERSITÉ DE BOURGOGNE
Spécialité PHYSIQUE

par

Colette Foulie

École doctorale : CARNOT

Laboratoire : ICB UMR CNRS 5209

Équipe d'accueil : Théorie de la Matière Condensée

Titre de la thèse :

**Dynamique et toxicité de protéines avec une haute
homologie de séquences : approche par dynamique
moléculaire et calculs ab-initio.**

Soutenue le 28 Octobre 2008 devant la commission d'examen composée de :

Mme : Carmen	GARRIDO	Président
Mme : Marie-Claire	BELLISSENT-FUNEL	Rapporteur
M. : Christian	VANALSENOY	Rapporteur
M. : Guillaume	MAURIN	Examineur
M. : Patrice	DELARUE	Examineur
M. : Patrick	SENET	Directeur de thèse

Remerciements

Le travail présenté dans ce mémoire a été mené au sein de l'Institut Carnot de Bourgogne dans l'équipe 'Théorie de la Matière Condensée'.

En premier lieu, je tiens à remercier mon directeur de thèse, Monsieur Patrick Senet, Professeur à l'Université de Bourgogne, pour m'avoir fait confiance tout au long de ces trois années. Il m'a laissée autonome tout en croyant en mes capacités. Sa confiance m'a apporté des compétences qui, j'en suis sûre, me seront très bénéfiques dans ma future vie active.

Mes remerciements vont aussi à Patrice Delarue, Maître de Conférences à l'Université de Bourgogne, pour la disponibilité et la patience qu'il a manifestées à mon égard, mais aussi pour tous les conseils qu'il m'a prodigués et les programmes qu'il a bien voulu m'envoyer.

J'ai été honorée d'apprendre que Madame Carmen Garrido, Directeur de recherches à l'INSERM, ait accepté le rôle d'examineur de ce travail et de présider le jury.

Madame Marie-Claire Bellissent-Funel, Directeur de recherches au CEA Saclay, m'a fait l'honneur de juger ce travail et d'en être rapporteur. Je la remercie d'avoir accepté et d'avoir mené à bien ce travail avec diligence et intérêt.

Monsieur Kris Van Alsenoy, Professeur à l'Université d'Anvers, m'a fait le plaisir d'être rapporteur de cette thèse. Je le remercie également de m'avoir accueillie au sein de son laboratoire en juillet dernier.

Monsieur Guillaume Maurin, Maître de Conférences à l'Université de Montpellier, a accepté de participer au jury. Je l'en remercie.

Je tiens également à remercier toutes les personnes que j'ai pu rencontrer durant ces trois années, au laboratoire ou pendant mes déplacements, qui ont su apporter leur contribution à mon travail par des discussions fructueuses.

Un grand merci à Nicole Rouget et Elly Lacroute pour leur aide administrative.

Tous mes remerciements s'adressent aussi à mes collègues thésards : tout d'abord Mélanie, avec qui j'ai partagé mon bureau pendant un an, très certainement la meilleure année de ma thèse. Je te souhaite beaucoup de réussite dans le futur. Maher, merci pour tous ces moments passés ensemble, ma thèse n'aurait pas été la même sans nos "pauses café" mais aussi pour ton écoute, j'espère que tu trouveras bientôt ta voie. Loïc, merci pour ton aide en Latex et ta patience. Mes pensées vont également vers Maxime, bon courage pour ta thèse qui débute, je te souhaite beaucoup de bonheur et de réussite. Je n'oublie pas non plus tous les autres thésards : à tous, bon courage pour la suite.

Je termine ces remerciements par les personnes qui me sont les plus chères : Ju, ta présence dans ma vie s'est imposée à moi comme une évidence. Tu as su apaiser mes tensions, répondre à mes interrogations et mes doutes avec une patience infinie. Merci, merci d'être là ; mes frères, JB et Denis, leur dérision et leur humour m'ont permis de prendre du

recul sur les aléas de la recherche, ils ont toujours eu les mots pour détendre l'atmosphère ; Emeline, un grand merci pour tout l'intérêt que tu as su porter à ma thèse mais aussi à ma vie de tous les jours ; Agnès, cela fait maintenant plusieurs années que tu es à mes côtés, je te remercie pour tous ces moments que nous avons passés ensemble et pour tous ceux à venir.

Je pense également à toute ma famille et mes amis mais tout particulièrement à mes parents. Je n'aurai jamais pu réaliser de telles études et aller au bout de mes ambitions sans leur soutien. Ils m'ont toujours encouragée, écoutée et soutenue. Je tiens à leur exprimer ma plus profonde gratitude pour la confiance aveugle qu'ils ont eue en moi. Cette thèse, je vous la dédie.

A mes parents.

Table des matières

Table des matières	i
Liste des figures	vii
Introduction	1
I Méthodes et systèmes étudiés	3
1 Techniques de simulations	5
1.1 Propriétés macroscopiques - physique statistique	7
1.1.1 Définitions d'un état microscopique et des ensembles statistiques . . .	7
1.1.2 Ergodicité et propriétés statistiques	8
1.2 Surface d'énergie potentielle	9
1.2.1 Approximation de Born-Oppenheimer	9
1.2.2 Définitions des différents termes d'énergie	10
1.2.3 Les différents types de champ de force	13
1.2.4 CHARMM	14
1.3 Dynamique et mécanique moléculaire	15
1.3.1 Equations de mouvement	16
1.3.2 Résolutions numériques des équations de mouvement : Algorithmes de Verlet	16
1.3.3 Algorithmes pour maintenir T ou P constantes	19
1.3.4 Conditions périodiques	22
1.3.5 Minimisation de l'énergie	23
2 Analyse des trajectoires et des fluctuations	25
2.1 Structure moyenne et trajectoire de référence	27
2.2 Fluctuations structurales	29
2.2.1 Variance des déplacements atomiques au Mean - Square - Displacement (MSD)	29
2.2.2 Coordonnées relatives au centre de masse	29
2.2.3 Rayon de giration et rayon de giration moyen	30
2.2.4 Déplacement quadratique moyen	31

2.2.5	Fonction de corrélation spatiale	31
2.3	Variables conformationnelles à gros grains γ caractérisant le backbone	32
2.4	Méthode d'analyse DFA (Detrended Fluctuation Analysis) et exposant fractal des séries temporelles $\gamma(t)$	32
2.5	Profil d'énergie libre pour les angles diédraux γ	34
3	Mesures expérimentales des fluctuations structurales	35
3.1	Diffusion incohérente de neutrons	37
3.2	Mesure de l'agitation thermique en diffraction X	39
3.3	Comparaison des fluctuations structurales mesurées et calculées	40
3.4	La Résonance Magnétique Nucléaire	41
4	Protéines étudiées et protocoles de simulations	43
4.1	Généralités sur les protéines	45
4.2	Structure type des thionines	46
4.3	Protocoles de simulation	49
4.3.1	Hydratation	49
4.3.2	Protocole dans le vide	51
4.3.3	Protocole dans l'eau	51
4.3.4	Viscotoxines avec tyrosines iodinisées	53
4.3.5	Tests sur les protocoles de simulation	55
5	Structures optimisées - Comparaisons à l'expérience	59
5.1	Calculs des structures de départ	61
5.1.1	Simulation de recuit et Calibrage	61
5.1.2	Résultats	63
5.2	Analyses des structures équilibrées	65
5.3	Comparaisons avec les techniques de RMN et de diffraction X	66
II	Dynamique conformationnelle et transition dynamique	69
1	Introduction	71
1.1	La transition dynamique des protéines	73
1.2	Hydratation des protéines	77
2	Dynamique des protéines déshydratées	79
2.1	Modèles simplifiés de la transition dynamique	81
2.1.1	Modèle de Doster	81
2.1.2	Modèle harmonique à deux états	84
2.2	Flexibilité conformationnelle de protéines déshydratées	86
2.2.1	Résultats pour la transition dans le vide des crambines et des viscotoxines	86
2.2.2	Désordre dynamique - inhomogénéité des fluctuations	100
2.2.3	Origine des effets induits par les mutations	115
2.3	Conclusion	117

3	Dynamique des protéines hydratées : rôle de l'eau	119
3.1	Résultats pour la transition dynamique des crambines et des viscotoxines . . .	121
3.1.1	Dépendance par rapport aux conditions initiales	121
3.1.2	Paramètres tirés des modèles	121
3.1.3	Séparation des contributions en type d'atomes	125
3.1.4	Dépendance en temps de simulations	129
3.1.5	Comparaisons vide et hydratation	130
3.1.6	Superpositions des structures à 100K et 300K	131
3.2	Analyse des types de mouvements et inhomogénéité des fluctuations	131
3.2.1	Dispersion des fluctuations	131
3.2.2	Identifications des atomes mobiles	133
3.2.3	Mobilité des résidus	136
3.2.4	Analyses des types de mouvements	139
3.3	Rôle de l'eau : analyse des paramètres macroscopiques	141
3.3.1	Rayons de Giration	141
3.3.2	Variation du volume	146
3.3.3	Variation du volume de la boîte de simulation	148
3.3.4	Etude de la répartition de l'eau autour des protéines	149
3.3.5	Variation du taux d'hydratation	152
3.3.6	Discussion	158
3.4	Origine de l'effet des mutations	159
3.4.1	Modification du réseau de liaisons hydrogène	159
3.5	Fluctuations structurales et paysage d'énergie libre en solution	161
3.5.1	Définition du paysage d'énergie libre	161
3.5.2	Paysage d'énergie libre et dynamique du backbone	163
3.5.3	Dynamique du backbone et PMF à 300K	165
3.5.4	Transition dynamique du backbone	173
3.5.5	Effet de l'hydratation sur la structure	182
3.6	Modèle de la transition dynamique	186
4	Les ions comme sondes de la transition dynamique	189
III	Dynamique et cytotoxicité	193
1	Données sur les interactions membrane/protéine	195
1.1	Point de vue expérimental	198
1.2	Point de vue théorique	199
2	Dynamique de protéines avec tyrosines iodinisées	201
2.1	Courbes moyennes générales	203
2.2	Mobilité des atomes	204
2.3	Etude détaillée de la tyrosine 13	204
2.4	Mobilité des résidus	207
2.5	Conclusion	207
	Conclusion	211

A	Protocoles dans le vide	215
A.1	Paramètres et programme utilisés pour l'étape de chauffage	217
A.1.1	Paramètres	217
A.1.2	Programme Charmm	218
A.2	Paramètres et programme utilisés pour l'étape d'équilibrage	219
A.2.1	Paramètres	219
A.2.2	Programme Charmm	219
A.3	Paramètres et programme utilisés pour l'étape de production	221
A.3.1	Paramètres	221
A.3.2	Programme Charmm	221
A.4	Programme Charmm pour le calcul du MSD des protéines dans le vide . . .	222
B	Protocoles des protéines hydratées	227
B.1	Programme Charmm utilisé pour l'insertion de la protéine dans la boîte d'eau	229
B.2	Programme Charmm pour l'insertion d'ions dans la boîte d'eau dans le cas des viscotoxines hydratées	232
B.3	Paramètres et programme utilisés pour l'étape de chauffage	233
B.3.1	Paramètres	233
B.3.2	Programme Charmm	234
B.4	Paramètres et programme utilisés pour l'étape d'équilibrage	236
B.4.1	Paramètres	236
B.4.2	Programme Charmm	236
B.5	Programme utilisé dans l'étape de production	238
B.6	Programme Charmm pour le calcul du MSD des protéines hydratées	240
B.6.1	Réorientation des trajectoires	240
B.6.2	Calcul du MSD	241
C	Simulation de recuit	243
C.1	Calibrage de la température maximale de chauffage	245
C.2	Résultats	245
C.2.1	Recuit simulé des viscotoxines	245
C.2.2	Recuit simulé des crambines	246
D	Protéines dans le vide : Séparation des contributions en type d'atomes	247
E	Protéines dans le vide : Dépendance en temps de simulations	251
F	Dispersion des fluctuations des protéines déshydratées	255
G	Augmentation de la mobilité des atomes dans le vide avec la température	259
H	Dépendance de la mobilité avec le temps de production des protéines dans le vide	267
I	Mobilité des résidus dans le vide	275

J	Modèle de Doster	279
J.1	Ajustement des paramètres des thionines hydratées	281
J.2	Deuxième ajustement pour VA3	282
K	Distribution des barrières d'énergie des méthyles	283
L	Variation de la mobilité des protéines hydratées en fonction des 5 conditions initiales utilisées sur des trajectoires de 1ns	287
M	Protéines hydratées : Séparation des contributions en type d'atomes	291
N	Protéines hydratées : Dépendance en temps de simulations	295
O	Dispersion des fluctuations des protéines hydratées	299
P	Augmentation de la mobilité des atomes des protéines hydratées avec la température	303
Q	Dépendance de la mobilité avec le temps de production des protéines hydratées	311
R	Mobilité des résidus des protéines hydratées	319
S	Répartition de l'eau autour des protéines	323
T	Programmes Fortran	327
T.1	Représentation des hydrogènes les plus mobiles par des sphères	329
T.2	Calcul du rayon de giration le long d'une trajectoire	331
	Bibliographie	339

Liste des figures

1.1	Conditions périodiques illustrées en deux dimensions avec une boîte cubique.	22
1.2	Inconvénient des méthodes de gradient : le minimum local atteint par les deux boules a et b est différent et dépend de la position (configuration) initiale.	23
3.1	Fenêtre de résolution expérimentale où T est le temps de résolution temporelle limité par la résolution en énergie du spectromètre.	38
3.2	Etude de la myoglobine : symboles non-colorés : Mesure de l'écart quadratique moyen par diffusion de neutrons : cercles : pour une résolution de 80ps, triangles : pour une résolution de 12ps. cercles noirs : Ecart quadratique moyen mesuré par spectroscopie de Mössbauer.	39
3.3	Dépendance en température de l'écart quadratique moyen des résidus de la myoglobine obtenue par les modes normaux [Chong <i>et al.</i> 01]. Symboles noirs et lignes continues : moyenne sur tous les atomes du backbone, symboles non-colorés et lignes en pointillés : moyenne sur les atomes des chaînes latérales [Parak 03].	40
4.1	Formule générale d'un acide aminé. C^α est un carbone asymétrique (sauf pour la Glycine) et R représente la chaîne latérale de l'acide aminé.	45
4.2	Backbone de la Viscotoxine A3, position des structures secondaires et structure typique en L, représentation avec le logiciel VMD. N indique l'extrémité de la protéine correspondante au résidu 1.	46
4.3	Localisation des acides aminés le long du squelette.	47
4.4	Réaction de mono-iodination sur une tyrosine.	53
4.5	Tests sur le temps de chauffe entre 0 et 220K pour VA3 isolée : la courbe noire correspond à un temps de chauffe de 2ps, la courbe rouge à 4ps, la courbe verte à 8ps et la courbe bleue à 16ps.	55
4.6	Comparaisons des protocoles de chauffe pour VA3 déshydratée (vide) : <i>chauffe de 200 à 220K</i> : la courbe noire correspond à un temps de chauffe de 2ps, la courbe rouge à 4ps ; <i>chauffe de 0 à 220K</i> : la courbe bleue correspond à un temps de chauffe de 2ps, la courbe verte à 4ps.	56
4.7	Comparaisons des protocoles de chauffe pour P22/L25 déshydratée. La courbe noire correspond à un chauffage par paliers et la courbe rouge à un chauffage de 0K à T_{voulue} .	57
4.8	Formation de "trous" dans l'ensemble NVT, cas de P22/L25 hydratée.	58

5.1	Principe de la simulation de recuit.	62
5.2	Recuit simulé de VA3 (représentation en ruban des squelettes) : (a) Superposition de la structure finale obtenue à 298K (en jaune) à la structure optimisée (en bleu). (b) Superposition de la structure finale obtenue à 298K (en jaune) à la structure expérimentale (en violet).	64
5.3	(a) Superposition des crambines : P22/L25 en bleu, S22/I25 en rouge, P22/I25 en vert et S22/L25 en violet. (b) Superposition des viscotoxines : VA2 en noir, VA3 en orange et VB en jaune.	65
5.4	Superposition de la viscotoxine A3 (en orange) et de la crambine P22/L25 (en bleu).	66
5.5	MSD des atomes lourds calculés pour P22/L25 hydratée à 300K par dynamique moléculaire à gauche et à partir des modèles RMN à droite.	67
5.6	MSD des carbones alpha calculés pour P22/L25 hydratée à 150K par dynamique moléculaire à gauche et à partir des structures DRX à droite.	67
1.1	Courbe typique représentant la transition dynamique, obtenue dans les premiers travaux de Parak [Parak <i>et al.</i> 80] où $\langle x^2 \rangle$ est le déplacement quadratique moyen de l'atome de Fer du groupe heme de la metmyoglobine. . .	74
1.2	Variation de la température de dénaturation T_d et de la largeur des pics de la température de dénaturation ΔT_d du lysozyme en fonction du taux d'hydratation [Fujita et Noda 78].	77
2.1	Barrière de potentiel ΔG séparant deux puits de potentiels. Dans l'ensemble NPT, $\Delta G = \Delta H - T\Delta S$, où $\Delta H = \Delta U + P\Delta V$. Dans l'ensemble NVT, l'énergie libre est notée F , telle que $\Delta F = \Delta U - T\Delta S$	81
2.2	Ajustement des paramètres du modèle de Doster pour la viscotoxine VA3 hydratée. La courbe noire est l'écart quadratique moyen calculé par nos simulations de dynamique moléculaire et la courbe rouge est le modèle de Doster.	83
2.3	Variation du puits de potentiel entre l'état R, à basses températures, et l'état F, à hautes températures.	85
2.4	Evolution de l'écart quadratique moyen de P22/L25 dans le vide, sur une trajectoire de 1ns, les lignes rouges en traits interrompus correspondent aux régressions linéaires dans les états R et F.	86
2.5	Evolution de l'écart quadratique moyen des thionines dans le vide pour une trajectoire de 1ns. Les lignes rouges en traits interrompus représentent la régression linéaire de la courbe pour l'état R, nécessaire pour le calcul des constantes effectives.	88
2.6	Evolution de l'écart quadratique moyen des différents types d'atomes pour les crambines dans le vide sur une trajectoire de 1ns. La courbe rose est l'écart quadratique moyen de tous les atomes qui ne sont pas des hydrogènes, la courbe noire est la même variation pour les hydrogènes et la ligne rouge en traits interrompus est la régression linéaire de la courbe dans l'état R.	90

2.7	Evolution de l'écart quadratique moyen des différents types d'atomes pour les viscotoxines dans le vide sur une trajectoire de 1ns. La courbe rose représente l'écart quadratique moyen de tous les atomes qui ne sont pas des hydrogènes, la courbe noire représente la variation pour les hydrogènes seulement et la ligne rouge en traits interrompus est la régression linéaire de la courbe dans l'état R.	91
2.8	Evolution de l'écart quadratique moyen de la crambine P22/L25 (courbe bleue), de ses atomes d'hydrogène (courbe noire) et de tous ses autres atomes (courbe rose), dans le vide et sur une trajectoire de 1ns.	92
2.9	Dépendance en temps de simulations de l'écart quadratique moyen de la crambine P22/L25 dans le vide, la courbe noire représente une trajectoire de 50ps, la courbe rouge de 100ps, la courbe bleue de 200ps et la courbe verte de 1ns.	93
2.10	Rayon de giration des crambines dans le vide, sur une trajectoire de 1ns : à gauche : Rgyr (Eq. 2.23), à droite : Rgmean (Eq. 2.24). Les triangles verts correspondent à S22/L25, les ronds rouges à S22/I25, les triangles bleus à P22/I25 et les carrés noirs à P22/L25.	95
2.11	Rayon de giration des crambines dans le vide, en éliminant tous les atomes d'hydrogène, sur une trajectoire de 1ns : à gauche : Rgyr (Eq. 2.23), à droite : Rgmean (Eq. 2.24). Les triangles verts correspondent à S22/L25, les ronds rouges à S22/I25, les triangles bleus à P22/I25 et les carrés noirs à P22/L25.	95
2.12	Rayon de giration de la viscotoxine A3 (carrés noirs), A2 (ronds rouges) et B (triangles bleus) dans le vide, sur une trajectoire de 1ns : à gauche : Rgyr, à droite : Rgmean.	96
2.13	Rayon de giration de la viscotoxine A3 (carrés noirs), A2 (ronds rouges) et B (triangles bleus) dans le vide, en éliminant tous les atomes d'hydrogène, sur une trajectoire de 1ns : à gauche : Rgyr, à droite : Rgmean.	96
2.14	Comparaison entre l'évolution de l'écart quadratique moyen (carrés bleus) de la viscotoxine A3 par rapport à son rayon de giration (carrés noirs), dans le vide et sur une trajectoire de 1ns.	97
2.15	Evolution de l'écart quadratique moyen de la crambine S22/I25 dans le vide, sur une trajectoire de 1ns et jusqu'à 450K.	99
2.16	Variation de $\frac{\sigma^2}{T^2}$ (carrés noirs) et de $\frac{MSD}{NT}$ (carrés bleus) en fonction de la température pour la crambine P22/L25 dans le vide, sur une trajectoire de 50ps à gauche et sur une trajectoire de 1ns à droite. Les lignes rouges en traits interrompus correspondent aux régressions linéaires des états R et F.	101
2.17	Comparaisons de la flexibilité globale des crambines (P22/L25 : carrés noirs, S22/I25 : ronds rouges, P22/I25 : triangles bleus, S22/L25 : triangles verts), à gauche, et des viscotoxines (VA3 : carrés roses, VA2 : ronds verts, VB : triangles oranges), à droite, dans le vide et sur une trajectoire de 1ns.	102
2.18	Superpositions des crambines équilibrées dans le vide, à 100K à gauche, et à 300K à droite : P22/L25 en bleu, S22/I25 en rouge, P22/I25 en vert et S22/L25 en violet. C indique l'extrémité de la protéine correspondante au résidu 46. La flèche localise l'acide aminé n°22.	103

2.19	Superpositions des viscotoxines équilibrées dans le vide, à 100K à gauche, et à 300K à droite : VA3 en orange, VA2 en noir et VB en jaune. C indique l'extrémité de la protéine correspondante au résidu 46. La flèche localise l'acide aminé n°22.	103
2.20	Variation de la mobilité des atomes avec la température de la crambine P22/L25 dans le vide sur une trajectoire de 1ns.	105
2.21	Variation de la mobilité des atomes avec le temps de simulation à 300K pour la crambine P22/L25 dans le vide sur une trajectoire de 1ns.	106
2.22	Variation du nombre de résidus suivant leur valeur du $\frac{MSD}{T}$ pour la crambine P22/L25 dans le vide, sur une trajectoire de 1ns. Les histogrammes correspondent aux températures suivantes : noir = 5K, rouge = 30K, bleu = 70K, bleu-vert = 100K, fushia = 130K, kaki = 150K, bleu marine = 170K, marron = 200K, rose = 220K, vert = 250K, gris = 280K, orange = 300K.	108
2.23	Variation du MSD de chaque acide aminé avec la température pour la crambine P22/L25 dans le vide, sur une trajectoire de 1ns.	109
2.24	Mouvements possibles avec l'augmentation de la température, les sphères bleues représentent des hydrogènes, les sphères rouges des atomes de carbone : (a) basculement d'un côté à l'autre, (b) rotation des 3 hydrogènes autour du carbone.	110
2.25	Localisation des hydrogènes mobiles par des sphères, de diamètre proportionnel à leur RMS, pour crambine P22/L25 déshydratée, sur une trajectoire de 1ns.	111
2.26	Localisation des hydrogènes mobiles par des sphères, de diamètre proportionnel à leur RMS, pour les thionines déshydratées à 300K, sur une trajectoire de 1ns.	112
2.27	Barrières d'énergie de rotation des méthyles de la crambine P22/L25 optimisée.	113
2.28	Distribution des barrières d'énergie pour la crambine P22/L25 optimisée.	114
2.29	Evolution du MSD moyen des groupements méthyles (calculés comme la moyenne sur ses 3 atomes d'hydrogène) en fonction de la hauteur moyenne de la barrière d'énergie de rotation. Les carrés bleus correspondent à 100K, les ronds verts à 200K et les triangles rouges à 300K.	114
3.1	Variation de l'écart quadratique moyen de P22/L25 hydratée sur une trajectoire de 1ns. Les carrés noirs correspondent à la 1ère condition initiale, les ronds rouges à la 2nde, les triangles bleus à la 3ième, les losanges verts à la 4ième et les triangles roses à la 5ième.	121
3.2	Evolution de la moyenne de l'écart quadratique moyen des crambines hydratées, calculée sur 5 conditions initiales. Les lignes rouges en traits interrompus correspondent aux régressions linéaires de la courbe pour l'état R.	123
3.3	Evolution de la moyenne de l'écart quadratique moyen des viscotoxines hydratées, calculée sur 5 conditions initiales. Les lignes rouges en traits interrompus correspondent aux régressions linéaires de la courbe pour l'état R.	124
3.4	A gauche, Flexibilité globale des crambines hydratées (P22/L25 : carrés noirs, S22/I25 : ronds rouges, P22/I25 : triangles bleus, S22/L25 : triangles verts), et à droite des viscotoxines hydratées (VA3 : carrés roses, VA2 : ronds verts, VB : triangles oranges). Moyenne du MSD calculée sur 5 trajectoires de 1ns.	125

3.5	Evolution de l'écart quadratique moyen des différents types d'atomes pour les crambines hydratées, moyennes calculées sur 5 trajectoires de 1ns. La courbe rose représente l'écart quadratique moyen de tous les atomes qui ne sont pas des hydrogènes, la courbe noire représente cette même variation pour les hydrogènes seulement et la ligne rouge en traits interrompus est la régression linéaire de la courbe dans l'état R.	127
3.6	Evolution de l'écart quadratique moyen des différents types d'atomes pour les viscotoxines hydratées, moyennes calculées sur 5 trajectoires de 1ns. La courbe rose représente l'écart quadratique moyen de tous les atomes qui ne sont pas des hydrogènes, la courbe noire représente cette même variation pour les hydrogènes seulement et la ligne rouge en traits interrompus est la régression linéaire de la courbe dans l'état R.	128
3.7	Evolution de l'écart quadratique moyen de la crambine P22/L25 (courbe bleue), de ses atomes d'hydrogène (courbe noire) et de tous ses autres atomes (courbe rose), hydratée et sur la moyenne des 5 trajectoires de 1ns.	129
3.8	Dépendance en temps de simulations de l'écart quadratique moyen de la crambine P22/L25 hydratée, la courbe noire représente une trajectoire de 50ps, la courbe rouge de 100ps, la courbe bleue de 200ps et la courbe verte de 1ns.	130
3.9	Comparaisons de la variation du MSD des protéines simulées dans le vide (courbes noires) et dans une boîte d'eau (courbes bleues) pour P22/L25 à gauche et VA3 à droite. Les lignes rouges en traits interrompus correspondent aux régressions linéaires de la courbe pour l'état R des protéines hydratées.	130
3.10	Variation de $\frac{\sigma^2}{T^2}$ (carrés noirs) et de $\frac{MSD}{NT}$ (carrés bleus) en fonction de la température pour la crambine P22/L25 hydratée, sur une trajectoire de 50ps à gauche et sur une trajectoire de 1ns à droite. Les lignes rouges en traits interrompus correspondent aux régressions linéaires des états R et F.	131
3.11	Superpositions des crambines équilibrées et hydratées, à 100K à gauche et à 300K à droite : P22/L25 en bleu, S22/I25 en rouge, P22/I25 en vert et S22/L25 en violet. C indique l'extrémité de la protéine correspondante au résidu 46. La flèche localise l'acide aminé n°22.	132
3.12	Superpositions des viscotoxines équilibrées et hydratées, à 100K à gauche et à 300K à droite : VA3 en orange, VA2 en noir et VB en jaune. C indique l'extrémité de la protéine correspondante au résidu 46. La flèche localise l'acide aminé n°22.	132
3.13	Variation de la mobilité des atomes avec la température de la crambine P22/L25 hydratée sur une trajectoire de 1ns.	134
3.14	Variation de la mobilité des atomes avec le temps de simulation pour la crambine P22/L25 hydratée à 300K.	135
3.15	Variation du nombre de résidus suivant leur valeur du $\frac{MSD}{T}$ pour la crambine P22/L25 hydratée, sur une trajectoire de 1ns. Les histogrammes correspondent aux températures suivantes : noir = 5K, rouge = 30K, bleu = 70K, bleu-vert = 100K, fushia = 130K, kaki = 150K, bleu marine = 170K, marron = 200K, rose = 220K, vert = 250K, gris = 280K, orange = 300K.	138
3.16	Variation du MSD de chaque acide aminé avec la température pour la crambine P22/L25 simulée dans le vide à gauche et hydratée à droite, sur une trajectoire de 1ns.	139

3.17	Comparaisons entre deux acides aminés, la tyrosine à gauche, l'alanine à droite, de l'encombrement spatial des chaînes latérales.	140
3.18	Variation du rayon de giration de la crambine P22/L25 dans l'eau (ronds bleus), et à partir des structures expérimentales de M. Teeter (carrés verts et rouges). Les résultats expérimentaux représentés par les carrés verts ont été obtenus sur le même échantillon.	142
3.19	Variation des rayons de giration pour P22/L25 à gauche, S22/I25 à droite : dans l'eau (ronds bleus), dans le vide (carrés violets) et à partir des structures DRX de M. Teeter (carrés verts et rouges), comparés à la variation de l'écart quadratique moyen (carrés noirs).	142
3.20	Variation du rayon de giration de P22/I25 à gauche et S22/L25 à droite dans le vide (carrés violets) et dans l'eau (ronds bleus).	143
3.21	Comparaison de la variation du rayon de giration à partir des résultats expérimentaux de M. Teeter (carrés verts et rouges) et des protéines hydratées : P22/L25 (ronds rouges), S22/I25 (triangles bleus), P22/I25 (losanges marrons), S22/L25 (étoiles roses)	144
3.22	Comparaison de l'évolution du rayon de giration sur une trajectoire de 100ps (courbe noire) et une trajectoire de 1ns (courbe rouge) pour crambine P22/L25 hydratée.	144
3.23	Variation du rayon de giration des viscotoxines dans l'eau (ronds bleus) et dans le vide (carrés violets).	145
3.24	Variation du rayon de giration des viscotoxines hydratées : VA3 (carrés noirs), VA2 (ronds oranges) et VB (triangles verts).	146
3.25	Variation du rapport entre le volume de la protéine (V_p) et le volume de la boîte ($V_{bo\grave{t}e}$) pour les crambines hydratées : P22/L25 (ronds rouges), S22/I25 (triangles bleus), P22/I25 (losanges marrons), S22/L25 (étoiles roses)	147
3.26	Variation du rapport entre le volume des viscotoxines hydratées et celui de la boîte d'eau : VA3 (carrés noirs), VA2 (ronds oranges) et VB (triangles verts).	147
3.27	Variation du volume des boîtes de simulation avec l'augmentation de la température pour les crambines à gauche : P22/L25 (carrés noirs), S22/I25 (ronds rouges), P22/I25 (triangles bleus), S22/L25 (triangles verts) et pour les viscotoxines à droite : VA3 (carrés roses), VA2 (ronds verts) et VB (triangles oranges).	148
3.28	Variation de la densité de l'eau avec la température, calculée à partir de nos simulations de dynamique moléculaire (courbe noire) et comparée à celle calculée d'après la littérature (courbe bleue) [Watanabe 91].	149
3.29	Schéma des couches d'hydratation, les couleurs des n couches correspondent à celles données dans le tableau 3.11.	150
3.30	Etude de la répartition de l'eau avec la distribution de probabilité de présence des molécules d'eau pour P22/L25 à gauche et VA3 à droite. Le calcul étant réalisé dans l'ensemble NPT, les couches ne seront pas les mêmes suivant le système étudié (cf Tab. 3.11).	151
3.31	Somme des distributions de probabilité de présence des molécules d'eau pour P22/L25 à gauche et VA3 à droite, à 150K (courbe noire), 220K (courbe bleue) et 300K (courbe rouge).	152

3.32	Comparaison de l'écart quadratique moyen de la crambine P22/L25 hydratée dans la boîte n°1, $h = 4$ (courbe rouge), dans la nouvelle boîte d'eau supérieure, $h = 7.5$ (courbe verte) et dans le cristal, $h = 0.9$ (courbe bleue) sur une trajectoire de 1ns.	153
3.33	Comparaison de l'écart quadratique moyen des deux protéines contenues dans le cristal de la boîte n°3, sur une trajectoire de 1ns.	154
3.34	Variation de l'écart quadratique moyen de la crambine P22/L25 hydratée dans la boîte d'eau de taille supérieure à gauche ($h = 7.5$) et dans le cristal à droite ($h = 0.9$), en fonction du temps de simulation : pour une trajectoire de 50ps (courbe noire), de 100ps (courbe rouge), de 200ps (courbe bleue) et de 1ns (courbe verte).	155
3.35	Mobilité des atomes pour la crambine P22/L25 simulée dans la boîte d'eau n°1 à gauche ($h = 4$) et dans la boîte d'eau n°2 à droite ($h = 7.5$), à 300K et sur une trajectoire de 1ns.	156
3.36	Mobilité des atomes pour la crambine P22/L25 simulée dans la boîte d'eau n°1 à gauche ($h = 4$) et dans la boîte d'eau n°3 à droite ($h = 0.9$), à 300K et sur une trajectoire de 1ns.	156
3.37	Comparaison des variations des volumes des boîtes d'eau avec l'augmentation de la température pour la première boîte, $h = 4$ (courbe bleue) et la seconde, $h = 7.5$ (courbe noire).	157
3.38	Somme des distributions de probabilité de présence des molécules d'eau pour P22/L25 dans la boîte n°1 ($h = 4$) à gauche et dans la boîte n°2 ($h = 7.5$) à droite, à 150K (courbe noire), 220K (courbe bleue) et 300K (courbe rouge).	158
3.39	Représentation schématique du paysage d'énergie libre [Ishikawa <i>et al.</i> 08].	162
3.40	Variation du MSD des gammas (courbe noire) comparée à celle des fluctuations des gammas (courbe bleue) pour la crambine P22/L25 hydratée à 300K, sur une trajectoire de 1ns.	164
3.41	Lois en puissance de γ_{11} , γ_{17} et γ_{37} de la crambine P22/L25 hydratée à 300K.	166
3.42	Représentation des 3 types de PMF possibles avec 3 angles gamma de la crambine P22/L25 hydratée à 300K sur une trajectoire de 5ns.	167
3.43	Variation du PMF de la crambine P22/L25 hydratée sur une trajectoire de 5ns à 300K. L'exposant α est donné pour chaque gamma, les hélices alpha sont localisées en jaune.	168
3.44	Séquence primaire des exposants de la crambine P22/L25 hydratée à 300K, sur une trajectoire de 5ns. Les labels rouges représentent les acides aminés de la séquence.	169
3.45	Séquences primaires des exposants des crambines hydratées à 300K sur une trajectoire de 1ns. Les labels rouges représentent les acides aminés de la séquence, la ligne noire en traits interrompus indique la limite pour les exposants $\alpha > 1$ et la structure secondaire est localisée.	170
3.46	Comparaisons des PMF de la crambine P22/L25 en haut et de la crambine P22/I25 en bas, hydratées sur une trajectoire de 1ns à 300K. L'exposant α est donné pour chaque gamma, les hélices alpha sont localisées en jaune.	171
3.47	Comparaisons des PMF de la crambine S22/I25 en haut et de la crambine S22/L25 en bas, hydratées sur une trajectoire de 1ns à 300K. L'exposant α est donné pour chaque gamma, les hélices alpha sont localisées en jaune.	172

3.48	Variation du PMF de la crambine P22/L25 hydratée sur une trajectoire de 1ns en fonction de la température : la première cartographie correspond à 150K, la seconde à 220K et la dernière à 300K. L'exposant α est donné pour chaque gamma, les hélices alpha sont localisées en jaune.	174
3.49	Variation du PMF de la viscotoxine B hydratée sur une trajectoire de 1ns en fonction de la température : la première cartographie correspond à 150K, la seconde à 220K et la dernière à 300K. L'exposant α est donné pour chaque gamma, les hélices alpha sont localisées en jaune.	175
3.50	Représentation des PMF de la crambine P22/L25 à 150K (en bleu), 220K (en rouge) et 300K (en noir) pour deux gammas harmoniques (en haut, n°10 et 11) et deux gammas anharmoniques (en bas, n°37 et 38).	177
3.51	Représentation des PMF de la viscotoxine B à 150K (en bleu), 220K (en rouge) et 300K (en noir) pour deux gammas harmoniques (en haut, n°28 et 29) et deux gammas anharmoniques (en bas, n°20 et 21).	178
3.52	Variation de l'écart quadratique moyen des gammas 10 et 11, équivalent aux gammas harmoniques et des gammas 37 et 38, pour les gammas anharmoniques, de la crambine P22/L25 hydratée sur une trajectoire de 1ns. Les lignes rouges en traits interrompus correspondent aux régressions linéaires de la courbe dans l'état R.	180
3.53	Variation de l'écart quadratique moyen des gammas 28 et 29, équivalent aux gammas harmoniques et des gammas 20 et 21, pour les gammas anharmoniques, de la viscotoxine B hydratée sur une trajectoire de 1ns. Les lignes rouges en traits interrompus correspondent aux régressions linéaires de la courbe dans l'état R.	181
3.54	Variation du PMF de la crambine P22/L25 à 150K sur une trajectoire de 1ns : la première cartographie correspond à la protéine dans le vide et la seconde à la protéine hydratée. L'exposant α est donné pour chaque gamma, les hélices alpha sont localisées en jaune.	183
3.55	Variation du PMF de la crambine P22/L25 à 220K sur une trajectoire de 1ns : la première cartographie correspond à la protéine dans le vide et la seconde à la protéine hydratée. L'exposant α est donné pour chaque gamma, les hélices alpha sont localisées en jaune.	184
3.56	Variation du PMF de la crambine P22/L25 à 300K sur une trajectoire de 1ns : la première cartographie correspond à la protéine dans le vide et la seconde à la protéine hydratée. L'exposant α est donné pour chaque gamma, les hélices alpha sont localisées en jaune.	185
3.57	Evolution du PMF entre 150K (courbe bleue en traits interrompus) et 300K (courbe rouge), pour deux gammas de la crambine P22/L25 hydratée sur une trajectoire de 5ns : g4, gamma harmonique, à gauche et g44, gamma anharmonique, à droite.	186
3.58	Barrières d'énergie pour deux méthyles de la structure optimisée de la crambine P22/L25 dans le vide : méthyle du résidu THR2 en noir et méthyle du résidu VAL8 en rouge.	187
4.1	Evolution de l'écart quadratique moyen des ions (courbe noire) comparée à celui des viscotoxines hydratées (courbe bleue), sur la moyenne des 5 trajectoires de 1ns. N correspond au nombre d'ions.	191

4.2	Evolution de la moyenne des écarts quadratiques moyens des ions mis en solution avec VA3 (carrés roses), VA2 (ronds bleu-verts) et VB (triangles oranges).	192
4.3	Dépendance en temps de simulations de l'écart quadratique moyen des ions simulés avec la viscotoxine A3, la courbe noire représente une trajectoire de 50ps, la courbe rouge de 100ps, la courbe bleue de 200ps et la courbe verte de 1ns.	192
1.1	Localisation des mutations entre VA3 et VB [Giudici <i>et al.</i> 03].	197
1.2	Modèle d'insertion probable des viscotoxines dans les membranes et localisation du résidu 25 lors de l'interaction [Coulon <i>et al.</i> 03].	198
1.3	Illustration des mécanismes d'insertion des thionines dans une membrane, le modèle <i>carpet</i> à gauche et le modèle <i>barrel-stave</i> à droite [Shai 99].	200
2.1	Comparaisons des écarts quadratiques moyens des viscotoxines (courbes rouges) et des viscotoxines iodinisées (courbes noires), simulées dans une boîte d'eau sur des trajectoires de 1ns.	203
2.2	Comparaisons de la mobilité des atomes entre les viscotoxines à gauche, et les viscotoxines iodinisées à droite à 300K, sur une trajectoire de 1ns.	205
2.3	Mobilité des atomes de la tyrosine 13 des viscotoxines à gauche et des viscotoxines iodinisées à droite, simulées dans une boîte d'eau à 300K et sur une trajectoire de 1ns.	206
2.4	Modèles utilisés lors des calculs de la densité électronique.	208
2.5	Variation des charges atomiques pour le modèle n°1 à gauche et le modèle n°2 à droite, en noir. Les calculs avec l'iodination de la TYR13 sont représentés en rouge.	209
D.1	Ecart quadratique moyen de la protéine globale (courbe noire) comparé aux différents types d'atomes : pour tous les hydrogènes (courbe rouge) et tous les atomes qui ne sont pas des hydrogènes (courbe bleue) dans le vide, sur une trajectoire de 1ns.	249
E.1	Variation de l'écart quadratique moyen selon le temps de production : 50ps (courbe noire), 100ps (courbe rouge), 200ps (courbe bleue) et 1ns (courbe verte) pour les thionines dans le vide.	253
F.1	Evolution de la dispersion des fluctuations (courbe noire) comparée à la variation du $\frac{MSD}{NT}$ des protéines (courbe bleue) en fonction de la température, dans le vide sur une trajectoire de 1ns. Les lignes rouges en traits interrompus correspondent aux régressions linéaires des états R et F.	257
G.1	Variation de la mobilité des atomes de la crambine S22/I25 en fonction de la température dans le vide, sur une trajectoire de 1ns.	261
G.2	Variation de la mobilité des atomes de la crambine P22/I25 en fonction de la température dans le vide, sur une trajectoire de 1ns.	262
G.3	Variation de la mobilité des atomes de la crambine S22/L25 en fonction de la température dans le vide, sur une trajectoire de 1ns.	263
G.4	Variation de la mobilité des atomes de la viscotoxine A3 en fonction de la température dans le vide, sur une trajectoire de 1ns.	264

G.5	Variation de la mobilité des atomes de la viscotoxine A2 en fonction de la température dans le vide, sur une trajectoire de 1ns.	265
G.6	Variation de la mobilité des atomes de la viscotoxine B en fonction de la température dans le vide, sur une trajectoire de 1ns.	266
H.1	Variation de la mobilité des atomes de la crambine S22/I25 en fonction du temps de production, à 300K.	269
H.2	Variation de la mobilité des atomes de la crambine P22/I25 en fonction du temps de production, à 300K.	270
H.3	Variation de la mobilité des atomes de la crambine S22/L25 en fonction du temps de production, à 300K.	271
H.4	Variation de la mobilité des atomes de la viscotoxine A3 en fonction du temps de production, à 300K.	272
H.5	Variation de la mobilité des atomes de la viscotoxine A2 en fonction du temps de production, à 300K.	273
H.6	Variation de la mobilité des atomes de la viscotoxine B en fonction du temps de production, à 300K.	274
I.1	Variation du nombre de résidus suivant leur valeur du $\frac{MSD}{T}$ pour les crambines S22/I25, P22/I25, S22/L25 et les 3 viscotoxines déshydratées, sur une trajectoire de 1ns. Les histogrammes correspondent aux températures suivantes : noir = 5K, rouge = 30K, bleu = 70K, bleu-vert = 100K, fushia = 130K, kaki = 150K, bleu marine = 170K, marron = 200K, rose = 220K, vert = 250K, gris = 280K, orange = 300K.	277
I.2	Variation du MSD de chaque acide aminé avec la température pour les autres thionines dans le vide, sur une trajectoire de 1ns.	278
J.1	Ajustement des paramètres du modèle de Doster pour les protéines hydratées. La courbe noire est l'écart quadratique moyen calculé par nos simulations de dynamique moléculaire et la courbe rouge est le modèle de Doster.	281
J.2	Second ajustement des paramètres du modèle de Doster pour la viscotoxine VA3 hydratée. La courbe noire est l'écart quadratique moyen calculé par nos simulations de dynamique moléculaire et la courbe rouge est le modèle de Doster.	282
K.1	Distribution des barrières d'énergie des méthyles pour la crambine P22/L25 à différentes températures.	285
L.1	Variation de l'écart quadratique moyen des thionines hydratées sur une trajectoire de 1ns. Les carrés noirs correspondent à la 1ère condition initiale, les ronds rouges à la 2nde, les triangles bleus à la 3ième, les losanges verts à la 4ième et les triangles roses à la 5ième.	289
M.1	Ecart quadratique moyen des protéines hydratées (courbe bleue) comparé aux différents types d'atomes : pour tous les atomes d'hydrogènes (courbe noire) et tous les atomes qui ne sont pas des hydrogènes (courbe rose), sur une trajectoire de 1ns.	293

N.1	Variation de l'écart quadratique moyen selon le temps de production : 50ps (courbe noire), 100ps (courbe rouge), 200ps (courbe bleue) et 1ns (courbe verte) pour les thionines hydratées.	297
O.1	Evolution de la dispersion des fluctuations (courbe noire) comparée à la variation du $\frac{MSD}{NT}$ des protéines hydratées (courbe bleue) en fonction de la température, sur une trajectoire de 1ns. Les lignes rouges en traits interrompus correspondent aux régressions linéaires des états R et F.	301
P.1	Variation de la mobilité des atomes de la crambine S22/I25 hydratée en fonction de la température, sur une trajectoire de 1ns.	305
P.2	Variation de la mobilité des atomes de la crambine P22/I25 hydratée en fonction de la température, sur une trajectoire de 1ns.	306
P.3	Variation de la mobilité des atomes de la crambine S22/L25 hydratée en fonction de la température, sur une trajectoire de 1ns.	307
P.4	Variation de la mobilité des atomes de la viscotoxine A3 hydratée en fonction de la température, sur une trajectoire de 1ns.	308
P.5	Variation de la mobilité des atomes de la viscotoxine A2 hydratée en fonction de la température, sur une trajectoire de 1ns.	309
P.6	Variation de la mobilité des atomes de la viscotoxine B hydratée en fonction de la température, sur une trajectoire de 1ns.	310
Q.1	Variation de la mobilité des atomes de la crambine S22/I25 en fonction du temps de production, à 300K.	313
Q.2	Variation de la mobilité des atomes de la crambine P22/I25 en fonction du temps de production, à 300K.	314
Q.3	Variation de la mobilité des atomes de la crambine S22/L25 en fonction du temps de production, à 300K.	315
Q.4	Variation de la mobilité des atomes de la viscotoxine A3 en fonction du temps de production, à 300K.	316
Q.5	Variation de la mobilité des atomes de la viscotoxine A2 en fonction du temps de production, à 300K.	317
Q.6	Variation de la mobilité des atomes de la viscotoxine B en fonction du temps de production, à 300K.	318
R.1	Variation du nombre de résidus suivant leur valeur du $\frac{MSD}{T}$ pour les crambines S22/I25, P22/I25, S22/L25 et les 3 viscotoxines hydratées, sur une trajectoire de 1ns. Les histogrammes correspondent aux températures suivantes : noir = 5K, rouge = 30K, bleu = 70K, bleu-vert = 100K, fushia = 130K, kaki = 150K, bleu marine = 170K, marron = 200K, rose = 220K, vert = 250K, gris = 280K, orange = 300K.	321
R.2	Variation du MSD de chaque acide aminé avec la température pour les autres thionines hydratées, sur une trajectoire de 1ns.	322
S.1	Somme des distributions de probabilité de présence des molécules d'eau pour les protéines hydratées, à 150K (courbe noire), 220K (courbe bleue) et 300K (courbe rouge).	325

Introduction - Objectifs

L'élucidation de la structure et de la fonction biologique des protéines est un enjeu majeur pour ses implications dans la recherche médicale et en biotechnologies. Les relations structure-fonction sont par ailleurs liées à la dynamique moléculaire : la structure d'une protéine n'est pas rigide, ses fluctuations et sa flexibilité sont des propriétés physiques importantes pour la compréhension des interactions entre biomolécules. Leur simulation et analyse sont au cœur de notre travail. L'objectif de notre thèse est la compréhension des relations entre structure, dynamique et flexibilité de protéines appartenant à la famille des thionines (crambines et viscotoxines). Ces protéines extraites des plantes ont des applications potentielles dans le traitement du cancer et comme antibiotiques [Romagnoli *et al.* 00]. Au-delà de leur intérêt biologique spécifique, leur comportement pose la question fondamentale de l'influence d'une perturbation locale de la structure (mutation) sur un comportement global (fonction). Les thionines possèdent en effet une haute homologie de séquence mais une toxicité très différente et la relation entre cette toxicité et la séquence de leurs acides aminés est loin d'être élucidée [Han *et al.* 96] : Pourquoi une seule mutation qui ne perturbe pas la structure tertiaire est-elle capable de modifier aussi fortement la toxicité de ces molécules vis-à-vis des membranes cellulaires ? Comment cet effet est-il relié aux mécanismes d'interactions protéines/membranes ? La réponse à ces deux premières questions est multifactorielle. Elle nécessite un modèle biochimique réaliste de la membrane cellulaire. Un tel modèle n'est actuellement pas disponible à l'échelle microscopique. Des modèles simplifiés des interactions protéines-membranes ont été développés ces dernières années [Giudici *et al.* 03, Coulon *et al.* 02] mais leur pertinence biologique pour les interactions thionines/membrane est loin d'être évidente. La construction théorique de modèles membranaires est par ailleurs une tâche ardue. On peut se demander dès lors si la question la plus pertinente n'est pas d'abord de tenter de comprendre comment les propriétés physiques et chimiques intrinsèques des thionines sont modifiées par une mutation ? En nous servant de la famille des thionines comme modèle, nous répondrons ainsi à des questions de portée générale : comment évolue la dynamique et la flexibilité d'une protéine modifiée chimiquement ou par mutation, comment se propage la modification locale des propriétés mécaniques et électroniques induites par une mutation ? Quel est le rôle de l'hydratation dans les relations séquence-structure-dynamique-fonction de protéines globulaires ?

Pour répondre à ces questions, nous avons utilisé des modèles physiques des interactions entre atomes (champ de force) et des modèles de la physique statistique afin de simuler les effets des mutations sur la dynamique de protéines isolées et hydratées. Nous tentons

ainsi de comprendre l'émergence de la fonction biologique de la protéine et sa relation à sa dynamique à partir d'une approche microscopique.

Il est acquis que la fonction biologique des protéines s'exprime au-delà de la température de transition dynamique où l'anharmonicité des fluctuations structurales (fluctuations de la position des atomes autour de leur valeur moyenne) entraîne un accroissement de flexibilité de la molécule. L'existence d'une telle transition de phase nous offre en principe un moyen d'étudier les perturbations de la fonction d'une protéine, due à une mutation, en analysant les modifications possibles de sa transition dynamique. Nous avons entrepris pour la première fois une telle simulation pour sept thionines (4 non toxiques : les crambines et 3 toxiques : les viscotoxines) en utilisant la dynamique moléculaire. Nous démontrons l'existence d'une transition pour cette famille de protéines. De plus, nous montrons que les mutations perturbent fortement la dynamique au-delà de la température de transition où s'exprime la fonction biologique. Une analyse détaillée des mouvements démontre un comportement complexe non-linéaire : une mutation modifie l'ensemble de la dynamique et ne se limite pas à l'environnement immédiat de la perturbation locale induite par le changement d'un acide aminé par un autre.

Les trois protéines toxiques étudiées, les viscotoxines, sont chargées positivement, l'électroneutralité locale du système est assurée par des contre-ions négatifs en solution. Chaque ion peut être considéré comme une sonde de la dynamique des protéines. Nos travaux théoriques suggèrent, en effet, que l'étude expérimentale de la dynamique de ces ions fournirait des informations pertinentes sur la transition dynamique et la mobilité.

A travers l'étude des protéines toxiques, les viscotoxines, nous avons voulu étudier tout d'abord l'effet des mutations et perturbations chimiques locales sur les propriétés électroniques globales (réactivité) des protéines et leur relation avec la dynamique. L'expérience montre en effet que le remplacement d'un seul atome d'hydrogène par un atome d'iode suffit à supprimer totalement la toxicité des viscotoxines (expérience réalisée par l'iodination du Tyr13). Ce changement induit-il une modification importante du réseau de liaisons hydrogène et de la dynamique et par delà de la fonction ? Observe-t-on par ailleurs une modification des propriétés électroniques/électrostatiques ? Pour répondre à ces questions, nous avons tout d'abord réalisé des simulations de dynamique moléculaire pour étudier le comportement dynamique de ces viscotoxines iodinisées puis nous avons appliqué la théorie de la fonctionnelle de la densité à l'étude de la réactivité des protéines modélisées par la structure fluctuante simulée précédemment par dynamique moléculaire. Les calculs quantiques ont été réalisés en collaboration avec l'Université d'Anvers avec qui le groupe de recherche auquel j'appartiens a développé des outils d'analyse de la densité électronique de macromolécules [Krishtal *et al.* 06, Senet et Aparicio 07].

Notre projet est pluridisciplinaire : l'effet des mutations sur le comportement d'une protéine nécessite :

- d'explorer l'espace des conformations en fonction de la température et d'analyser les fluctuations (physique, mathématique),
- de modéliser la réactivité chimique à partir des propriétés électroniques (chimie),
- d'étudier les implications biologiques (biologie moléculaire).

Nos travaux sur les relations entre structure-dynamique-réactivité, l'influence d'une mutation sur le comportement global, pourraient avoir un intérêt pour l'étude d'autres familles de protéines globulaires.

PREMIÈRE PARTIE

Méthodes et systèmes étudiés

Techniques de simulations

1.1 Propriétés macroscopiques - physique statistique

1.1.1 Définitions d'un état microscopique et des ensembles statistiques

En physique classique, un état microscopique (appelé aussi micro-état) d'un système est défini par la position $\vec{r}(t)$ et la quantité de mouvement $\vec{p}(t)$ de ses constituants à un instant t donné. \vec{r} et \vec{p} représentent des vecteurs dont les composantes sont $r_{i\alpha}$ et $p_{i\alpha}$ pour la particule i des systèmes dans la direction cartésienne α . Le système possède une énergie potentielle $V(\vec{r})$ et cinétique $T(\vec{p})$ dont la somme forme l'hamiltonien $H(\vec{r}, \vec{p})$. Dans un système isolé, H ne dépend pas explicitement du temps et la valeur de $H(\vec{r}, \vec{p})$ est l'énergie du micro-état visité au temps t . C'est aussi l'énergie totale E du système fixée à δE près. L'évolution temporelle du système consiste au passage d'un micro-état à un autre de même énergie.

Un système macroscopique isolé a un nombre gigantesque d'états (de l'ordre du nombre d'Avogadro). Sa description microscopique est statistique et fait appel à la notion d'ensemble statistique. Maxwell, Boltzmann et Gibbs furent les premiers à introduire ce concept d'ensemble statistique. Un système est représenté par un ensemble, soit une infinité de répliqués de lui-même dans les mêmes conditions macroscopiques. On peut rendre concrète cette notion en imaginant un grand nombre de systèmes identiques dans des conditions initiales différentes. A chaque instant de leur évolution, la probabilité d'un micro-état peut être calculée en examinant les états de l'ensemble. A l'équilibre, cette probabilité n'évolue pas avec le temps. Pour un système isolé à l'équilibre, nous pouvons postuler que tous les micro-états discernables et accessibles d'un système isolé sont équiprobables (premier postulat de la mécanique statistique énoncé par Gibbs).

Lorsque le système est fermé en contact avec un thermostat à la température T , les probabilités dépendent de l'énergie du micro-état :

$$p_I = \frac{e^{-\frac{E_I}{kT}}}{\sum_I e^{-\frac{E_I}{kT}}} = \frac{1}{Z} e^{-\frac{E_I}{kT}}, \quad (1.1)$$

où la somme (Z) dans l'Eq. 1.1 est sur tous les états possibles et $E_I(\vec{r}, \vec{p})$ un micro-état.

Cette loi de Boltzmann (Eq. 1.1) s'applique aussi bien à un système macroscopique qu'à une molécule à température constante.

Il existe 4 ensembles statistiques très utilisés en fonction des paramètres macroscopiques (N, V, E, T, P, μ) qui sont fixés :

1. L'ensemble NVE (microcanonique),
2. L'ensemble NVT (canonique),
3. L'ensemble NPT (isobare/isotherme),
4. L'ensemble μVT (grand canonique).

Les notations N, V, E, T, P, μ signifient respectivement, le nombre de particules N , le volume V , l'énergie totale E , la température T , la pression P ou le potentiel chimique μ .

Les ensembles NVT et NPT sont utilisés dans notre travail.

1.1.2 Ergodicité et propriétés statistiques

Dans les différents ensembles statistiques que nous avons présentés précédemment, les moyennes ont un sens statistique et elles sont définies sur l'ensemble des états accessibles au système. Ceci ne correspond pas à la moyenne que l'on effectue dans une expérience ou dans une simulation de dynamique moléculaire. En effet dans ces cas là, on réalise une série de mesures ou de simulations durant des intervalles de temps finis ou sur un nombre de configurations fini puis on détermine la moyenne sur ces mesures ou simulations.

L'hypothèse ergodique stipule alors que les quantités calculées sur une moyenne d'ensemble seront équivalentes aux quantités moyennées sur le temps.

Si X est une propriété du système $X(\vec{r}(t), \vec{p}(t))$ et $\Gamma(\vec{r}(t), \vec{p}(t))_\tau$ représente la trajectoire pendant un temps τ [Diu *et al.* 89], on suppose :

$$\langle X \rangle_\Gamma = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau dt X(\vec{r}(t), \vec{p}(t))_\Gamma, \quad (1.2)$$

$$\langle X \rangle_{ensemble} = \sum_I p_I X_I, \quad (1.3)$$

où $\langle \dots \rangle_\Gamma$ représente une moyenne temporelle sur les états visités au cours de la trajectoire Γ alors que $\langle \dots \rangle_{ensemble}$ représente une moyenne sur les états accessibles au système. p_I est la probabilité du micro-état I et X_I la valeur de X dans cet état $X(\vec{r}_I, \vec{p}_I)$, où \vec{r}_I et \vec{p}_I sont les vecteurs représentant les positions et quantités de mouvement dans l'état I .

L'hypothèse ergodique suppose une trajectoire de longueur infinie. Plus précisément, elle suppose que le système a eu le temps de visiter tous les états possibles. Bien entendu, dans nos simulations numériques, rien ne nous le garantit et une trajectoire est toujours de longueur finie. En pratique, les moyennes temporelles sont souvent réalisées sur plusieurs trajectoires plus courtes [Caves *et al.* 98].

$$\langle X \rangle_{temps} = \frac{\sum_{\Gamma_i} \frac{1}{\tau_i} \int_0^{\tau_i} dt X(\vec{r}(t), \vec{p}(t))_{\Gamma_i}}{\sum_{\Gamma_i}}. \quad (1.4)$$

Les trajectoires Γ seront calculées par dynamique moléculaire (cf chapitre 1.3, partie I) à T ou P constantes.

La plupart des propriétés que nous mesurerons par dynamique moléculaire seront des propriétés structurales qui dépendent des configurations moléculaires $(\vec{r}(t))$, c'est à dire qu'elles reflètent les propriétés de l'énergie potentielle $V(\vec{r})$. Si $X(\vec{r}(t))$ est une propriété structurale, la probabilité de trouver X entre X et $X + dX$ est donnée par $p_X dX$ où p_X est la distribution de probabilité de X . Si on suppose p_X gaussien, elle sera entièrement caractérisée par son premier et second moments, soit $\langle X \rangle$ et σ_X^2 . Dans notre travail, nous utiliserons ces deux mesures statistiques :

1. La valeur moyenne temporelle d'une propriété :

$$\langle X \rangle_\Gamma = \frac{1}{n} \sum_{I=1}^n X_I, \quad (1.5)$$

où n représente le nombre de valeurs de X calculées le long de la trajectoire Γ et X_I sa valeur à l'étape I .

2. Les fluctuations structurales :

$$\begin{aligned}
 \sigma^2 &= \langle \delta X^2 \rangle_{\Gamma} \\
 &= \frac{1}{n} \sum_{I=1}^n \delta X_I^2 \\
 &= \frac{1}{n} \sum_{I=1}^n (X_I - \langle X \rangle_{\Gamma})^2 \\
 &= \frac{1}{n} \left(\sum_{I=1}^n X_I^2 \right) - (\langle X \rangle_{\Gamma})^2.
 \end{aligned} \tag{1.6}$$

La variance des fluctuations structurales, ou MSD (Mean Square Displacement), sera souvent utilisée dans toute notre étude afin d'étudier les fluctuations des atomes des thionines avec l'élévation de la température.

Les structures moyennes calculées peuvent être comparées aux données RMN ou de diffraction X. Par ailleurs, les fluctuations structurales sont mesurées via le facteur de Debye-Waller (noté B), établi par Peter Debye et Ivar Waller [Debye 14, Waller 23]. Il décrit l'atténuation de l'intensité de diffraction X ou de la diffusion de neutrons due à des mouvements thermiques et au désordre. Il est également appelé B-facteur ou facteur de température.

Ce facteur peut être comparé approximativement à nos calculs du MSD. En effet, si on néglige le désordre statique, nous avons la relation suivante :

$$B = \frac{8\pi^2}{3} \times MSD. \tag{1.7}$$

Dans le tableau suivant, nous résumons les informations (sur les moyennes, fluctuations et rapports à l'expérience) pour différentes propriétés physiques :

Propriété Physique	$\langle \Gamma \rangle$	Fluctuations (NVT)	Expérience
E	Enthalpie	C_v (Capacité calorifique)	Calorimétrie
\vec{r}	Structure moyenne	Flexibilité	Diffraction X, Neutrons, RMN (facteur de Debye-Waller B)

TAB. 1.1 – *Quantités physiques calculées dans notre travail.*

1.2 Surface d'énergie potentielle

1.2.1 Approximation de Born-Oppenheimer

La résolution de l'équation de Schrödinger pour une molécule complexe avec une précision raisonnable est impossible sans faire l'approximation de Born et Oppenheimer établie en 1927 [Born et Oppenheimer 27].

Dans cette approximation, la masse du noyau atomique étant largement plus grande que la masse de l'électron, l'énergie cinétique des noyaux est d'abord négligée et l'équation de

Schrödinger est résolue pour les électrons. La fonction d'onde des électrons peut être considérée comme découplée de celle des noyaux. L'énergie des électrons pour une configuration moléculaire définit ensuite une énergie potentielle effective pour les atomes. Cette énergie potentielle est une donnée essentielle et doit être modélisée pour une protéine.

1.2.2 Définitions des différents termes d'énergie

Le modèle pour l'énergie potentielle de la protéine est appelé "champ de force". Un champ de force est une collection d'atomes types, de paramètres et d'équations qui calculent l'énergie d'une molécule. En mécanique moléculaire, l'énergie potentielle d'une molécule proche de sa configuration d'équilibre est une simple fonction des distances, liaisons et angles qui relient les atomes d'une molécule.

Basée sur la loi de Hooke, la mécanique moléculaire représente un atome comme une sphère et les liaisons comme des ressorts qui peuvent être tirés, comprimés, repliés et tordus avec un angle de torsion; les interactions non-liantes entre les atomes sont aussi prises en considération. L'énergie potentielle totale E_{tot} d'une molécule est divisée en 5 parties qui sont attribuées aux liaisons ("stretch"), aux angles ("bond"), aux torsions (les angles diédraux) et aux interactions non-liantes (forces de Van der Waals et électrostatiques), comme dans l'équation 1.8.

$$E_{tot} = E_{str} + E_{bond} + E_{tors} + E_{vdw} + E_{elec}. \quad (1.8)$$

Cette équation définit la surface d'énergie potentielle de la molécule, ou encore le paysage d'énergie potentielle représentée par de multiples cols et vallées (appelé aussi "energy landscape"). L'additivité des termes de l'Eq. 1.8 est justifiée par les calculs quantiques sur de petites molécules.

1.2.2.1 Energies des liaisons covalentes

Energie d'élongation La liaison chimique entre deux atomes est décrite comme un puits de potentiel harmonique. Cette approximation n'est valable qu'au voisinage de l'équilibre. La liaison ne peut se briser et elle est matérialisée par un ressort : un ressort comprimé ou dilaté emmagasine en effet de l'énergie potentielle élastique. Cette énergie potentielle emmagasinée par le ressort est égale au travail effectué par la force qui a permis de le comprimer ou de le dilater. L'énergie d'élongation est :

$$E_{str} = \frac{1}{2} \sum_{l=1}^n k_l (R_l - R_{0l})^2, \quad (1.9)$$

où l représente la liaison étudiée, n le nombre de liaisons, k_l est la constante de force associée à cette liaison, R_{0l} la longueur d'équilibre de cette liaison et R_l la longueur de la liaison à l'instant t .

Mais cette équation n'est pas toujours assez précise pour représenter les liaisons au voisinage de l'équilibre. C'est pourquoi Allinger développa dans les années 90 un potentiel anharmonique représenté par un polynôme d'ordre 4 afin d'augmenter la précision, [Allinger *et al.* 89, Lii et Allinger 89a, Lii et Allinger 89b] :

$$E_{str} = \frac{1}{2} \sum_{l=1}^n \left[k_l^2 (R_l - R_{0l})^2 + k_l'^3 (R_l - R_{0l})^3 + k_l''^4 (R_l - R_{0l})^4 \right]. \quad (1.10)$$

Energie de flexion L'énergie de flexion peut être écrite avec un simple potentiel harmonique, comme dans l'équation 1.11 :

$$E_{bond} = \frac{1}{2} \sum_{i=1}^m K_i (\theta - \theta_0)^2, \quad (1.11)$$

où m est le nombre d'angles θ , K_i est la constante de force de l'angle étudié, θ_0 l'angle de référence et θ l'angle à l'instant t .

De la même façon, un polynôme d'ordre supérieur peut être utilisé pour décrire avec plus de précision l'énergie de déformation des angles :

$$E_{bond} = \frac{1}{2} \sum_{i=1}^m \left[K_i (\theta - \theta_0)^2 + K_i' (\theta - \theta_0)^3 + K_i'' (\theta - \theta_0)^4 \right]. \quad (1.12)$$

Energie de torsion La torsion d'une liaison définit un angle diédral τ autour de la liaison centrale entre les atomes B et C (en considérant un enchaînement d'atomes A-B-C-D). L'expression de l'énergie de torsion est donnée dans l'équation 1.13 :

$$E_{tors} = \sum_{i=1}^p V_{tors}. \quad (1.13)$$

Chaque potentiel de torsion V_{tors} a la forme suivante :

$$V_{tors} = \sum_{n=0}^p A_n [1 + \cos(n\tau - s)], \quad (1.14)$$

où A_n est une constante qui détermine la barrière de rotation autour de la liaison B-C, s représente la phase et le nombre de termes p la multiplicité.

Ce potentiel est anharmonique et représenté par plusieurs minima.

1.2.2.2 Energies des liaisons non-covalentes

Energie de Van der Walls Elle est représentée par une somme de potentiels à deux corps i et j (cf Eq. 1.15). Chaque potentiel de paire V est la somme de deux termes (cf Eq. 1.16). Le premier correspond à des forces de dispersion, le second représente les forces répulsives.

Le potentiel total est souvent représenté par un terme de Lennard-Jones (LJ) 12 – 6, où 12 – 6 représentent les puissances des deux termes (mais il peut également être un terme LJ 9 – 6 ou LJ 10 – 6 suivant le champ de force utilisé).

$$E_{vdw} = \frac{1}{2} \sum_i \sum_j V(r_{ij}), \quad (1.15)$$

$$V = 2\varepsilon \left[\left(\frac{r_0}{r} \right)^{12} - \left(\frac{r_0}{r} \right)^6 \right], \quad (1.16)$$

où r est la distance, r_0 est la distance pour laquelle l'énergie est nulle et ε est la profondeur du puits d'énergie minimale.

Energie électrostatique Le potentiel électrostatique entre deux charges q_i et q_j de deux atomes i et j est donnée par la loi de Coulomb et l'énergie coulombienne est :

$$E_{coulomb} = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{q_i q_j}{D r_{ij}}, \quad (1.17)$$

où n est le nombre de charge, D est une constante effective diélectrique et la distance r_{ij} sépare les deux charges q_i et q_j .

Termes croisés L'indépendance de ces différents termes énergétiques représente une des approximations utilisées en mécanique moléculaire. En réalité, l'énergie de la molécule ne dépend pas de manière totalement indépendante des longueurs de liaisons et des angles de ces liaisons. Le couplage entre ces degrés de liberté peut être décrit par de nouveaux termes croisés, qui augmentent ainsi la précision du champ de force. Normalement, ces termes sont représentés par des potentiels harmoniques. Les plus importants couplages sont :

- un terme élongation-élongation : ces interactions sont importantes dans les systèmes contenant des liaisons π . Le potentiel a la forme suivante :

$$V_{str-str} = \frac{1}{2} K_{bb'} [(r - r_0) (r' - r'_0)], \quad (1.18)$$

où $K_{bb'}$ la constante de force associée aux ressorts b et b' , r_0 sa longueur d'équilibre, r sa longueur à l'instant t et r'_0 la longueur d'équilibre du ressort b' et r' sa longueur à l'instant t .

- un terme élongation-flexion : l'utilisation de ce terme croisé est limitée à un angle et une liaison entre deux atomes constituant cet angle. Il a une influence importante sur la structure et les fréquences vibrationnelles du système. Sa forme est similaire à celle montrée dans l'équation 1.18.

- un terme élongation-torsion : ce terme croisé entre un angle diédral et sa liaison centrale est décrite par un des potentiels suivants :

$$V_{str-torsion} = \sum_{i=1}^n K_{b\tau} (r - r_0) \cos n\tau, \quad (1.19)$$

$$V_{str-torsion} = \sum_{i=1}^n K_{b\tau} (r - r_0) (1 + \cos n\tau), \quad (1.20)$$

où n le nombre de termes possibles, $K_{b\tau}$ la constante de force du ressort b situé sur la liaison centrale de l'angle de torsion τ , r_0 sa longueur d'équilibre et r sa longueur à l'instant t .

• un terme flexion-flexion-torsion : ces termes couplent un angle diédral A-B-C-D aux deux angles A-B-C et B-C-D.

$$V_{bond-bond-torsion} = \frac{1}{2} \sum_{i=1}^n k_{\theta\theta'\tau} (\theta - \theta_0) (\theta' - \theta'_0) \cos \tau, \quad (1.21)$$

où i représente le terme croisé étudié, n le nombre de termes possibles, $k_{\theta\theta'\tau}$ la constante de force de l'angle θ (avec θ_0 , son angle de référence), θ' (avec θ'_0 , son angle de référence) et l'angle diédral τ .

1.2.3 Les différents types de champ de force

Le développement des champs de force et de leurs paramètres a commencé dans les années 60 et a nécessité plusieurs étapes. Les champs de force dans lesquels a été décrite la surface d'énergie potentielle des hydrocarbures [Allinger *et al.* 89], des protéines [Hagler *et al.* 74b, Hagler *et al.* 74a], des acides nucléiques [Nilsson et Karplus 86], et aussi pour d'autres composants [Clark *et al.* 89], ont été introduits et appliqués à beaucoup d'études. Différents champs de force existent, ils diffèrent par trois aspects :

1. La définition de chaque terme énergétique,
2. Le nombre de termes croisés ajoutés,
3. Le type d'informations utilisés pour adapter les paramètres.

Les champs de force sont divisés en 2 classes : la classe I et la classe II.

1.2.3.1 Les champs de force classe I

Ce type de champ de force est utilisé pour traiter des systèmes larges, comme les protéines ou l'ADN. La forme des termes énergétiques est aussi simple que possible. Les premiers champs de force développés furent des classes I. Ils sont constitués de termes harmoniques qui décrivent les interactions des liaisons covalentes et sont paramétrés sur des observations expérimentales de spectroscopie, propriétés thermodynamiques et structures cristallines. Plus tard, des termes plus complexes furent ajoutés mais toujours paramétrés sur l'expérience. Se situent dans cette classe, les champs de force suivants :

- MM1 : premier champ de force développé par Allinger [Allinger et Sprague 73].
- CFF-3 : champ de force développé par Ermer et Lifson [Ermer et Lifson 73].
- EAS : "Engler-Androse-Schleyer", basé sur le champ de force MM1 avec des termes pour les interactions des liaisons non-covalentes [Engler *et al.* 73].
- ECEPP/2 : "Empirical Conformational Energy Program for Peptides", développé pour des simulations d'acides nucléiques et de protéines [Momany *et al.* 75].
- MM2 : Basé sur MM1, il contient des termes paramétrés pour les composés organiques [Allinger 77].
- MMP2 : Basé sur MM2, il repose sur des relations linéaires entre longueurs des liaisons et angles [Sprague *et al.* 87].
- QCFF/PI : une extension du champ de force CFF pour les systèmes π [Warshel 73].
- CVFF : "Consistent Valence Force field", il inclut des termes anharmoniques et croisés. Ce champ de force a été paramétré pour 20 acides aminés et quelques hydrocarbures [Dauber-Osguthorpe *et al.* 88].
- AMBER : "Assisted Model Building with Energy Refinement", ce champ de force est paramétré pour l'ADN et les protéines [Weiner *et al.* 84].

- OPLS : Il utilise le potentiel TIP3P pour l'eau et n'a pas de potentiel pour les liaisons hydrogène [Jorgensen et Tirado-Rives 88].
- MMX : une extension de MM2 et MMP1 pour les systèmes organiques et inorganiques [Gajewski *et al.* 90].
- CHARMM : "Chemistry at HARvard Molecular Mechanics", il est également paramétré pour les protéines, l'ADN mais aussi pour les lipides et d'autres petites molécules organiques [Brooks *et al.* 83].
- GROMOS : "GRONingen MOlecular Simulation", ce champ de force est paramétré pour les protéines, les nucléotides, les sucres et d'autres biomolécules [Hermans *et al.* 84].
- GROMACS : "GRONingen MACHine for Chemical Simulations", il est dérivé du champ de force GROMOS et simule les équations de mouvements de Newton pour des systèmes comprenant des centaines jusqu'à des millions de particules [Bekker *et al.* 93].
- DREIDING : un champ de force où les constantes de force et les paramètres de géométrie dépendent de l'hybridation au lieu de la constitution des atomes [Mayo *et al.* 90].

1.2.3.2 Les champs de force classe II

Les champs de force de classe II ont été développés pour simuler de petites molécules dans des environnements variés. Les calculs *ab-initio* sont utilisés pour déterminer tous les paramètres des champs de force. Cette précision est meilleure que pour ceux de la classe I. Ils sont la plupart du temps paramétrés sur de petites molécules.

Se situent dans la classe II les champs de force suivants :

- MM3 et MM4 : ces champs de force proviennent de MM2 avec des potentiels plus complexes. Ils sont paramétrés pour des composants organiques et pour reproduire leurs fréquences de vibration [Allinger *et al.* 89, Allinger *et al.* 96].
- CFF93-CFF94 : ce champ de force, similaire à MM4, utilise une méthodologie rigoureuse pour générer les paramètres des termes croisés [Hwang *et al.* 94].
- MMFF : "Merck Molecular Force Field" est semblable au MM3. Seuls les paramètres de Van der Waals ont utilisés des données expérimentales [Halgren 92].
- UFF : "Universal Force Field" représente tout le tableau périodique. Les paramètres sont basés sur l'hybridation atomique [Rappe *et al.* 92].

Dans le domaine des protéines, les meilleurs résultats sont obtenus avec les champs de force OPLS, CHARMM et AMBER. Dans notre travail, nous avons donc choisi d'étudier les thionines avec le champ de force CHARMM.

1.2.4 CHARMM

Nous avons choisi d'utiliser le champ de force CHARMM puisqu'il présente les avantages suivants :

- il permet l'utilisation de contraintes sur les liaisons et angles diédraux pour les atomes sélectionnés,
- il permet l'utilisation de contraintes harmoniques sur les groupes d'atomes sélectionnés (eau, protéine, backbone et/ou chaîne latérale),
- il permet de suivre dans le temps les modifications des angles, distances et angles diédraux,
- il permet des simulations dans le vide et avec des simulations implicites ou explicites,

- les méthodes d'hydratation explicite incluent les conditions périodiques,
- il existe plusieurs algorithmes d'optimisation : Steepest Descent, Conjugate Gradient, Powell et Newton-Raphson,
- il permet l'édition de rapports détaillés pour chaque simulation (sur les structures, contraintes et paramètres),
- les trajectoires peuvent être analysées et visualisées avec des logiciels tels que VMD [Humphrey *et al.* 96].

Une étude comparative de 4 champs de force (OPLS, CHARMM, AMBER et GROMOS) fut réalisée par Manuel Rueda en 2007. Ses analyses prouvèrent également que parmi tous les champs de force étudiés, CHARMM donnait de très bons résultats, conformes à la réalité [Rueda *et al.* 07]. Par exemple, CHARMM donne des résultats conformes aux expériences sur les calculs des B-facteurs et la mobilité des structures (études également réalisées dans nos simulations) alors que Gromos est toujours le plus éloigné.

1.3 Dynamique et mécanique moléculaire

La dynamique moléculaire (MD) consiste à simuler le comportement d'un système moléculaire, dépendant du temps en se basant sur les équations de mouvement de Newton et un modèle de la fonction d'énergie potentielle (champ de force). La dynamique moléculaire permet de prédire les propriétés structurales, dynamiques et thermodynamiques des systèmes. En effet, les équations de Newton peuvent être couplées à un thermostat (température constante) ou barostat (pression constante) afin de simuler la molécule dans un ensemble statistique approprié.

La dynamique moléculaire a été utilisée pour la première fois dans les années 50 par B. Alder [Alder et Wainwright 59]. En 1964, A. Rahman réalisa la première simulation utilisant un potentiel réel pour un liquide, l'Argon [Rahman 64]. La première simulation de MD pour un système réel a été faite par Rahman et Stillinger dans leur étude de l'eau liquide en 1974 [Stillinger et Rahman 74]. La première simulation d'une protéine fut réalisée par J. Andrew McCammon en 1977 avec l'étude de l'inhibiteur pancréatique de la trypsine de bovin [McCammon *et al.* 77]. La dynamique moléculaire est maintenant utilisée dans la science des matériaux [Nakano *et al.* 93], la chimie [Hautman *et al.* 89], la biologie [Brooks *et al.* 88] et la physique [Abraham 86].

La méthode utilisée pour calculer les forces d'interaction (ou le potentiel dont elles dérivent) caractérise une simulation. Par exemple, la dynamique moléculaire classique représente l'évolution gouvernée par la loi de Newton (hypothèse de la mécanique classique). La dynamique moléculaire quantique représente l'évolution temporelle gouvernée par l'équation de Schrödinger pour les noyaux, en général dans l'approximation de Born-Oppenheimer. Le potentiel effectif dans les deux approches est évalué empiriquement et/ou par des méthodes *ab initio*. Enfin, la dynamique moléculaire *ab initio* de type Car Parrinello [Marx et Hutter 00] consiste à calculer les forces entre atomes à partir de la résolution de l'équation de Schrödinger pour les électrons (approximation de Born-Oppenheimer) à chaque pas de temps : il n'y a pas de modélisation du potentiel.

Toutes les quantités observables peuvent être calculées numériquement, mais les temps de calculs et l'espace mémoire disponible des ordinateurs restent des facteurs limitants. Les

capacités de calcul étant limitées, le nombre de particules et la durée des trajectoires dans une simulation numérique le sont aussi. Typiquement dans notre travail, pour générer une trajectoire de $100ps$, les calculs ont duré environ 8 heures. Pour simuler un matériau infini dans une, deux ou trois dimensions, on placera les particules dans un espace périodique : on parlera alors d'une boîte de simulation. Toutes ces notions seront développées dans cette partie.

1.3.1 Equations de mouvement

Dans un système isolé avec un Hamiltonien $H(\vec{r}(t), \vec{p}(t))$, le changement d'état du système au cours du temps est gouverné par les équations d'Hamilton :

$$\dot{r}_{i\alpha} = \frac{\partial H}{\partial p_{i\alpha}} = \frac{p_{i\alpha}}{m_i}, \quad (1.22)$$

$$\dot{p}_{i\alpha} = -\frac{\partial H}{\partial r_{i\alpha}} = F_{i\alpha} = -\frac{\partial V(\vec{r})}{\partial r_{i\alpha}}, \quad (1.23)$$

où m_i est la masse de la particule i , $\dot{r}_{i\alpha}$ sa vitesse dans la direction cartésienne α et $\dot{p}_{i\alpha}$ le taux de variation de sa quantité de mouvement, égale à la force $F_{i\alpha}$. $V(\vec{r})$ est l'énergie potentielle.

Les deux équations 1.22 et 1.23 se résument à la seconde loi de Newton :

$$F_{i\alpha} = m_i \frac{d^2 r_{i\alpha}}{dt^2} = -\frac{\partial V(\vec{r})}{\partial r_{i\alpha}}. \quad (1.24)$$

Le système d'équations différentielles couplées (Eq. 1.24) doit être résolu numériquement pas à pas en discrétisant le temps. L'état initial du système est fixé en donnant les positions et vitesses à un instant initial t_0 . Les forces sont calculées (avec l'Eq. 1.24) et les équations intégrées pour générer la trajectoire. Pour cela, plusieurs algorithmes ont été développés.

1.3.2 Résolutions numériques des équations de mouvement : Algorithmes de Verlet

Ces algorithmes, nécessaires pour intégrer les équations de mouvement des particules le long de la trajectoire, ont tous pour origine les lois de la dynamique classique de Newton discrétisées. Le principe est d'obtenir depuis les forces (et potentiels) d'interaction entre les particules, les accélérations, puis les vitesses et les positions.

L'algorithme Verlet est le plus utilisé en dynamique moléculaire et cet algorithme peut être divisé en trois formes :

1. l'algorithme Verlet basique,
2. l'algorithme Leapfrog (algorithme saute-mouton),
3. l'algorithme de vitesse Verlet.

1.3.2.1 L'algorithme Verlet

L'idée principale de l'algorithme de Verlet est d'écrire un développement de Taylor à l'ordre 3 pour les positions, à un instant $t - \Delta t$ et un instant $t + \Delta t$, comme dans les équations suivantes :

$$\vec{r}(t + \Delta t) = \vec{r}(t) + \vec{v}(t)\Delta t + \frac{1}{2}\vec{a}(t)\Delta t^2 + \frac{1}{6}\vec{b}(t)\Delta t^3 + O(\Delta t^4), \quad (1.25)$$

$$\vec{r}(t - \Delta t) = \vec{r}(t) - \vec{v}(t)\Delta t + \frac{1}{2}\vec{a}(t)\Delta t^2 - \frac{1}{6}\vec{b}(t)\Delta t^3 + O(\Delta t^4), \quad (1.26)$$

où \vec{v} , \vec{a} et \vec{b} représentent la vitesse, l'accélération et la dérivée troisième de \vec{r} par rapport à t , respectivement. \vec{r} est un vecteur colonne à $3N$ composantes : $\vec{r} = (\vec{r}_1, \vec{r}_2, \vec{r}_3, \dots, \vec{r}_N)$ où \vec{r}_i désigne le vecteur position de la particule i et N est le nombre de particules.

En additionnant les deux expressions ci-dessus, nous obtenons la forme simple de l'algorithme de Verlet :

$$\vec{r}(t + \Delta t) = 2\vec{r}(t) - \vec{r}(t - \Delta t) + \vec{a}(t)\Delta t^2 + O(\Delta t^4), \quad (1.27)$$

$\vec{a}(t)$ représente la force divisée par la masse, elle est fonction de la position $\vec{r}(t)$:

$$\vec{a}(t) = - \left(\frac{1}{m} \right) \nabla V(\vec{r}(t)). \quad (1.28)$$

Les composantes du vecteur vitesse \vec{v} peuvent être calculées à partir des équations 1.25 et 1.26 :

$$\vec{v}(t) = \frac{\vec{r}(t + \Delta t) - \vec{r}(t - \Delta t)}{2\Delta t} + O(\Delta t)^3. \quad (1.29)$$

L'algorithme de Verlet nécessite un calcul des forces (Eq. 1.28) à chaque pas de temps et le stockage des positions au pas de temps t et $t - \Delta t$ (Eq. 1.27). Sa précision sur les vitesses (Eq. 1.29) peut être améliorée.

1.3.2.2 L'algorithme de Verlet Leapfrog

Le manque de précision sur les vitesses dans l'algorithme de Verlet peut être remédié par l'algorithme Leapfrog [Verlet 67]. Son nom vient du fait que les vitesses sont calculées pour des intervalles de temps demi-entiers alors que les positions sont obtenues pour des intervalles de temps entiers.

Nous pouvons ainsi approximer les vitesses pour les temps $t + \frac{\Delta t}{2}$ et $t - \frac{\Delta t}{2}$ par :

$$\vec{v} \left(t + \frac{\Delta t}{2} \right) = \frac{\vec{r}(t + \Delta t) - \vec{r}(t)}{\Delta t}, \quad (1.30)$$

$$\vec{v} \left(t - \frac{\Delta t}{2} \right) = \frac{\vec{r}(t) - \vec{r}(t - \Delta t)}{\Delta t}. \quad (1.31)$$

On obtient alors :

$$\vec{r}(t + \Delta t) = \vec{r}(t) + \vec{v}\left(t + \frac{\Delta t}{2}\right) \times \Delta t, \quad (1.32)$$

$$\vec{r}(t - \Delta t) = \vec{r}(t) - \vec{v}\left(t - \frac{\Delta t}{2}\right) \times \Delta t. \quad (1.33)$$

A partir de l'équation 1.27, on obtient la relation suivante :

$$\vec{v}\left(t + \frac{\Delta t}{2}\right) = \vec{v}\left(t - \frac{\Delta t}{2}\right) + \vec{a}(t)\Delta t + O(\Delta t^4). \quad (1.34)$$

Seul problème, les vitesses et les positions ne sont pas connues au même temps t , elles sont toujours déphasées d'un demi-temps.

1.3.2.3 L'algorithme de vitesse Verlet

Les positions, vitesses et accélérations au temps $t + \Delta t$ sont obtenues au même instant t , de la façon suivante [Swope *et al.* 82] :

1. Dans la première étape, le vecteur position \vec{r} au temps $t + \Delta t$ est calculé d'après la position, les vitesses \vec{v} et les accélérations \vec{a} au temps t , comme le montre l'équation suivante :

$$\vec{r}(t + \Delta t) = \vec{r}(t) + \vec{v}(t)\Delta t + \frac{1}{2}\vec{a}(t)\Delta t^2 + O(\Delta t^3), \quad (1.35)$$

2. Puis on calcule l'expression de $\vec{v}\left(t + \frac{\Delta t}{2}\right)$:

$$\vec{v}\left(t + \frac{\Delta t}{2}\right) = \vec{v}(t) + \frac{1}{2}\vec{a}(t)\Delta t + O(\Delta t^2), \quad (1.36)$$

3. L'accélération au temps $t + \Delta t$ peut alors être calculée :

$$\vec{a}(t + \Delta t) = -\left(\frac{1}{m}\right) \nabla V(\vec{r}(t + \Delta t)), \quad (1.37)$$

4. Et enfin, la vitesse au temps $t + \Delta t$ peut être évaluée de la façon suivante :

$$\vec{v}(t + \Delta t) = \vec{v}\left(t + \frac{1}{2}\Delta t\right) + \frac{\vec{a}(t + \Delta t)\Delta t}{2} + O(\Delta t^2). \quad (1.38)$$

L'avantage de cet algorithme est qu'il nécessite moins de mémoire vive, puisqu'un seul ensemble de positions, forces et vitesses sont nécessaires pour chaque pas. Afin de pouvoir tenir compte de toutes les interactions entre les atomes, le pas d'intégration pour l'algorithme Verlet est généralement d'une femtoseconde ($1fs = 10^{-15}s$).

1.3.3 Algorithmes pour maintenir T ou P constantes

1.3.3.1 Thermostat de Nosé Hoover

Pour nos simulations, nous avons chauffé les différentes protéines à l'aide du thermostat de Nosé Hoover [Evans et Holian 85]. Nosé décrit le thermostat par un degré de liberté supplémentaire que l'on ajoute au système mécanique. Ce degré de liberté permet de modifier la dynamique hamiltonienne en introduisant une force de "friction" qui va soit augmenter les vitesses des particules soit les diminuer.

Tout d'abord, considérons le Lagrangien suivant :

$$\mathcal{L} = \sum_{i=1}^N \frac{m_i s^2}{2} \left(\frac{dr_i}{dt} \right)^2 - V(\vec{r}) + \frac{Q}{2} \left(\frac{ds}{dt} \right)^2 - \frac{M}{\beta} \ln(s), \quad (1.39)$$

où M est un paramètre dont la valeur est à déterminer et $V(\vec{r})$ le potentiel entre les N particules. Q représente la masse effective de la variable spatiale unidimensionnelle s . Les équations de Lagrange de ce système s'écrivent :

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{r}_i} = m_i s^2 \dot{r}_i, \quad (1.40)$$

$$p_s = \frac{\partial \mathcal{L}}{\partial \dot{s}} = Q \dot{s}. \quad (1.41)$$

Cela conduit à un Hamiltonien :

$$H = \sum_{i=1}^N \frac{p_i^2}{2m_i s^2} + V(\vec{r}) + \frac{p_s^2}{2Q} + \frac{M}{\beta} \ln(s). \quad (1.42)$$

Nous allons exprimer la fonction de partition microcanonique d'un système à N particules couplées à ce degré de liberté supplémentaire :

$$Z = \frac{1}{N!} \int dp_s \int ds \int d\vec{p} \int d\vec{r} \delta(H - E), \quad (1.43)$$

qui traduit le fait que nous avons N particules indiscernables dans l'ensemble microcanonique classique où les configurations acceptables sont celles qui ont la même énergie et sont alors comptées avec un poids égal dans la fonction de partition.

En posant $\vec{p}' = \frac{\vec{p}}{s}$, on obtient :

$$Z = \frac{1}{N!} \int dp_s \int ds s^{3N} \int d\vec{p}' \int d\vec{r} \delta \left(\sum_{i=1}^N \frac{p_i'^2}{2m_i} + V(\vec{r}) + \frac{p_s^2}{2Q} + \frac{M}{\beta} \ln(s) - E \right), \quad (1.44)$$

En définissant H' comme :

$$H' = \sum_{i=1}^N \frac{p_i^2}{2m_i} + V(\vec{r}), \quad (1.45)$$

la fonction de partition devient

$$Z' = \frac{1}{N!} \int dp_s \int ds s^{3N} \int d\vec{p}' \int d\vec{r} \delta \left(H' + \frac{p_s^2}{2Q} + \frac{M}{\beta} \ln(s) - E \right). \quad (1.46)$$

On intègre alors sur la variable s en se servant de la propriété suivante :

$$\delta(h(s)) = \frac{\delta(s - s_0)}{h'(s_0)}, \quad (1.47)$$

où $h(s)$ est une fonction ne s'annulant qu'une seule fois sur l'axe réel pour la valeur $s = s_0$.

Ainsi on a :

$$\delta \left(H' + \frac{p_s^2}{2Q} + \frac{M}{\beta} \ln(s) - E \right) = \frac{\beta s}{M} \delta \left(s - \exp \left(-\frac{\beta}{M} \left(H' + \frac{p_s^2}{2Q} - E \right) \right) \right). \quad (1.48)$$

Cela donne finalement la fonction de partition :

$$Z' = \frac{\beta \exp \left(\frac{E(3N+1)}{M} \right)}{MN!} \int dp_s \int d\vec{p}' \int d\vec{r} \exp \left(\frac{\beta(3N+1)}{M} \left(H' + \frac{p_s^2}{2Q} \right) \right). \quad (1.49)$$

Avec le choix $M = 3N + 1$ la fonction de partition Z de l'ensemble microcanonique de départ, (système des N particules et de la variable supplémentaire), correspond à un système de N particules dans l'ensemble canonique avec les variables \vec{r} , \vec{p}' et t .

Une telle procédure conduit à un algorithme avec un pas de temps variable, ce qui n'est pas facilement manipulable dans une simulation de dynamique moléculaire. Il est possible d'éliminer le pas de temps variable et de travailler avec un pas de temps fixe en effectuant des transformations supplémentaires. Ce travail a été réalisé par Hoover et consiste à dériver à partir de ces équations établies par Nosé, un système d'équations qui peut être intégré avec un pas de temps constant. On a alors le système d'équations suivant :

$$\dot{r}_i = \frac{p_i}{m_i}, \quad (1.50)$$

$$\dot{p}_i = -\frac{\partial V(\vec{r})}{\partial r_i} - \xi p_i, \quad (1.51)$$

$$\dot{\xi} = \left(\sum_{i=1}^N \frac{p_i^2}{2m_i} - \frac{M}{\beta} \right) \frac{1}{Q}, \quad (1.52)$$

$$\frac{\dot{s}}{s} = \xi. \quad (1.53)$$

Les deux premières équations constituent un ensemble fermé d'équations pour l'évolution du système à N particules.

1.3.3.2 Pression constante dans l'ensemble NPT

Andersen développa en 1980 une méthode pour maintenir la pression constante dans un système [Andersen 80]. Cette méthode consiste à coupler le système à un volume externe variable V , équivalent au volume de la boîte de simulation. Ceci représentera ainsi l'action d'un piston, appelé le piston de Langevin [Feller *et al.* 95].

Le piston a une masse Q qui est associée à une énergie cinétique K_V :

$$K_V = \frac{1}{2}Q\dot{V}^2. \quad (1.54)$$

L'énergie potentielle associée E_{pV} est définie par :

$$E_{pV} = PV, \quad (1.55)$$

où P est la pression voulue.

Les énergies potentielles et cinétiques associées aux molécules sont données en fonction des positions r et vitesses v :

$$r = V^{\frac{1}{3}}s, \quad (1.56)$$

$$v = V^{\frac{1}{3}}\dot{s}, \quad (1.57)$$

où s représente les degrés de liberté.

Ainsi

$$E_p = E_p\left(V^{\frac{1}{3}}s\right), \quad (1.58)$$

$$K = \frac{1}{2}m_i \sum_{i=1}^N v_i^2 = \frac{1}{2}m_i V^{\frac{2}{3}} \sum_{i=1}^N \dot{s}_i^2. \quad (1.59)$$

où N représente le nombre de molécules i , de masse m_i .

Durant la simulation, les paramètres de la maille seront recalculés à chaque pas afin de définir le volume pour lequel la pression sera conservée.

1.3.4 Conditions périodiques

Pour réduire les effets de bord dans un système fini, on utilise souvent les conditions périodiques. Les atomes du système simulé sont introduits dans une boîte cubique entourée d'images identiques translattées de R_{box} .

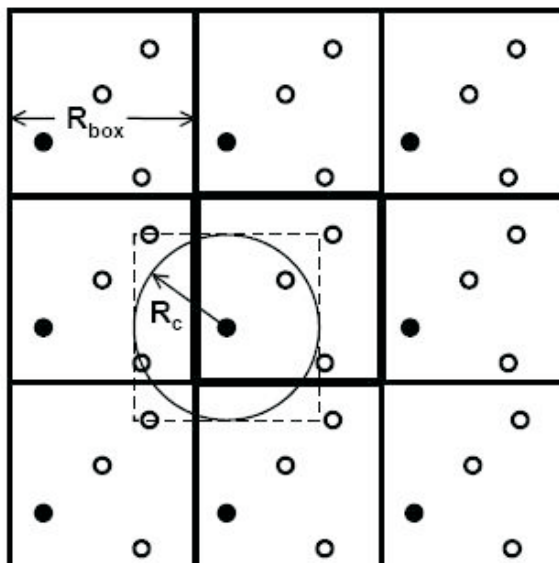


FIG. 1.1 – Conditions périodiques illustrées en deux dimensions avec une boîte cubique.

Le calcul des forces sur l'atome noir dans la boîte centrale s'effectue à partir des contributions des autres atomes de cette boîte ou des boîtes avoisinantes qui se trouvent à l'intérieur de la distance de troncature R_c . A cause de la périodicité, quand un atome quitte la boîte d'un côté, un autre atome identique entre en même temps du côté opposé, après translation de R_{box} , et avec la même vitesse. Le maintien d'une image minimale nécessite la définition d'un rayon de coupure sphérique (appelé "*cutoff*") pour les forces non liantes qui soit d'au moins la moitié de la longueur d'un côté de la boîte. L'application des conditions périodiques introduit un ordre à longue distance conforme ou non à la réalité. Elle permet de simuler un milieu condensé macroscopique.

Les conditions périodiques aux limites facilitent le calcul des forces électrostatiques en utilisant la sommation d'Ewald (appelée aussi somme d'Ewald). L'avantage de cette approche repose sur la convergence rapide de la sommation du potentiel électrostatique dans l'espace réciproque. La méthode fut développée par Paul Peter Ewald en 1921 afin de déterminer l'énergie électrostatique (et ainsi la constante de Madelung) des cristaux ioniques [Ewald 21]. Cependant, les conditions périodiques introduisent des artefacts [Cheatham *et al.* 95], comme par exemple, la stabilisation artificielle des hélices alpha lors des travaux de Mac Cammon [McCammon *et al.* 77].

1.3.5 Minimisation de l'énergie

La recherche des minima consiste à déterminer les valeurs des coordonnées $\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N$ pour lesquelles la fonction d'énergie potentielle $V(\vec{r})$ est minimum. Soit la fonction d'énergie potentielle $V(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)$, on cherche les solutions telles que :

$$\frac{\partial V}{\partial r_{i\alpha}} = 0 \text{ et } \frac{\partial^2 V}{\partial r_{i\alpha}^2} > 0, \forall i, \alpha. \quad (1.60)$$

La résolution analytique de l'Eq. 1.60 n'est pas envisageable car trop complexe. On envisagera plutôt des méthodes numériques qui changent graduellement les coordonnées pour produire des conformations d'énergie de plus en plus basse. De cette façon, on se déplace presque exclusivement vers le bas de l'hypersurface d'énergie et le minimum atteint est en général le minimum local le plus proche de la configuration initiale (cf figure 1.2).

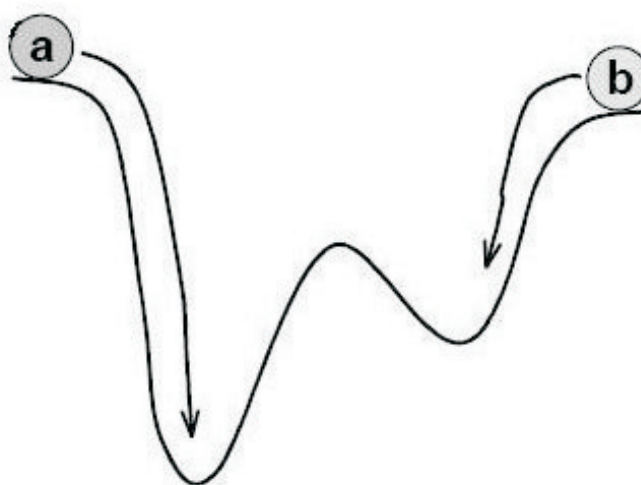


FIG. 1.2 – Inconvénient des méthodes de gradient : le minimum local atteint par les deux boules a et b est différent et dépend de la position (configuration) initiale.

Pour le choix de la méthode de minimisation, nous avons pris la plus rapide avec une utilisation minimale de mémoire vive. Il n'existe pas de méthodes qui soient systématiquement meilleure qu'une autre, cela dépend du système. Dans notre étude, l'optimisation de l'énergie se fera par la méthode des gradients conjugués [Fletcher et Reeves 64].

La méthode des gradients conjugués est un algorithme qui trouve le minimum local le plus proche d'une fonction à n variables. Cet algorithme consiste à minimiser une fonction selon une première direction \vec{u}_k . On arrive alors à un point \vec{X}_k qui est le minimum le long de cette direction. Pour continuer à minimiser la fonction $f(\vec{X}_k)$ dans une deuxième direction \vec{u}_{k+1} et sans détruire le minimum \vec{X}_k , il faudra ensuite prendre une direction perpendiculaire à \vec{u}_k .

L'algorithme peut donc être écrit de la façon suivante :

$$\vec{X}_{k+1} = \vec{X}_k + d_k \vec{u}_k, \quad (1.61)$$

où d_k est la distance de descente et \vec{u}_k est la direction de descente. Ainsi, la direction de descente \vec{u}_k variera à chaque étape en ayant toujours $\vec{u}_k \perp \vec{u}_{k+1}$.

L'algorithme des gradients conjugués peut se résumer de la sorte :

1. Choisir un point de départ \vec{X}_k ,
2. Construire le nouveau vecteur \vec{u}_k , défini par :

$$\vec{u}_k = \vec{\nabla} f(\vec{X}_k) + \frac{\|\vec{\nabla} f(\vec{X}_k)\|}{\|\vec{\nabla} f(\vec{X}_{k-1})\|} \vec{u}_{k-1}, \quad (1.62)$$

3. Calculer la longueur de descente d_k :

$$d_k = \frac{-\vec{\nabla} f(\vec{X}_k) \cdot \vec{u}_k}{(A\vec{u}_k) \cdot \vec{u}_k}, \quad (1.63)$$

où A est appelé le "Hessian" de $f(x)$, il représente la matrice des dérivées secondes au point P . Il est donc défini par :

$$A = \left[\begin{array}{ccc} \frac{\partial^2 f}{\partial x \partial x} & \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial x \partial z} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y \partial y} & \frac{\partial^2 f}{\partial y \partial z} \\ \frac{\partial^2 f}{\partial z \partial x} & \frac{\partial^2 f}{\partial z \partial y} & \frac{\partial^2 f}{\partial z \partial z} \end{array} \right] \Bigg|_P, \text{ calculée au point } P, \quad (1.64)$$

4. Calculer le nouveau vecteur \vec{X}_{k+1} , défini par l'équation 1.61.

L'algorithme s'arrête soit lorsque le gradient devient très petit, soit quand les \vec{X}_k n'évoluent plus.

Analyse des trajectoires et des fluctuations

2.1 Structure moyenne et trajectoire de référence

La trajectoire, que l'on notera $\overleftrightarrow{\Gamma}_r(\vec{r}(t))$, est un vecteur ligne défini par la série temporelle $[\vec{r}(t_0), \vec{r}(t_0 + \Delta t), \dots, \vec{r}(t_0 + n\Delta t)]$ où n est le nombre de fois où les coordonnées de \vec{r} sont mesurées et stockées. En général, $\Delta t \simeq 0,5 - 1ps$ alors que le pas d'intégration $\delta t \simeq 1fs$. Pour rappel, \vec{r} est un vecteur colonne contenant toutes les positions des atomes de la molécule. L'information structurale d'une trajectoire de dynamique moléculaire est ainsi contenue en réalité dans une matrice $\overleftrightarrow{\Gamma}_r(\vec{r}(t))$ de $3N$ lignes et $n + 1$ colonnes où N est le nombre d'atomes. L'ensemble des états visités est représenté par l'union de cette matrice et la matrice correspondante des vitesses $\overleftrightarrow{\Gamma}_v(\vec{v}(t))$. Pour une molécule dans un solvant, et qui est capable de diffuser, la trajectoire $\overleftrightarrow{\Gamma}_r$ peut être formellement décomposée en deux termes : les mouvements diffusifs globaux de translation et de rotation moléculaire et les fluctuations structurales internes. Pour une molécule qui serait rigide, cette séparation est exacte et unique : les fluctuations structurales internes peuvent se calculer en définissant un repère moléculaire rigide attaché à la molécule. Pour une molécule flexible, la séparation des mouvements globaux et internes est réalisée de manière approximative, la meilleure séparation possible entre mouvements locaux et globaux est obtenue de manière variationnelle.

Dans notre travail, nous ne nous intéressons pas aux mouvements globaux diffusifs des protéines qui dépendent essentiellement de leur taille et de la viscosité du solvant. Nous voulons au contraire étudier les fluctuations structurales par rapport à une structure moyenne de référence. La structure s'obtient en général à partir d'une trajectoire de référence $\overleftrightarrow{\Gamma}_r^{ref}(\vec{r}(t))$ obtenue en soustrayant numériquement les rotations et translations globales.

La trajectoire de référence est obtenue à partir de l'algorithme suivant [Kabsch 76] :

1. Soit une structure choisie à un instant t_0 (1ère colonne de la matrice $\overleftrightarrow{\Gamma}_r$),
 2. La structure considérée à l'instant ultérieur $t_0 + \Delta t$ (2ième colonne de la matrice $\overleftrightarrow{\Gamma}_r$) est superposée à la structure choisie au temps t_0 en lui appliquant la translation \vec{T} et la rotation rigide \vec{R} qui minimise les écarts entre les structures ainsi superposées,
 3. L'étape 2 est répétée pour tous les instants supérieurs en $t_0 + k\Delta t$, où $k = 2, \dots, n$.
- A l'étape 2, la translation \vec{T} est obtenue en superposant les centres de masses :

$$\vec{r}(t_0 + k\Delta t) = \vec{r}(t_0 + k\Delta t) + \vec{T}, \quad (2.1)$$

$$\vec{T} = \sum_{i=1}^N w_i (\vec{r}_i(t_0) - \vec{r}_i(t_0 + k\Delta t)), \quad (2.2)$$

$$w_i = \frac{m_i}{\sum_{j=1}^N m_j}, \quad (2.3)$$

où m_i est la masse de l'atome i . et donc

$$\sum_{i=1}^N w_i \vec{r}_i(t_0 + k\Delta t) = \sum_{i=1}^N w_i \vec{r}_i(t_0). \quad (2.4)$$

La rotation \overleftrightarrow{R} est obtenue en minimisant la fonction suivante [Kneller et Hinsen 04, Flower 99] :

$$m(\overrightarrow{q}) = \sum_{i=1}^N \beta_i \left(\overleftrightarrow{R}(\overrightarrow{q}) \tilde{r}_i(t_0 + k\Delta t) - \tilde{r}_i(t_0) \right)^2, \quad (2.5)$$

où $\tilde{r}_i(t_0 + k\Delta t)$ est la position de l'atome i au temps $k\Delta t$ translaturée (cf Eq. 2.1), β_i est le poids associé à l'atome i ($\sum_{i=1}^N \beta_i = 1$). La matrice $\overleftrightarrow{R}(\overrightarrow{q})$ dépend de quatre paramètres réels (q_0, q_1, q_2, q_3) définissant un quaternion \overrightarrow{q} . Le premier terme est un scalaire, alors que les trois derniers forment un vecteur de l'espace 3D. Le quaternion associé à la rotation ϕ autour d'un axe de rotation n d'un vecteur de l'espace 3D est :

$$q_0 = \cos \frac{\phi}{2}, \quad (2.6)$$

$$\begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} = \sin \frac{\phi}{2} \vec{n}. \quad (2.7)$$

La matrice de rotation \overleftrightarrow{R} est :

$$\overleftrightarrow{R}(\overrightarrow{q}) = \begin{pmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(-q_0q_3 + q_1q_2) & 2(q_0q_2 + q_1q_3) \\ 2(q_0q_3 + q_1q_2) & q_0^2 + q_2^2 - q_1^2 - q_3^2 & 2(-q_0q_1 + q_2q_3) \\ 2(-q_0q_2 + q_1q_3) & 2(q_0q_1 + q_2q_3) & q_0^2 + q_3^2 - q_1^2 - q_2^2 \end{pmatrix}. \quad (2.8)$$

Le fait que $\overleftrightarrow{R}(\overrightarrow{q})$ représente une rotation est imposé par la contrainte :

$$q_0^2 + q_1^2 + q_2^2 + q_3^2 = +1. \quad (2.9)$$

Les moyennes statistiques (cf chapitre 1) sont calculées sur la trajectoire de référence $\overleftrightarrow{\Gamma}_r^{ref}(\tilde{r}(t))$ ainsi obtenue. On peut se passer de cette étape lorsque les variables fluctuantes sont directement des coordonnées internes de la molécule (par exemple : un angle diédral ou une distance entre résidus). La trajectoire de référence est aussi appelée trajectoire "réorientée".

Afin de vérifier que l'algorithme implémenté dans CHARMM réoriente correctement les trajectoires, nous avons réalisé la même étude avec le package MMTSB [Feig *et al.* 04]. Ce package nous a ainsi permis de calculer le MSD des atomes avec deux algorithmes différents. La réorientation des trajectoires est automatique avec MMTSB. Nous avons alors calculé les RMS des atomes de la crambine S22/I25 sur une trajectoire réorientée par CHARMM et avec MMTSB (les résultats pour le 1er atome de la protéine sont donnés dans le tableau suivant).

	CHARMM	MMTSB
RMS du 1er atome (Å)	0.612	0.607

TAB. 2.1 – Comparaison du RMS obtenu avec CHARMM et celui généré par MMTSB.

Nous pouvons remarquer que les différences obtenues entre ces deux calculs sont très faibles. Nous pouvons donc conclure que l'algorithme utilisé par CHARMM réoriente correctement les trajectoires simulées.

2.2 Fluctuations structurales

Toutes les propriétés statistiques structurales sont calculées sur la trajectoire de référence (excepté si les fluctuations concernent des coordonnées internes). Dans ce paragraphe, nous définissons les grandeurs statistiques utilisées dans ce travail. Pour simplifier les notations, nous omettrons le symbole \sim qui identifie les coordonnées des atomes de la trajectoire de référence \vec{r}_i dans la section 2.1 et les noterons simplement \vec{r}_i dans cette section.

2.2.1 Variance des déplacements atomiques au Mean - Square - Displacement (MSD)

Pour un atome :

$$\sigma_i^2 = \langle \delta \vec{r}_i \cdot \delta \vec{r}_i \rangle = \langle \vec{r}_i \cdot \vec{r}_i \rangle - \langle \vec{r}_i \rangle \cdot \langle \vec{r}_i \rangle, \quad (2.10)$$

Pour la protéine :

$$\sigma_\alpha^2 = \sum_{i=1}^N \alpha_i \sigma_i^2, \quad (2.11)$$

α_i est une fonction poids. Le MSD moyen est obtenu en posant $\alpha_i = \frac{1}{N}$. Il sera noté dans notre thèse $\frac{MSD}{N}$.

Le MSD des atomes d'hydrogène est obtenu en calculant la moyenne sur ces atomes d'hydrogène. α_i peut représenter également un rapport de masse $\alpha_i = \frac{m_i}{\sum_j m_j}$.

2.2.2 Coordonnées relatives au centre de masse

On définit le vecteur $\vec{O}l$ qui représente le centre de masse :

$$\vec{O}l = \sum_{i=1}^N w_i \vec{r}_i, \quad (2.12)$$

où

$$w_i = \frac{m_i}{\sum_{j=1}^N m_j} \equiv \frac{m_i}{M}. \quad (2.13)$$

On définit ensuite les positions des atomes par rapport au centre de masse :

$$\vec{r}_i l = \vec{r}_i - \vec{O}l. \quad (2.14)$$

Notons que

$$\sigma_i^2 = \langle \delta \vec{r}_i \cdot \delta \vec{r}_i \rangle = \langle \delta \vec{r}_i l \cdot \delta \vec{r}_i l \rangle. \quad (2.15)$$

2.2.3 Rayon de giration et rayon de giration moyen

Mathématiquement, le rayon de giration se définit de la manière suivante :

$$R_{gyr}^2 = \sum_{i=1}^N w_i \langle (\vec{r}_i')^2 \rangle. \quad (2.16)$$

Le rayon de giration R_{gyr} représente l'état de compaction d'une protéine au cours d'une simulation. Plus R_{gyr} est grand, plus la protéine est étendue. Au contraire, plus R_{gyr} est petit, plus la protéine est "compacte". Le rayon de giration est lié au tenseur \overleftrightarrow{I} d'inertie :

$$R_{gyr}^2 = \frac{1}{M} \left(\sum_{i=1}^N m_i \langle x_i'^2 + y_i'^2 + z_i'^2 \rangle \right), \quad (2.17)$$

où M est la masse totale de la protéine.

Soit

$$\begin{aligned} R_{gyr}^2 &= \frac{1}{2M} \left\langle \sum_{i=1}^N m_i (y_i'^2 + z_i'^2) + \sum_{i=1}^N m_i (x_i'^2 + y_i'^2) + \sum_{i=1}^N m_i (x_i'^2 + z_i'^2) \right\rangle, \\ &= \frac{1}{2M} \langle I_{xx} + I_{yy} + I_{zz} \rangle = \frac{1}{2M} \text{Trace} \left(\langle \overleftrightarrow{I} \rangle \right). \end{aligned} \quad (2.18)$$

Selon l'équation 2.18, le rayon de giration ne dépend pas de l'orientation de la molécule. Il peut en effet être calculé à partir des distances entre atomes. On montre aisément que :

$$R_{gyr}^2 = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N w_i w_j \langle d_{ij}^2 \rangle, \quad (2.19)$$

où

$$d_{ij} = |\vec{r}_i' - \vec{r}_j'| = |\vec{r}_i - \vec{r}_j|. \quad (2.20)$$

On définit ensuite le déplacement atomique $\delta \vec{r}_i'(t)$:

$$\delta \vec{r}_i'(t) = \vec{r}_i'(t) - \langle \vec{r}_i' \rangle. \quad (2.21)$$

Donc, d'après l'équation 2.16 et 2.21, on trouve :

$$R_{gyr}^2 = \sum_{i=1}^N w_i \langle (\delta \vec{r}_i'(t) + \langle \vec{r}_i' \rangle)^2 \rangle. \quad (2.22)$$

Soit finalement :

$$R_{gyr}^2 = \underbrace{\sum_{i=1}^N w_i \langle \delta \vec{r}_i'^2 \rangle}_{MSD} + \underbrace{\sum_{i=1}^N w_i \langle (\langle \vec{r}_i' \rangle)^2 \rangle}_{\bar{R}_g^2}. \quad (2.23)$$

Le premier terme de l'équation 2.23 est le MSD (cf Eq. 2.10) et le second terme définit ce que nous appelons dans notre thèse le rayon de giration moyen \bar{R}_g^2 . R_g peut être comparé à l'expérience. En effet, en cristallographie on connaît $\langle \vec{r}_i' \rangle$ à partir de la structure moyenne. On peut alors en principe calculer \bar{R}_g^2 .

De plus, si les facteurs de températures B_i sont également connus, on peut calculer R_g^2 d'après l'équation 1.7 et via la formule :

$$R_{gyr}^2 = \bar{R}_g^2 + \sum_{i=1}^N w_i \frac{B_i}{8\pi^2}. \quad (2.24)$$

La précision sur les B-facteurs obtenus expérimentalement peut cependant varier considérablement. En effet, les B-facteurs sont des paramètres qui prennent en compte l'agitation thermique de l'atome mais aussi de toute la densité électronique non-paramétrée par le reste du modèle. Par exemple, si aucun modèle n'est appliqué pour éliminer la contribution aux B facteurs due à la liaison, la précision sur les B facteurs ne sera pas bonne.

2.2.4 Déplacement quadratique moyen

Le déplacement quadratique moyen d'un atome i au temps t , $\Delta_i(t)$, représente une mesure de la diffusion. Pour l'atome, il est défini par :

$$\Delta_i(t) = \langle [\vec{r}_i(t) - \vec{r}_i(0)]^2 \rangle, \quad (2.25)$$

où $\vec{r}_i(0)$ est la position à l'instant initial et $\vec{r}_i(t)$ est la position après un temps t . Les parenthèses $\langle \rangle$ représentent une moyenne sur toutes les conditions initiales $\vec{r}_i(0)$. Pour une trajectoire de référence, représentée par la série temporelle $[\vec{r}_i(t_0), \dots, \vec{r}_i(t_0 + n\Delta t)]$ ($\Delta t = 0.5ps$), la moyenne est calculée numériquement de la manière suivante :

$$\Delta_i(t) = \frac{1}{n-t} \left(\sum_{k=0}^{n-t} [\vec{r}_i(t+k) - \vec{r}_i(k)]^2 \right), \quad (2.26)$$

où $\vec{r}_i(k)$ représente la position de l'atome i à l'instant $t_0 + k\Delta t$.

2.2.5 Fonction de corrélation spatiale

La fonction de corrélation spatiale $C_i(t)$ pour un atome est définie par :

$$C_i(t) = \langle \delta\vec{r}_i(t) \cdot \delta\vec{r}_i(0) \rangle, \quad (2.27)$$

$C_i(t)$ mesure la corrélation du déplacement de l'atome i au temps t avec le déplacement à un temps antérieur initial $t = 0$.

Pour la protéine :

$$C_\alpha(t) = \sum_{i=1}^N \alpha_i C_i(t). \quad (2.28)$$

Cette fonction est directement reliée au MSD.

$$\sigma_i^2 = C_i(0), \quad (2.29)$$

$$\sigma_\alpha^2 = C_\alpha(0). \quad (2.30)$$

2.3 Variables conformationnelles à gros grains γ caractérisant le backbone

Pour chaque résidu d'une chaîne polypeptidique, les fluctuations structurales des atomes formant le backbone (N, O', C $^\alpha$) (cf chapitre 4, partie I) sont habituellement décrites par les deux angles diédraux de Ramachandran ϕ and ψ [Creighton 94].

Dans notre travail de thèse (chapitre 3, partie II), nous allons nous intéresser aux fluctuations du backbone représentées par le seul mouvement des atomes C $^\alpha$. Par conséquent, nous utiliserons un seul angle diédral, appelé γ , pour chaque résidu [Nishikawa *et al.* 74]. Cette représentation simplifiée, où le nombre de variables est réduit, est appelée un modèle à gros grains de la protéine. La quantité γ_n pour le résidu n est l'angle diédral formé par les vecteurs (appelés liaisons virtuelles) joignant quatre atomes C $^\alpha$ ($n - 1, n, n + 1, n + 2$) le long de la séquence primaire. Chaque résidu n , pour n compris entre 2 et $M - 2$, où M est le nombre total de résidus de la protéine, peut être caractérisé par un angle diédral γ_n (γ_1, γ_{M-1} et γ_M n'existent pas). Les trajectoires des angles γ sont calculées numériquement à partir des coordonnées des atomes obtenues dans nos simulations de dynamique moléculaire.

2.4 Méthode d'analyse DFA (Detrended Fluctuation Analysis) et exposant fractal des séries temporelles $\gamma(t)$

Le déplacement quadratique moyen au temps t de l'angle diédral γ_p associé au résidu p s'écrit :

$$\Delta_{\gamma_p}(t) = \left\langle [\gamma_p(t) - \gamma_p(0)]^2 \right\rangle, \quad (2.31)$$

où la moyenne est calculée sur toutes les conditions initiales, $\gamma_p(0)$, de la trajectoire de la même manière que dans l'équation 2.26. L'analyse de la loi d'évolution temporelle (Eq. 2.31) sera effectuée en appliquant un algorithme appelé DFA pour Detrended Fluctuation Analysis [Peng *et al.* 94]. Nous sommes les premiers à avoir appliqué cet algorithme aux protéines. L'algorithme DFA que nous allons décrire permet d'analyser les propriétés statistiques d'une série temporelle non-stationnaire, c'est-à-dire dont la valeur moyenne varie au cours du temps. C'est toujours le cas des trajectoires de dynamique moléculaire de protéines car d'une part, les temps de simulations sont souvent trop courts pour atteindre un équilibre stationnaire et d'autre part, car la nature anharmonique de la dynamique d'une protéine implique un tel comportement.

L'algorithme DFA est appliqué à la trajectoire $\gamma_n(t)$ de la manière suivante. La trajectoire est représentée par sa série temporelle $[\gamma_p(t_0), \gamma_p(t_0 + \Delta t), \dots, \gamma_p(t_0 + n\Delta t)]$ de longueur $n + 1$ où $n + 1$ est le nombre de fois que les angles diédraux sont calculés et enregistrés en mémoire. Comme pour la position de l'atome au cours du temps, la notation $\gamma_p(i)$ représentera la valeur de l'angle au temps $t_0 + i\Delta t$. La durée de la trajectoire est $n\Delta t$ où $\Delta t = 0.5ps$. La séquence des données $\gamma_p(i)$ est divisée en n/m fenêtres temporelles qui ne se recouvrent pas, chacune contenant $m + 1$ valeurs de l'angle γ_p et de durée $m\Delta t$. Nous avons choisi une valeur minimale de m égale à 3. Dans chacune des n/m fenêtres, les données sont approximées par

une loi linéaire, $y_p(i) = ai + b$, en utilisant la méthode d'ajustement par moindres carrés. La valeur des coefficients a et b ajustés dépendent de la position de la fenêtre pour une valeur de m donnée.

Pour chaque valeur de m , les fluctuations $f_p^2(m)$ relatives à la meilleure approximation linéaire, $y_p(i) = ai + b$, dans chaque fenêtre, sont calculées à partir de l'équation suivante :

$$f_p^2(m) = \frac{1}{m+1} \sum_{i=(l-1)m}^{ml} [\gamma_p(i) - y_p(i)]^2, \quad l = 1, 2, \dots, n/m. \quad (2.32)$$

La moyenne de la fonction $f^2(m)$ sur les n/m fenêtres donne le MSD, $F_p^2(m) = \langle f_p^2(m) \rangle$, pour une taille de fenêtre m . La quantité $F^2(m)$ est équivalente au déplacement quadratique moyen de l'angle $\Delta_{\gamma_p}(t)$ au temps $t = m\Delta t$ mais calculée pour une trajectoire dont on a enlevé les "tendances". Plus précisément, les fluctuations calculées à partir de l'équation 2.32 sont relatives aux fluctuations par rapport à la valeur moyenne locale calculée dans chaque fenêtre temporelle de longueur $m\Delta t$.

L'algorithme DFA permet de mesurer l'exposant self-affine d'une série temporelle. Une série temporelle aléatoire stationnaire est dite self-affine (ou fractale) si la distribution de probabilité des fluctuations observées à des temps courts est similaire à celles observées sur des temps longs [Mandelbrot et Ness 68]. Pour de telles séries temporelles, le déplacement quadratique moyen suit la loi suivante [Mandelbrot et Ness 68] :

$$MSD(t) = A t^{2\beta}. \quad (2.33)$$

La quantité β est la dimension fractionnaire et est aussi appelée dimension de Hurst [Mandelbrot et Ness 68]. La trajectoire d'un mouvement Brownien est self-affine avec un exposant $\beta = 0.5$ et avec $A = 2D$ où D est le coefficient de diffusion. Dans ce cas, l'équation 2.33 représente la loi classique de diffusion d'Einstein pour une particule libre à une dimension. Pour cette raison, des séries temporelles self-affines qui obéissent à l'équation 2.33 mais avec un exposant $\beta \neq 0.5$ représentent un mouvement diffusif dit "anormal", la diffusion "anormale" se rencontre en réalité dans de nombreuses situations en physique, en particulier pour la diffusion de particules dans les solides complexes [Bouchaud et Georges 90]. Lorsque l'exposant β est inférieur à 0.5, l'exploration aléatoire de l'espace est dite sous-diffusive (situation rencontrée dans les protéines, cf chapitre 3, partie II). Lorsque l'exposant β est supérieur à 0.5, l'exploration de l'espace est dite super-diffusive (situation rencontrée dans les phénomènes de turbulence). La loi de diffusion "anormale" (Eq. 2.33 avec $\beta \neq 0.5$) implique des corrélations temporelles [Mandelbrot et Ness 68].

Dans notre travail (chapitre 3, partie II), nous avons représenté le logarithme de la racine carrée de la fonction $F_p^2(m)$ en fonction du logarithme de m et nous avons ajusté le résultat par une loi linéaire en utilisant la méthode des moindres carrés afin de vérifier si la loi de diffusion des angles diédraux suit une loi en puissance et afin d'en déterminer l'exposant β (Eq. 2.33).

En pratique, l'algorithme DFA produit une fonction $F_p(m)$ correcte, c'est-à-dire qui a une signification statistique, entre une valeur de m minimale et une valeur de m maximale [Hu *et al.* 01]. Pour des valeurs de m trop petites, il n'y a pas assez de points dans chaque

fenêtre pour que les fluctuations calculées dans chacune des fenêtres aient un sens. Pour des valeurs de m trop élevées, il n'y a pas assez de fenêtres pour que la moyenne calculée dans l'équation 2.32 soit reproductible. La valeur de m maximale est environ de l'ordre de $n/10$. La valeur de m minimale dépend fortement de la valeur de l'exposant β .

Lorsque les exposants sont faibles $\beta \ll 0.5$, il est utile d'appliquer l'algorithme DFA, non pas directement à la série temporelle mais à son intégrale sur le temps. L'exposant obtenu en appliquant l'algorithme à l'intégrale du signal est appelé α dans notre travail. On doit s'attendre à la relation suivante entre les deux exposants :

$$\alpha = \beta + 1, \quad (2.34)$$

Puisque nous avons :

$$\int t^\beta dt = \frac{1}{\beta + 1} t^{\beta+1}. \quad (2.35)$$

Ce sont en réalité les exposants α qui sont calculés et présentés dans notre travail. Il faut noter que le régime de diffusion normale correspond ainsi à $\alpha = 1.5$ et le régime sous-diffusif à $\alpha < 1.5$. Dans une publication à venir, nous avons également calculé directement les exposants β et nous avons pu vérifier que la relation de l'équation 2.34 est en effet approximativement correcte.

2.5 Profil d'énergie libre pour les angles diédraux γ

L'hypothèse ergodique permet de calculer la densité de probabilité de chaque coordonnée interne d'une protéine et d'en déduire le profil d'énergie libre de la molécule le long de cette coordonnée. Au chapitre 3, nous présenterons ce profil d'énergie libre, appelé potentiel de force moyen (Potential of Mean Force ou PMF), pour chaque angle γ_p . Nous l'avons calculé de la manière suivante :

La série temporelle $[\gamma_p(t_0), \gamma_p(t_0 + \Delta t), \dots, \gamma_p(t_0 + n\Delta t)]$ est d'abord réordonnée par ordre croissant des valeurs des angles par un algorithme de tri. L'étendue des valeurs est ensuite divisée en fenêtres de 1 degré. La densité de probabilité $P(\gamma_p)$ pour une valeur d'angle donnée, soit $\gamma_p = x$, est calculée en comptant le nombre de fois que cette valeur apparaît dans la trajectoire dans l'intervalle $[x, x + 1^\circ[$, divisé par $n + 1$ et par 1 degré. Le potentiel de force moyen $V(\gamma_p)$ est calculé par la formule standard de Boltzmann [Landau et Lifshitz 84].

$$V(\gamma_p) = -k_B T \ln[P(\gamma_p)], \quad (2.36)$$

où $P(\gamma_p)$ est une densité de probabilité renormalisée, T est la température et k_B la constante de Boltzmann.

Mesures expérimentales des fluctuations structurales

3.1 Diffusion incohérente de neutrons

Un faisceau de neutrons de longueur d'onde typique de 1\AA a une énergie de l'ordre de kT à température ambiante, soit de l'ordre de 30 meV . La quantité mesurée expérimentalement est le facteur de structure dynamique $S(\vec{q}, \omega)$ auquel est proportionnelle l'intensité détectée.

$$S(\vec{q}, \omega) = S_{coh}(\vec{q}, \omega) + S_{incoh}(\vec{q}, \omega), \quad (3.1)$$

où \vec{q} est le vecteur d'onde échangé ($\hbar\vec{q} = \hbar\vec{k}' - \hbar\vec{k}$, avec \vec{k} et \vec{k}' respectivement les vecteurs d'onde incident et diffusé et \hbar est la constante de Planck) et $\hbar\omega$ est l'énergie des quanta absorbés ou excités par les particules sondes. La diffusion élastique correspond à $\omega = 0$. S_{coh} est la contribution cohérente de la diffusion due aux mouvements corrélés des atomes du cristal tandis que S_{incoh} est liée à la fonction d'autocorrélation du déplacement individuel des atomes.

La diffusion incohérente est une méthode importante d'analyse de la dynamique des protéines car elle est sensible aux déplacements des atomes d'hydrogène, constituant environ la moitié des atomes d'une protéine. La diffusion des neutrons par le noyau d'un atome est caractérisée par une longueur de diffusion ou par une section efficace σ .

L'atome d'hydrogène (1H) a une section efficace $\sigma_{inc} = 80.27\text{ barn}$ qui est largement supérieure de plusieurs ordres de grandeur à la section efficace σ_{coh} de tous les autres atomes d'une protéine. La diffusion incohérente des neutrons par les 1H domine donc le spectre des macromolécules biologiques. De plus, la contribution des atomes d'hydrogène de l'eau d'hydratation peut être séparée de celle des hydrogènes de la protéine en utilisant de l'eau deutérée. En effet, l'isotope 2H a une section efficace très petite $\sigma_{inc} = 2\text{ barn}$. De surcroît, cette différence de sections efficaces peut être exploitée pour étudier séparément la dynamique de différentes parties d'une structure moléculaire en suivant la dynamique des hydrogènes 1H liés à des groupements chimiques spécifiques de la protéine.

La partie incohérente $S_{incoh}(\vec{q}, \omega)$ est la transformée de Fourier de la fonction intermédiaire de diffusion $I(\vec{q}, t)$ et dépend de la résolution expérimentale.

$$S_{incoh}(\vec{q}, \omega) = \int_{-\infty}^{+\infty} d\omega' R(\omega - \omega') \frac{1}{\pi} \int_0^{+\infty} I(\vec{q}, t) \cos(\omega' t) dt, \quad (3.2)$$

où $R(\omega - \omega')$ est la fonction de résolution expérimentale (avec $\int_{-\infty}^{+\infty} d\omega R(\omega) = 1$) qui peut être modélisée schématiquement par un signal rectangulaire (cf Fig. 3.1) [Bicout 05].

La résolution temporelle T varie entre quelques dizaines de picosecondes à 1 ns : les mouvements plus lents ne peuvent être détectés. Par exemple, un spectromètre avec une résolution d'énergie de 100 meV correspond à des fenêtres temporelles de 10 ps , ou encore le spectromètre IN13 ($10\mu\text{eV}$), à l'Institut Langevin de Grenoble, correspond à des dynamiques de 400 ps [Hayward *et al.* 03].

Tous les atomes d'hydrogène d'une protéine ne sont pas dynamiquement équivalents : les atomes des groupes méthyles ou ceux participant aux liaisons hydrogène du backbone ont certainement une dynamique différente. La fonction $I(\vec{q}, t)$ moyenne ces contributions :

$$I(\vec{q}, t) = \sum_{\alpha} I_{\alpha} \langle e^{i\vec{q}\cdot\vec{u}_{\alpha}(t)} e^{-i\vec{q}\cdot\vec{u}_{\alpha}(0)} \rangle, \quad (3.3)$$

où I_{α} donne la fraction des hydrogènes de type α contribuant à l'intensité ($\sum_{\alpha} I_{\alpha} = 1$), $\vec{u}_{\alpha}(t)$ est le déplacement de l'atome α au temps t .

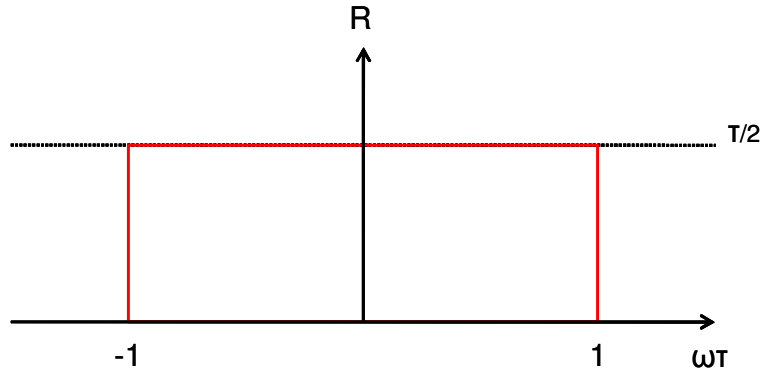


FIG. 3.1 – Fenêtre de résolution expérimentale où T est le temps de résolution temporelle limité par la résolution en énergie du spectromètre.

Lorsque le vecteur \vec{q} diffusé est assez petit et proche de zéro, l'expression 3.3 peut se simplifier en moyennant sur toutes les directions angulaires dans l'espace des vecteurs \vec{q} [Callaway 74].

$$I(\vec{q}, t) \approx \sum_{\alpha=1}^n I_{\alpha} e^{-\frac{q^2}{2}(\langle |\vec{u}_{\alpha}(t)|^2 \rangle + \langle |\vec{u}_{\alpha}(0)|^2 \rangle)} \times e^{-q^2 \langle \vec{u}(t) \vec{u}(0) \rangle}, \quad (3.4)$$

et à la limite où $t \rightarrow \infty$, on trouve :

$$I(\vec{q}, t \rightarrow \infty) \approx \sum_{\alpha=1}^n I_{\alpha} e^{-q^2 \langle x_{\alpha}^2 \rangle}. \quad (3.5)$$

En posant $\langle x_{\alpha}^2 \rangle = \frac{1}{3} \langle |\vec{u}_{\alpha}(t)|^2 \rangle$.

On peut donc définir les fluctuations structurales moyennes par :

$$\langle x^2 \rangle = \frac{1}{N} \sum_{\alpha=1}^n I_{\alpha} \langle x_{\alpha}^2 \rangle = -\frac{1}{N} \frac{d}{d(q^2)} (I(\vec{q}, \infty)) \Big|_{q=0}. \quad (3.6)$$

Chaque déplacement $\langle x_{\alpha}^2 \rangle$ dépend de la température. En pratique, $I(\vec{q}, t)$ dépend également de la fonction de corrélation des déplacements car la fenêtre temporelle de mesure est limitée par la résolution en énergie du spectromètre (cf Fig. 3.1).

L'influence de la résolution temporelle sur la mesure de $\langle x^2 \rangle$ est illustrée par la figure extraite de [Parak 03].

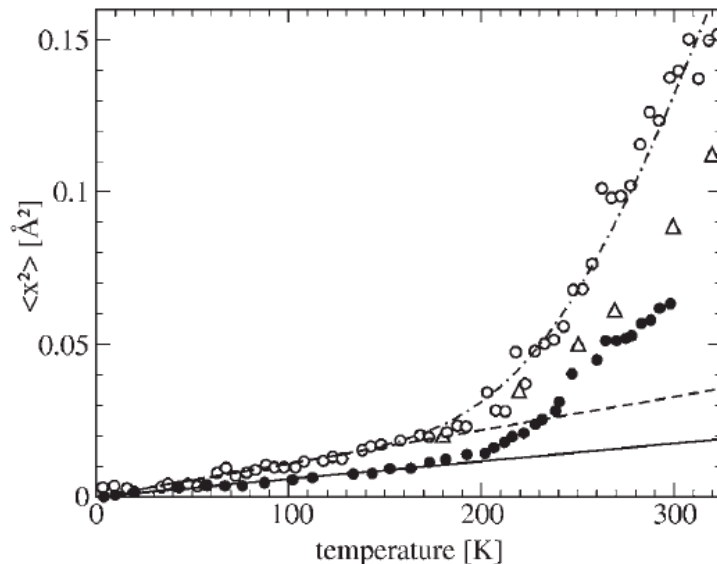


FIG. 3.2 – Etude de la myoglobine : symboles non-colorés : Mesure de l'écart quadratique moyen par diffusion de neutrons : cercles : pour une résolution de 80ps, triangles : pour une résolution de 12ps. cercles noirs : Ecart quadratique moyen mesuré par spectroscopie de Mössbauer.

3.2 Mesure de l'agitation thermique en diffraction X

La diffractométrie de rayons X produit une image en trois dimensions des composantes de Fourier (en réalité de leur module) de la densité d'électrons dans le réseau réciproque d'un cristal, de laquelle les positions atomiques, les liaisons, le désordre et d'autres informations peuvent être tirées.

L'énergie des photons X est largement plus élevée que celle des neutrons. Pour une même longueur d'onde $\lambda = 1\text{\AA}$, l'énergie incidente est de l'ordre de $\hbar\omega = \frac{hc}{\lambda} \simeq 12.5\text{keV}$. Le rayonnement est diffusé par la densité électronique. Contrairement aux neutrons, la section efficace de diffusion est plus élevée pour les atomes lourds et la résolution spatiale expérimentale ne permet pas en général de détecter la position des atomes d'hydrogène. L'intensité diffusée est également modulée par les fluctuations structurales. Dans l'approximation gaussienne [Borwein et Borwein 87], elle est quantifiée par le facteur atomique B_α :

$$B_\alpha = 8\pi \langle x_\alpha^2 \rangle, \quad (3.7)$$

où $\langle x_\alpha^2 \rangle = \frac{1}{3} \langle u_\alpha^2 \rangle$, où u_α^2 est le déplacement quadratique moyen ($\langle u_\alpha^2 \rangle = \langle r_\alpha^2 \rangle - \langle r_\alpha \rangle^2$).

Les B_α sont sensibles à la fois aux fluctuations de structure statiques et dynamiques et aux approximations liées au modèle de la densité électronique utilisé. Les fluctuations statiques sont dues aux variations d'orientations et de positions des protéines dans les différentes mailles du cristal. Ce désordre statique est pratiquement indépendant de la température et peut être estimé en extrapolant les B_α à $T \rightarrow 0$. Les fluctuations dynamiques proviennent

de la distribution structurale des molécules dans les différentes mailles qui sont dues à la dynamique intramoléculaire. L'ajustement des structures protéiques par la méthode des modes normaux permet d'extraire la contribution dynamique [Kidera et Go 92a, Kidera et Go 92b].

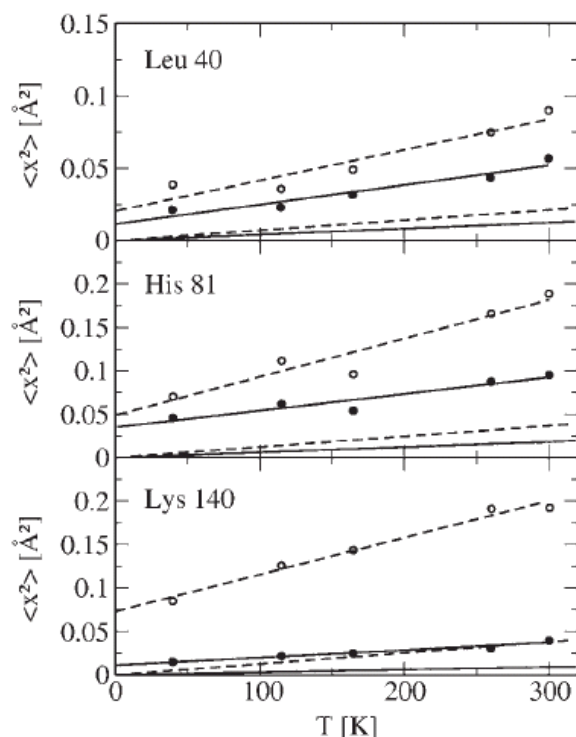


FIG. 3.3 – Dépendance en température de l'écart quadratique moyen des résidus de la myoglobine obtenue par les modes normaux [Chong *et al.* 01]. Symboles noirs et lignes continues : moyenne sur tous les atomes du backbone, symboles non-colorés et lignes en pointillés : moyenne sur les atomes des chaînes latérales [Parak 03].

3.3 Comparaison des fluctuations structurales mesurées et calculées

Les mouvements des protéines couvrent différentes échelles de temps : de la picoseconde pour de petites fluctuations des groupes d'atomes, jusqu'aux nano ou microsecondes, et même plus, pour les mouvements collectifs [Ringe et Petsko 03]. Nous nous sommes donc intéressés à des mouvements de l'ordre de la picoseconde jusqu'à 1 nanoseconde pour visualiser les fluctuations que subissent les atomes de la protéine, mais également car une étude pour un temps de simulation compris entre $50ps$ et $1ns$ correspond à des fenêtres expérimentales utilisables en neutrons sur des spectromètres. Contrairement aux cristaux réels, nos simulations emploient des conditions périodiques parfaites : aucun désordre "statique" ne contribue donc aux fluctuations.

3.4 La Résonance Magnétique Nucléaire

La RMN est l'outil d'analyse actuellement le plus utilisé en chimie organique. Cette technique a été mise au point par, entre autres, Kurt Wüthrich, qui reçut le prix Nobel de Chimie en 2002 [Wuthrich 90]. Elle permet d'obtenir des informations qualitatives ou quantitatives sur l'échantillon analysé, suivant la technique employée. Les noyaux les plus souvent étudiés sont le proton ^1H , le ^{13}C et le ^{31}P . La RMN est utilisée pour obtenir des informations sur la structure et la dynamique des protéines. La RMN peut également être employée pour déterminer les structures tertiaires des protéines, mais cette technique ne peut être employée que pour des molécules relativement petites.

Dans la banque de données PDB (Protein Data Bank ; www.rcsb.org/pdb/), les structures obtenues par RMN sont représentées par la superposition de plusieurs modèles compatibles avec l'ensemble des données. Ces modèles permettent d'estimer les fluctuations structurales de la protéine. Ceci est possible par le calcul du RMS entre les différents modèles. La RMN sera donc utilisée dans la suite de notre étude afin de contrôler nos résultats (cf chapitre 5, partie I).

Protéines étudiées et protocoles de simulations

4.1 Généralités sur les protéines

Les protéines sont des polymères d'acides aminés, reliés entre eux par des liaisons peptidiques. La structure primaire, ou séquence, d'une protéine correspond à la succession linéaire des acides aminés (ou résidus) la constituant. Par convention, on écrit le nom des acides aminés selon un code à trois lettres ou un autre code à une lettre.

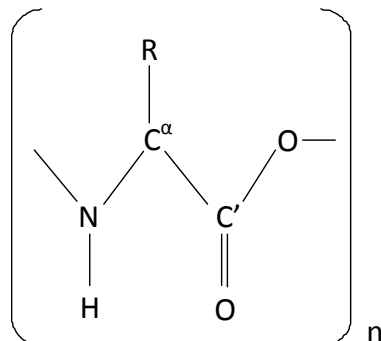


FIG. 4.1 – Formule générale d'un acide aminé. C^α est un carbone asymétrique (sauf pour la Glycine) et R représente la chaîne latérale de l'acide aminé.

La structure secondaire décrit le repliement local de la chaîne principale d'une protéine. L'existence de structures secondaires vient du fait que les repliements énergétiquement favorables de la chaîne peptidique sont limités et que seules certaines conformations sont possibles. Ainsi, une protéine peut être décrite par une séquence d'acides aminés mais aussi par un enchaînement d'éléments de structure secondaire. De plus, certaines conformations se trouvent nettement favorisées car stabilisées par des liaisons hydrogène entre les groupements amide (-NH) et carbonyle (-CO) du squelette peptidique [figure 4.1]. Il existe trois principales catégories de structures secondaires selon l'échafaudage de liaisons hydrogène, et donc selon le repliement des liaisons peptidiques : les hélices, les feuillets et les coudes.

La structure tertiaire est le résultat de liaisons diverses (hydrogène, hydrophobes, électrostatiques, covalentes,...) entre des acides aminés de la même chaîne peptidique voisins ou non dans la structure primaire. Les liaisons de la structure tertiaire sont plus ou moins stables : il y a des liaisons faibles de type liaisons hydrogène ou liaisons hydrophobes. Les liaisons électrostatiques sont dues principalement aux charges électriques des radicaux des acides aminés. Des liaisons covalentes établissent des liaisons de forte énergie entre les acides aminés : ce sont les ponts disulfures entre deux cystéines.

Le squelette (backbone) de la structure tertiaire est défini par la position dans l'espace des atomes N, C_α et C' des plans peptidiques (cf Fig.4.1).

4.2 Structure type des thionines

Les thionines sont de petites protéines basiques riches en cystéine, généralement toxiques pour les cellules animales. Elles sont constituées d'une chaîne polypeptidique de 45 à 50 acides aminés avec 3 à 5 ponts disulfures. La stabilité des protéines provient de ces ponts disulfures qui les rendent résistantes à l'élévation de la température. Il existe six formes différentes de viscotoxines, de la famille des α - β -thionines. Les protéines de cette famille peuvent être extraites de sources naturelles par des bains d'acides et par extraction de solvant organique [Pelegri et Franco 05].

La plupart des thionines sont naturellement chargées positivement. Elles ont une structure tertiaire caractéristique de la forme de la lettre grecque Γ ou de la lettre latine L avec deux paires d'hélices alpha anti-parallèles et un court feuillet beta anti-parallèle. D'autres thionines ont une structure similaire : la crambine ou encore l'alpha1-purothionine par exemple. Leur helix-turn-helix (HTH) motif pourrait représenter un motif de liaison pour l'ADN (Acide DésoxyriboNucléique, ce motif représente deux hélices alpha anti-parallèles reliées entre elles par une simple boucle non structurée). Elles sont, de plus, amphiphatiques (sauf la crambine) [Romagnoli *et al.* 00]. La structure de la crambine est d'une grande précision puisqu'il existe des données à ultra haute résolution (0,54 Å) obtenues par diffraction X [Jelsch *et al.* 00].

Les thionines sont donc composées de deux hélices α , de deux feuillets β anti-parallèles et de trois ponts disulfures (résidus 3 à 40, 4 à 32 et 16 à 26) ; toutes les protéines étudiées possèdent cette structure type.

La lettre L, symbole de la famille des thionines, est représentée en rouge sur la figure 4.2.

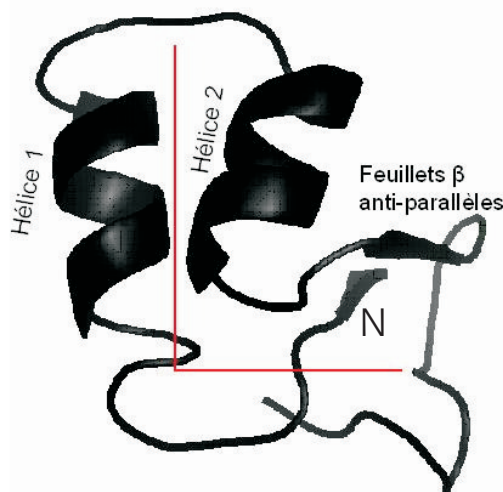


FIG. 4.2 – Backbone de la Viscotoxine A3, position des structures secondaires et structure typique en L, représentation avec le logiciel VMD. N indique l'extrémité de la protéine correspondante au résidu 1.

Dans le tableau suivant, nous répertorions les différences structurales entre les protéines étudiées ; les codes donnés correspondent aux codes de la banque de données PDB.

	α	β	charge	code PDB	technique
VA3	7 – 18 23 – 30	2 – 3 33 – 34	+6	1ED0 [Romagnoli <i>et al.</i> 00]	RMN
VB	6 – 19 22 – 30		+5	1JMP [Coulon <i>et al.</i> 03]	RMN
VA2	6 – 18 23 – 31	2 – 3 33 – 34	+5	1JMN [Mosbah <i>et al.</i> 01]	RMN
Crambine n°1	7 – 18 23 – 30	1 – 4 32 – 35	neutre	1CCM [Bonvin <i>et al.</i> 93]	RMN
Crambine n°2	7 – 16		neutre	1CXR [Xu <i>et al.</i> 99]	RMN

TAB. 4.1 – Données expérimentales sur les propriétés structurales des thionines.

La localisation des acides aminés sur le backbone des thionines est représentée sur la figure 4.3. Grâce à cette représentation, nous pourrions ainsi mettre en évidence dans notre travail les interactions possibles entre des résidus éloignés dans la séquence mais voisins dans l'espace.

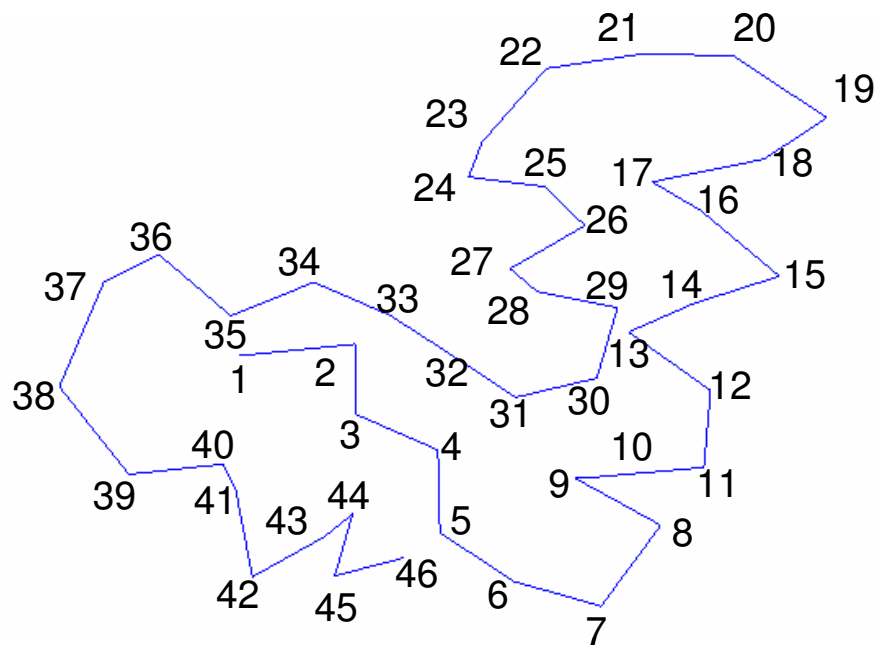


FIG. 4.3 – Localisation des acides aminés le long du squelette.

Les protéines diffèrent également par leur composition en acide aminé et par leur séquence primaire, listée dans le tableau 4.3. Dans le tableau ci-dessous, les labels des différents acides aminés sont définis.

Acide Aminé	1	2	3	4	5	6	7	8
Code 1 lettre	A	R	N	D	C	E	G	I
Code 3 lettres	Ala	Arg	Asn	Asp	Cys	Glu	Gly	Ile
Acide Aminé	9	10	11	12	13	14	15	16
Code 1 lettre	L	K	F	P	S	T	V	Y
Code 3 lettres	Leu	Lys	Phe	Pro	Ser	Thr	Val	Tyr

TAB. 4.2 – Labels des acides aminés présents dans les protéines étudiées.

Seulement quatre acides aminés sont absents de la composition des protéines que nous avons étudiés (la Glutamine (notée Q), l’Histidine (H), la Méthionine (M) et le Tryptophane (W)).

A partir du tableau 4.3, on observe qu’il n’y a que deux mutations différenciant les deux crambines :

- Pour l’acide aminé 22, il s’agit de la sérine pour la crambine 1CXR et de la proline pour la crambine 1CCM.
- Pour l’acide aminé 25, il s’agit de l’isoleucine pour 1CXR et de la leucine pour 1CCM.

Afin d’affiner notre étude des crambines, nous avons créé, à l’aide du logiciel Swiss PDB Viewer [Guex et Peitsch 97] deux mélanges de crambines, qui ne diffèrent que par une seule mutation. Nous avons défini ainsi la S22/L25, cette crambine n’a donc qu’une seule mutation par rapport à 1CXR (que nous appellerons S22/I25), et la P22/I25, cette crambine n’a donc qu’une seule mutation par rapport à 1CCM (que nous appellerons P22/L25).

En analysant ces séquences, nous avons pu remarquer que la plupart des mutations entre VA3 et VB sont localisées dans l’hélice 2 et la boucle reliant les deux hélices (au niveau du motif HTH). Nous observons un résultat intéressant pour VB, dans l’hélice 2 se trouvent 3 mutations sur des résidus chargés, et on a une suite de trois résidus chargés RER (n°23, 24, 25) alors que dans VA3, on a RPT (et donc un seul résidu chargé, de même pour VA2 avec un enchaînement RQV). VA2 et VB ne diffèrent que par 3 résidus dans les hélices.

De plus, les viscotoxines sont chargées et possèdent un caractère amphiphatique. VA3 a 8 résidus chargés et 13 hydrophobes (dont 3 dans l’hélice 1 et 2 dans l’hélice 2), VB a 7 résidus chargés et 10 hydrophobes (dont 2 dans l’hélice 1 et 1 dans l’hélice 2), VA2 a 5 résidus chargés et 11 hydrophobes (dont 1 dans l’hélice 1 et 2 dans l’hélice 2).

La crambine a 4 résidus chargés et 18 hydrophobes (dont 6 dans l’hélice 1 et 3 dans l’hélice 2), elle est neutre avec un caractère hydrophobe. Alors que les résidus hydrophobes de la crambine se trouvent tout au long de ses 46 acides aminés, pour les viscotoxines ces résidus sont plus concentrés.

Les crambines diffèrent de VA3, VA2 ou VB par de nombreux résidus, la plupart se trouvant dans les hélices alpha.

Acide Aminé	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
VA3	K	S	C	C	P	N	T	T	G	R	N	I	Y	N	A	C
VB	K	S	C	C	P	N	T	T	G	R	N	I	Y	N	T	C
VA2	K	S	C	C	P	N	T	T	G	R	N	I	Y	N	T	C
Crambines	T	T	C	C	P	S	I	V	A	R	S	N	F	N	V	C

Acide Aminé	17	18	19	20	21	22	23	24	25	26	27	28	29	30
VA3	R	L	T	G	A	P	R	P	T	C	A	K	L	S
VB	R	L	G	G	G	S	R	E	R	C	A	S	L	S
VA2	R	F	G	G	G	S	R	Q	V	C	A	S	L	S
Crambines	R	L	P	G	T	P/S	E	A	L/I	C	A	T	Y	T

Acide Aminé	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46
VA3	G	C	K	I	I	S	G	S	T	C	P	S	D	Y	P	K
VB	G	C	K	I	I	S	A	S	T	C	P	S	D	Y	P	K
VA2	G	C	K	I	I	S	A	S	T	C	P	S	D	Y	P	K
Crambines	G	C	I	I	I	P	G	A	T	C	P	G	D	Y	A	N

TAB. 4.3 – Séquence des protéines étudiées : localisation des mutations et des structures secondaires avec, en bleu, les feuilletts beta et, en jaune, les hélices alpha pour la viscotoxine A3 (en effet, les structures secondaires sont à différents endroits, cf tableau 4.1).

4.3 Protocoles de simulation

4.3.1 Hydratation

Les simulations des thionines ont été réalisées avec le champ de force Charmm, version *c30b1*. Les structures de départ ont été tirées de la Protein Data Bank, puis optimisées par une méthode de recuit simulé, décrite dans le chapitre suivant.

4.3.1.1 Potentiel TIP3P

Un solvant est qualifié d'implicite s'il est représenté par un champ diélectrique continu autour du soluté. Le solvant agit comme une perturbation continue sur le comportement du soluté dans le vide. Ce modèle est nommé GB/SA (Generalized Born / Surface Area) [Stil *et al.* 90], puisqu'il est fondé sur l'approche de Born pour la contribution électrostatique. L'avantage majeur de ce modèle d'hydratation est le temps de calcul qui est 20 à 30 fois inférieur à une simulation utilisant un modèle d'hydratation explicite. Néanmoins, ce modèle n'est pas bien adapté à des solutés polaires car les liaisons hydrogène ne sont pas correctement représentées. De plus, les liaisons hydrogène intermoléculaires entre plusieurs molécules de solutés ne seront pas reproduites de façon correcte. Pour l'étude de notre système, nous avons donc utilisé une hydratation explicite.

Les modèles d'hydratation explicite sont divisés en plusieurs catégories suivant la manière de représenter les molécules d'eau : la catégorie représentant un modèle rigide et prenant

en compte les interactions coulombiennes et les expressions de Lennard-Jones, la catégorie ayant un modèle flexible, avec ajout de l'effet de polarisation.

Nous avons utilisé une hydratation explicite avec un modèle rigide de molécules de solvant de type TIP3P [Jorgensen *et al.* 83]. Ce modèle utilise 3 sites d'interactions électrostatiques (correspondant aux 3 atomes de la molécule d'eau), c'est à dire la charge partielle positive des hydrogènes et la charge négative de l'oxygène. L'interaction de Van der Walls entre deux molécules d'eau sera calculée d'après la fonction de Lennard-Jones, représentée par l'équation 4.1, avec un seul point d'interaction par molécule centré sur l'atome d'oxygène.

$$E_{ab} = \sum_i^a \sum_j^b \frac{q_i q_j e^2}{r_{ij}} + \frac{A}{r_{OO}^{12}} - \frac{B}{r_{OO}^6}. \quad (4.1)$$

Les sommes dans l'Eq. 4.1 portent sur tous les atomes i appartenant à la molécule a , et sur tous les atomes j appartenant à la molécule b .

Le tableau suivant résume les paramètres du modèle TIP3P :

Paramètres	TIP3P
r(OH), Å	0.9572
HOH, degré	104.52
A*10 ⁻³ (coeff. A de l'équation de Lennard-Jones), kcal.Å ¹² .mol ⁻¹	582.0
B (coeff. B de l'équation de Lennard-Jones), kcal.Å ⁶ .mol ⁻¹	595.0
q(O) (charge atomique)	-0.834
q(H)	0.417

TAB. 4.4 – Paramètres du modèle TIP3P.

4.3.1.2 Boîtes d'eau simulées

Le programme CHARMM permet l'utilisation du modèle d'hydratation TIP3P et d'ajuster différents paramètres tels que la taille de la boîte de solvant, le nombre de molécules de solvant, etc. Pour notre étude, nous avons utilisé 3 boîtes de simulations suivant les conditions voulues : premier cas, pour une densité de l'eau = 0.032 mol/Å³ (équivalent à la densité de l'eau liquide à 0°C), deuxième cas pour une boîte d'eau de taille supérieure au premier cas mais toujours avec la même densité de l'eau et dernier cas reproduisant les données DRX de la crambine dans un cristal [Jelsch *et al.* 00], soit une densité de 0.026 mol/Å³. Ces trois cas sont détaillés dans le tableau suivant :

	Cas 1	Cas 2	Cas 3
Dimensions (Å ³)	48 × 34 × 25	46.55 × 40.82 × 37.10	40.824 × 18.498 × 22.371
Ensemble	NPT	NPT	NVT
Nombre de molécules d'eau pour Crambine P22/L25	1085	2009	196
Densité (mol/Å ³)	0.032	0.032	0.026

TAB. 4.5 – Paramètres des boîtes d'eau simulées.

Les dimensions des boîtes d'eau simulées dans l'ensemble NPT sont données pour $t = 0$. Ces paramètres évolueront avec le temps et l'augmentation de la température. De plus, le nombre de molécules d'eau pour les deux premières simulations varie suivant la protéine étudiée. En effet, ces thionines n'ayant pas le même volume, pour conserver la densité de l'eau, le nombre de molécule d'eau doit être modifié.

4.3.2 Protocole dans le vide

A partir des structures obtenues avec le recuit simulé (cf chapitre 5), nous avons suivi le protocole suivant pour les 7 protéines :

- 1) Optimisation de la structure du recuit avec l'algorithme des gradients conjugués,
 - 2) Chauffage de $0K$ à différentes températures (5, 30, 70, 100, 130, 150, 170, 200, 220, 250, 280 et $300K$) de cette structure optimisée avec le thermostat de Nosé Hoover,
 - 3) Equilibrage pendant $600ps$ avec l'algorithme Verlet de la structure chauffée,
 - 4) Production des trajectoires sur différents temps : 50, 100, $200ps$ et $1ns$.
- Tous les paramètres utilisés sont définis dans l'annexe A.

4.3.3 Protocole dans l'eau

Les simulations des protéines hydratées ont été réalisées dans l'ensemble NPT, avec les conditions périodiques, la sommation d'Ewald et de nouveaux paramètres définis ci-dessous. De la même façon, tous les paramètres utilisés dans l'hydratation sont définis dans l'annexe B.

Algorithme Shake La méthode SHAKE est utilisée pour appliquer des contraintes de distance au système moléculaire [Ryckaert *et al.* 77]. Des contraintes sur les longueurs de liaisons ou les angles peuvent être mis sous la forme d'une contrainte de distance entre les atomes n_1 et n_2 , de la façon suivante :

$$\sigma_N(r) = r_{r_1r_2}^2 - d_{r_1r_2}^2 = 0. \quad (4.2)$$

Les deux atomes n_1 et n_2 , séparés de la distance $r_{r_1r_2}$, sont donc restreints à la distance de référence $d_{r_1r_2}$.

Nous avons appliqué SHAKE sur toutes les liaisons faisant intervenir un atome d'hydrogène pour éliminer les plus hautes fréquences de vibration présentes dans nos systèmes dans tous les programmes des protéines hydratées. Afin de nous assurer que ces contraintes ne modifient pas nos conclusions sur la dynamique des protéines hydratées, nous avons aussi simulé l'hydratation de la crambine P22/L25 sans ces contraintes. L'étude détaillée de la mobilité de la protéine nous montre que l'application ou non de l'algorithme SHAKE ne change pas la distribution de la mobilité le long de la séquence. L'écart quadratique moyen reste du même ordre de grandeur, par exemple à $300K$ le $\frac{MSD}{N} = 1.3\text{\AA}^2$ avec SHAKE et 1.1\AA^2 sans SHAKE.

Contraintes harmoniques Les contraintes harmoniques peuvent être utilisées sur une sélection d'atomes qui les forceront à se déplacer dans un puits harmonique, empêchant leurs mouvements diffusionnels. Dans notre étude, nous avons appliqué des contraintes harmoniques à la protéine jusqu'à la fin de l'étape d'équilibration (étape 6 de la simulation,

cf ci-dessous). Puis la protéine doit se relaxer une fois les contraintes harmoniques retirées. Cette période de relaxation ne doit pas être prise en compte dans l'étape de production (étape 7 de la simulation, cf ci-dessous). Les contraintes harmoniques permettent ainsi d'équilibrer le solvant liquide d'abord puis l'ensemble (eau+protéine).

Cristal et Image Afin de simuler les conditions en phase condensée, nous avons créé un cristal. Dans notre étude, nous avons utilisé deux cristaux orthorhombiques (tel que $\alpha, \beta, \gamma = 90^\circ$ et les côtés a, b et c soient de longueurs différentes) et un monoclinique (tel que $\alpha, \gamma = 90^\circ \neq \beta$ et les côtés a, b et c soient de longueurs différentes). L'ajout des termes "crystal" et "image" dans les programmes Charmm donnés en annexe B permettent de créer la symétrie voulue et une reproduction à l'infini de la protéine étudiée.

A partir des structures obtenues avec le recuit simulé, nous avons suivi le protocole suivant pour les 4 crambines :

- 1) Optimisation de la structure du recuit avec l'algorithme des gradients conjugués,
- 2) Création de la boîte d'eau aux dimensions voulues,
- 3) Insertion de la protéine au milieu de cette boîte d'eau, en éliminant toutes les molécules d'eau situées à une distance inférieure à 2.8\AA de l'atome le plus proche de la protéine,
- 4) Optimisation de la protéine hydratée avec l'algorithme des gradients conjugués,
- 5) Chauffage à différentes températures avec le thermostat de Nosé Hoover : le chauffage a été réalisé par étapes afin de préserver la stabilité du système. C'est pourquoi nous avons pris comme structure de départ la structure chauffée et équilibrée pendant $600ps$ de la température inférieure. (soit un chauffage de 0 à $5K$, de 5 à $30K$, de 30 à $70K$, de 70 à $100K$, de 100 à $130K$, de 130 à $150K$, de 150 à $170K$, de 170 à $200K$, de 200 à $220K$, de 220 à $250K$, de 250 à $280K$ et de 280 à $300K$).
- 6) Equilibrage pendant $600ps$ de la structure chauffée avec l'algorithme leap-frog de Verlet,
- 7) Production pendant $100ps$ de la structure chauffée et équilibrée avec l'algorithme leap-frog de Verlet, et SANS les contraintes harmoniques : cette étape donnera $100ps$ à la protéine pour se stabiliser à ces pressions et températures avant de produire les trajectoires qui seront analysées par la suite.
- 8) Production sur différents temps à partir de la structure précédente (stabilisée sans contraintes pendant $100ps$) : $50, 100, 200ps$ et $1ns$.

Afin de vérifier le comportement de ces protéines, plusieurs conditions initiales ont été utilisées pour créer les trajectoires (paramètre ISEED, choisi au hasard), elles sont référencées dans le tableau suivant :

	1	2	3	4	5
ISEED	314159	251008	112739	441825	160412

TAB. 4.6 – Valeurs du paramètre ISEED suivant les conditions initiales.

4.3.3.1 Ajouts des contre-ions

Dans l'étude des viscotoxines, il a fallu ajouter des contre-ions dans la boîte d'eau afin de neutraliser le système. En effet, la viscotoxine A3 est chargée +6 tandis que les viscotoxines A2 et B sont chargées +5. Nous avons alors ajouté le nombre correspondant d'atomes de Chlore (noté Cl^- et chargé -1).

Le programme Charmm consiste à choisir au hasard un atome d'oxygène dans la boîte d'eau puis à éliminer la molécule d'eau correspondante et la remplacer par un atome de Chlore autant de fois que nécessaire pour arriver à une charge nulle.

Une fois cette boîte d'eau créée, nous avons suivi le protocole suivant pour les viscotoxines hydratées (les paramètres utilisés dans les simulations sont les mêmes que ceux décrits dans l'annexe B) :

- 1) Optimisation de la structure (boîte d'eau contenant une viscotoxine et les atomes de Cl^-) avec l'algorithme des gradients conjugués,
- 2) De la même façon que précédemment, chauffage par étapes à différentes températures avec le thermostat de Nosé Hoover,
- 3) Equilibrage pendant $600ps$ de la structure chauffée avec l'algorithme leap-frog de Verlet,
- 4) Production pendant $100ps$ de la structure chauffée et équilibrée avec l'algorithme leap-frog de Verlet, et sans les contraintes harmoniques, cette étape donne de la même façon que précédemment $100ps$ à la protéine pour se stabiliser,
- 5) Production sur différents temps à partir de la structure précédente (stabilisée sans contraintes pendant $100ps$) : $50, 100, 200ps$ et $1ns$.

4.3.4 Viscotoxines avec tyrosines iodinisées

Des expériences ont prouvé que l'iodination de la tyrosine 13 conduit à une perte totale de la toxicité des viscotoxines [Stec *et al.* 04]. Une réaction d'iodination consiste à remplacer un atome d'hydrogène d'un cycle aromatique par un atome d'iode (cf Fig. 4.4). Nous avons alors construit des viscotoxines dont la tyrosine 13 est iodinisée afin de simuler les différences sur le comportement dynamique due à l'influence éventuelle de cette modification chimique.

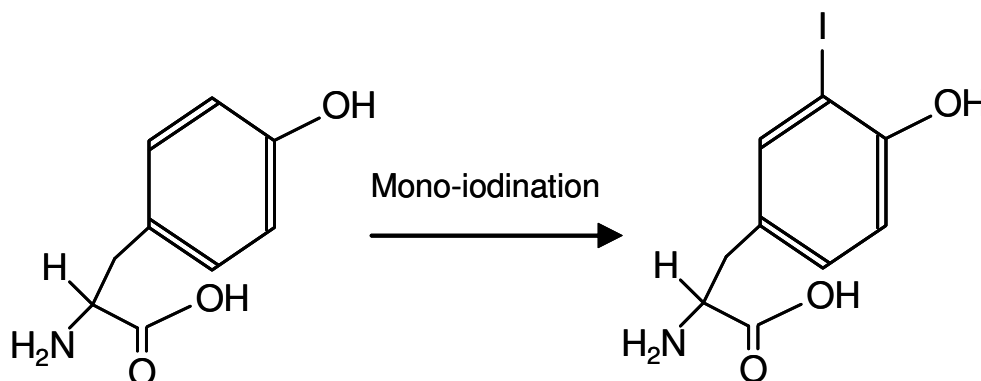


FIG. 4.4 – Réaction de mono-iodination sur une tyrosine.

Les paramètres de Charmm qui définissent tous les atomes, liaisons, angles, charges et masses n'existent pas pour l'iode. A partir des travaux de MacKerell [MacKerell], nous avons créé de nouveaux fichiers de paramètres afin de définir des tyrosines modifiées. Dans la description de la tyrosine, les atomes d'hydrogène sont référencés sous le label HGR61, les carbones alpha sous CG2R61. Nous avons alors remplacé tout ce qui décrivait une liaison CG2R61-HGR61 par une liaison CG2R61-IGR1 (où IGR1 correspond à l'iode), avec les paramètres donnés dans le tableau 4.7.

Les paramètres qui définissent le potentiel équivalent à la liaison C-I correspondent à ceux utilisés dans l'équation 1.9, ceux de l'angle C-C-I correspondent à l'équation 1.11, les angles diédraux sont décrits par l'équation 1.19 et le potentiel de Lennard-Jones par l'équation 1.16.

Masses des atomes	Mot clé	Type d'atome	Masse molaire (g)				
Masse de l'iode	Mass	IGR1	126.90447				
Liaisons	Type d'atome	Type d'atome	$K(N/m)$	$r_0(\text{Å})$			
Liaison C-I	CG2R61	IGR1	190.00	2.1150			
Angles	Type d'atome	Type d'atome	Type d'atome	$K(N/^\circ)$	$\theta_0(^\circ)$		
Angle C-C-I	CG2R61	CG2R61	IGR1	45.00	120.00		
Angles diédraux	Type d'atome	Type d'atome	Type d'atome	Type d'atome	An ($Kcal/mol$)	n	$s(^\circ)$
Angle C-C-C-I	CG2R61	CG2R61	CG2R61	IGR1	2.10	2	180.00
Angle I-C-C-H	IGR1	CG2R61	CG2R61	HGR61	3.00	2	180.00
Angle I-C-C-H2	IGR1	CG2R61	CG2R61	HGR62	3.00	2	180.00
Potentiel de Lennard Jones	Type d'atome	Valeur ignorée	ϵ ($Kcal/mol$)	$r_o(m)$			
I	IGR1	0.0	-0.55	2.190			

TAB. 4.7 – Paramètres de la tyrosine iodinisée.

4.3.5 Tests sur les protocoles de simulation

4.3.5.1 Vitesse de chauffe

Nous avons choisi de chauffer les protéines déshydratées (dans le vide) pendant $2ps$. Mais afin de nous assurer d'un temps de chauffe suffisant, nous avons réalisé des tests à $220K$ sur des temps de chauffe de 2, 4, 8 et $16ps$ sur la viscotoxine A3.

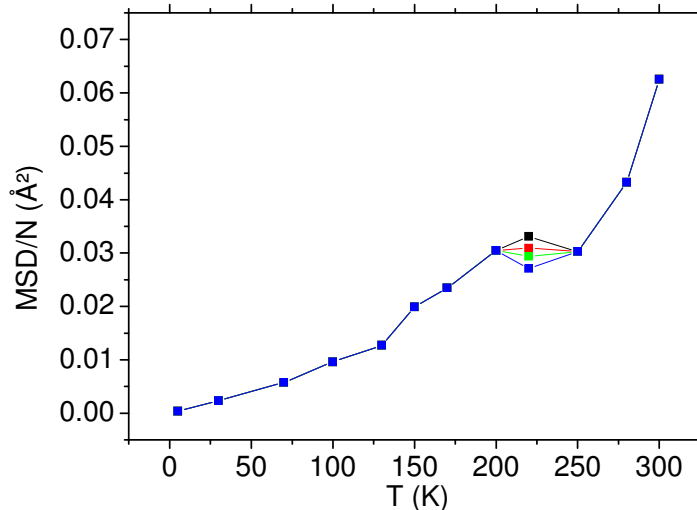


FIG. 4.5 – Tests sur le temps de chauffe entre 0 et $220K$ pour VA3 isolée : la courbe noire correspond à un temps de chauffe de $2ps$, la courbe rouge à $4ps$, la courbe verte à $8ps$ et la courbe bleue à $16ps$.

D'après la Fig. 4.5, nous pouvons constater qu'augmenter le temps de chauffe revient à diminuer la mobilité globale de la protéine. Mais de manière générale, les différences obtenues sont très faibles.

Puis nous avons calculé le RMS (Root-Mean-Square, cf Eq. 4.3) entre la structure chauffée à $220K$, pour un temps de chauffe donné, et la structure obtenue après une étape de production de $50ps$ pour ce même temps de chauffe, et ce, pour les 4 temps de chauffe étudiés. En effet, si la structure chauffée est bien stable, elle ne doit pas subir de modifications trop importantes lors de son étape de production.

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N \|r_i^A - r_i^B\|^2}, \quad (4.3)$$

où r_i^A et r_i^B représentent les positions d'un atome i dans les structures A et B , respectivement et N est le nombre d'atomes.

temps de chauffe (<i>ps</i>)	2	4	8	16
RMS (\AA)	0.1780	0.1748	0.2419	0.2176

TAB. 4.8 – Test sur les temps de chauffe pour VA3 déshydratée.

En calculant le RMS entre la structure de départ et la structure de production, nous avons pu constater qu'un temps de chauffe supérieur ou égal à $8ps$ engendre des modifications dans la structure qui sont non-négligeables. Les résultats obtenus pour un chauffage de $2ps$ ou $4ps$ sont similaires. Il est possible qu'un temps de chauffe plus long permette à la protéine d'échantillonner un autre sous-état conformationnel. La statistique de nos tests n'est pas suffisante pour conclure (en effet, une seule condition initiale a été étudiée). Les différences entre les temps de chauffe étant petites, nous avons choisi de chauffer pendant $2ps$ dans le vide et le double dans l'eau.

4.3.5.2 Chauffage par paliers

Pour les simulations dans le vide, nous avons également voulu vérifier qu'un "chauffage par paliers" n'était pas nécessaire. C'est pourquoi nous avons chauffé VA3 de 200 à 220K, pour un temps de chauffe de 2 et 4ps, et nous avons comparé ces résultats à ceux obtenus en chauffant directement de 0 à 220K, en 2 et 4ps.

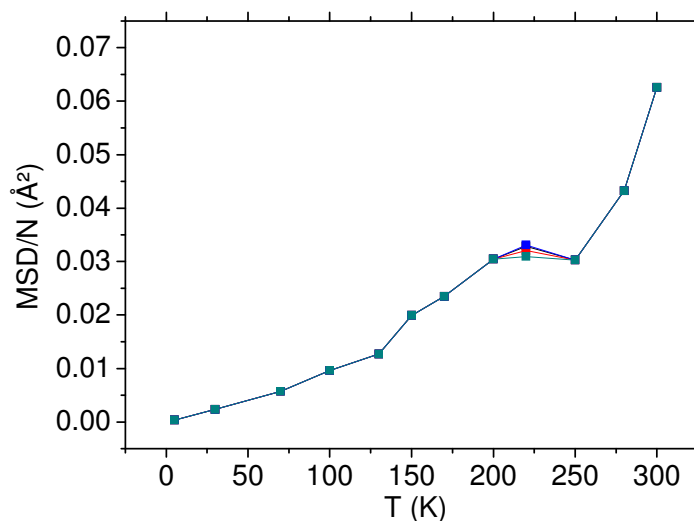


FIG. 4.6 – Comparaisons des protocoles de chauffe pour VA3 déshydratée (vide) : *chauffe de 200 à 220K* : la courbe noire correspond à un temps de chauffe de $2ps$, la courbe rouge à $4ps$; *chauffe de 0 à 220K* : la courbe bleue correspond à un temps de chauffe de $2ps$, la courbe verte à $4ps$.

Ce graphe nous montre qu'un chauffage par paliers serait inutile pour les viscotoxines dans le vide. En effet, en chauffant pendant $2ps$ à partir de $200K$ (courbe noire) ou à partir de $0K$ (courbe bleue), les résultats sont quasi-identiques. De plus, ce graphe confirme une fois

encore qu'un temps de chauffe plus long ne modifie pas de manière appréciable la flexibilité de la protéine. Nous avons calculé le RMS entre la structure chauffée ($t = 0$) et celle obtenue après une trajectoire, pour un même protocole de chauffe ($t = 50ps$, cf tableau 4.9).

temps de chauffe	0 à 220K, 2ps	0 à 220K, 4ps	200 à 220K, 2ps	200 à 220K, 4ps
RMS (\AA)	0.1780	0.1748	0.2692	0.2583

TAB. 4.9 – Test sur les temps de chauffe.

Le RMS est supérieur en chauffant plus doucement.

Nous avons également fait ce calcul pour toutes les températures comprises entre 0 et 300K pour une crambine simulée dans le vide. Nous avons ainsi comparé la mobilité globale de cette crambine obtenue avec un chauffage par paliers et un chauffage de 0K à la température voulue.

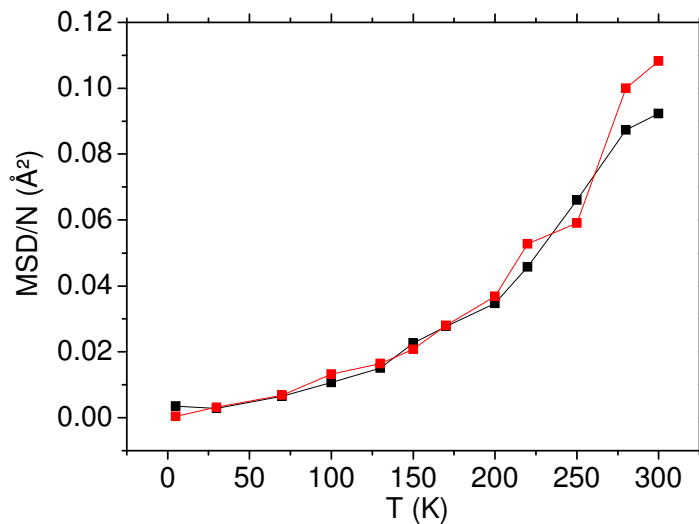


FIG. 4.7 – Comparaisons des protocoles de chauffe pour P22/L25 déshydratée. La courbe noire correspond à un chauffage par paliers et la courbe rouge à un chauffage de 0K à T_{voulue} .

Cette courbe montre clairement que, pour les crambines également, un chauffage par paliers est inutile. C'est pourquoi dans nos protocoles de simulations des thionines déshydratées, nous avons finalement chauffé directement de 0K à la température voulue en 2ps.

4.3.5.3 Ensemble NVT

Comme le montre la figure 4.8, les simulations des protéines hydratées dans l'ensemble NVT peuvent provoquer la formation de "trous" dans la boîte d'eau. Après vérification de l'algorithme, de la densité et des différents paramètres utilisés, nous avons identifié l'origine de cet artéfact. Le volume étant fixé, il est crucial que la densité de l'eau corresponde à la densité d'équilibre du modèle à la température donnée. Or le nombre de molécules d'eau

initial dans la boîte est contrôlé dans la construction de la protéine hydratée par une distance de coupure d_c . Cette distance définit un espace vide ($d < d_c$) autour de la protéine où les molécules d'eau de la phase liquide sont éliminées lors de l'insertion de la macromolécule en solution. Les trous disparaissent si cette distance est fixée à 2.8\AA et apparaissent pour $d > 2.8\text{\AA}$ où les molécules d'eau dans la boîte ne sont plus en nombre suffisant pour maintenir une phase liquide homogène. Cet artefact est éliminé automatiquement en travaillant dans l'ensemble NPT où le volume s'ajuste à la densité du liquide initial. Nous avons choisi de simuler l'hydratation dans l'ensemble NPT avec $d_c = 2.8\text{\AA}$.

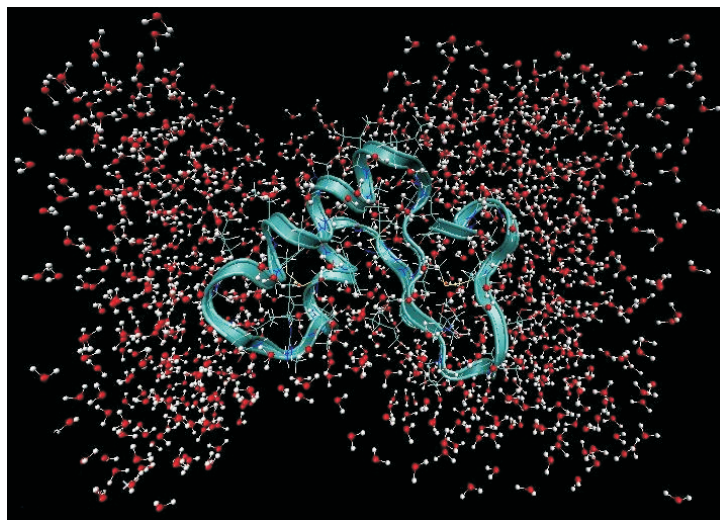


FIG. 4.8 – Formation de "trous" dans l'ensemble NVT, cas de P22/L25 hydratée.

4.3.5.4 Structure de référence

Pour calculer la variance des fluctuations structurales d'une trajectoire, il faut déterminer une structure de référence sur laquelle CHARMM va aligner l'ensemble des structures. La première possibilité consiste à prendre la structure de départ de cette trajectoire ($t = 0$), alors que la seconde consiste à prendre la structure moyenne de cette même trajectoire. Mais ce dernier choix impliquerait un calcul supplémentaire. En effet, il faudrait tout d'abord calculer la structure moyenne de la trajectoire, puis relancer ce même calcul avec comme structure de référence la structure moyenne générée.

Nous avons alors calculé le RMS de P22/L25 hydratée à $30K$ avec ces deux possibilités. Les résultats sont donnés dans le tableau suivant :

	1er cas	2ième cas
RMS 1er atome (\AA)	0.84653	0.84694
RMS dernier atome (\AA)	0.09715	0.09747

TAB. 4.10 – Test sur la structure de référence nécessaire au calcul du RMS.

Nous pouvons ainsi constater que quelque soit la structure de référence utilisée, les valeurs du RMS sont quasi-identiques. Afin de prendre le cas où les temps de calculs sont les moins élevés, nous avons choisi comme structure de référence la structure de départ ($t = 0$).

Structures optimisées - Comparaisons à l'expérience

Les protéines sont des molécules dont les fonctions biologiques dérivent de leur structure tertiaire et de leur flexibilité. La structure tertiaire d'une protéine correspond à la conformation de la plus faible énergie libre dans les conditions du milieu. Afin de valider le champ de force, nous avons simulé les structures d'équilibre des thionines à $T = 298K$ par une méthode de recuit simulé. Ensuite, nous avons optimisé les structures obtenues ("structures à 0K"). Les structures obtenues seront comparées aux fichiers expérimentaux bruts de la Protein Data Bank (données d'entrée de nos calculs ; pour les fichiers PDB obtenus par RMN et qui contiennent plusieurs modèles, nos fichiers de départ sont la moyenne de ces modèles).

5.1 Calculs des structures de départ

5.1.1 Simulation de recuit et Calibrage

5.1.1.1 Principe de la simulation de recuit

La méthode de Simulated Annealing (simulation de recuit) est souvent utilisée comme une méthode d'optimisation pour prédire la structure tertiaire des protéines en minimisant leur énergie [Yoshida *et al.* 02]. Le terme 'annealing' dérive d'un procédé physique qui consiste à chauffer un système puis à le refroidir. 'Simulated Annealing' est une méthode d'optimisation pour simuler ce procédé physique par ordinateur. Il est basé sur un algorithme de type Monte Carlo. Nous l'appliquerons sur les structures dans le vide à 298K. On considère trois étapes :

- La première modifie une structure x et génère une autre structure x' .
- L'acceptation ou non de la structure x' est déterminée par la différence d'énergie potentielle $\Delta E (= E' - E)$, avec $E = f(x)$ et $E' = f(x')$:
 - Si $\Delta E \leq 0$, la modification est acceptée, x est remplacée par x' ,
 - Si $\Delta E > 0$, la modification est acceptée avec la probabilité :

$$P_{accept} = \begin{cases} 1 & \text{si } \Delta E \leq 0 \\ \exp\left(-\frac{\Delta E}{k_B T}\right) & \text{sinon.} \end{cases}$$

La température T est un paramètre de contrôle. Refroidir consiste à passer de la température T_{k+1} à la température T_k .

La simulation commence donc avec une solution initiale x avec une énergie E , et une température initiale T . La solution est modifiée en x' , avec une énergie E' . L'acceptation de cette modification est calculée par la moyenne RMS à la température T_k . Si elle est acceptée, la solution x' devient le point de départ de la prochaine étape. Ces opérations sont répétées à chaque température pour le système jusqu'à atteindre l'état final ou l'équilibre.

La première étape consiste à calibrer le système pour déterminer une température maximale (théorique) de chauffage de la protéine, soit une température où le paramètre RMS, défini par l'Eq. 4.3, reste inférieur à 1\AA (ce critère sera justifié dans la partie 5.1.2). Ensuite, on refroidit lentement, par paliers.

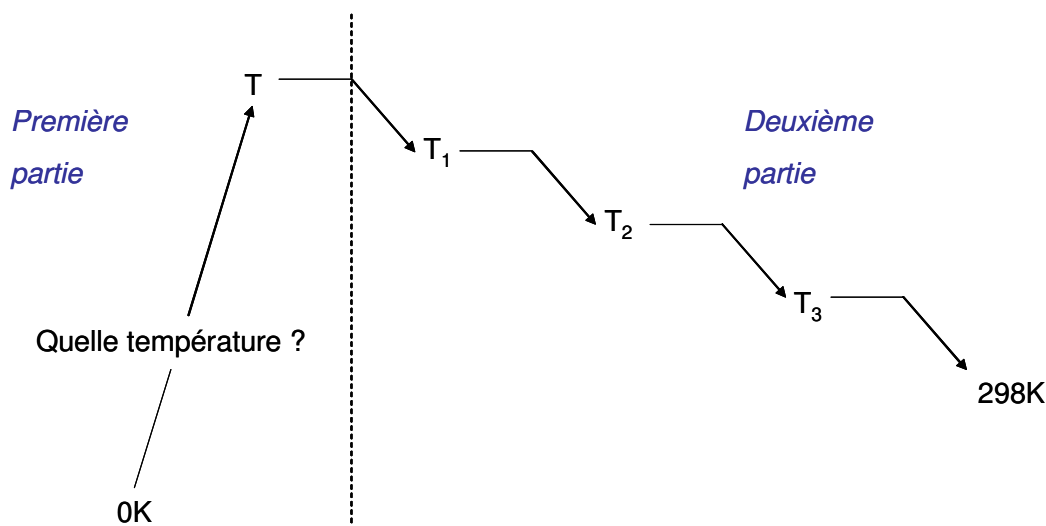


FIG. 5.1 – Principe de la simulation de recuit.

5.1.1.2 Calibrage

Nous avons tout d'abord calibré la température maximale de chauffage. Une façon d'évaluer la modification de structure provoquée par le chauffage, et dans une étape suivante, le refroidissement, consiste à calculer la moyenne entre la structure initiale A, et la structure obtenue B, nous obtiendrons alors le paramètre RMS (Root Mean Square) défini par l'équation 4.3.

Si $0 < RMS < 1\text{\AA}$, la nouvelle structure est acceptée. Si par contre, $RMS > 1\text{\AA}$, la nouvelle structure présente de nombreuses différences, acceptables avec certaines probabilités. Nous utiliserons ce paramètre RMS dans la suite des calculs.

La détermination de la température de chauffage est une étape essentielle puisque si l'on chauffe trop, nous nous écarterons de résultats ayant un sens physique. Dans la nature, la protéine subit le phénomène physique de dénaturation à la température de fusion de la molécule T_m . Dans notre cas, la molécule, à cause des ponts disulfures, ne peut se dénaturer.

Après avoir chauffé la protéine, nous abaisserons doucement la température jusqu'à 298K par paliers. A chaque palier, nous pourrons faire évoluer librement le système équilibré à la température T (on obtiendra ainsi la trajectoire) puis calculer l'énergie moyenne du système.

Pour chaque température d'équilibration de cette simulation, nous obtiendrons donc trois résultats :

- L'énergie moyenne :

$$\bar{E} = \frac{E_1 + E_2 + \dots + E_M}{M}, \quad (5.1)$$

où E_1, E_2, \dots, E_M représentent les énergies obtenues lors de la trajectoire et M le nombre d'étapes dans cette trajectoire.

- Les fluctuations de la trajectoire :

$$F = \left(\frac{\sum (E_i - \bar{E})^2}{M} \right)^{\frac{1}{2}}, \quad (5.2)$$

où E_i est l'énergie à l'étape i , \bar{E} est l'énergie moyenne calculée précédemment et M le nombre d'étapes.

- Le ratio :

$$R = \frac{F}{(\bar{E}^2)^{\frac{1}{2}}}. \quad (5.3)$$

Plus les fluctuations seront faibles, plus notre système sera stable. En plus du calcul du RMS, nous vérifierons le ratio et nous considérerons le système stable si $10^{-3} < R < 10^{-4}$. Nous avons choisi ce niveau de précision pour nous garantir la stabilité de notre système [Aparicio 05].

Lors du calibrage de la température de chauffe, nous avons suivi ce même protocole pour chacune des températures étudiées :

- 1) Optimisation du système avec l'algorithme des gradients conjugués,
- 2) Equilibrage du système pendant 600 ps, avec l'algorithme Verlet,
- 3) Production du fichier PDB de la nouvelle structure pendant 20ps, avec l'algorithme Verlet,
- 4) Calcul de la moyenne entre la structure expérimentale et la structure générée (RMS).

A partir du fichier PDB expérimental de la VA3, nous avons donc calibré la température de chauffage pour laquelle $10^{-3} < R < 10^{-4}$ et $RMS < 1\text{\AA}$. Si le système est stable, l'énergie doit être négative. Après différents tests, nous avons décidé de chauffer les protéines jusqu'à 600K. Les résultats sont donnés dans l'annexe C.

5.1.2 Résultats

Après avoir chauffé jusqu'à 600K, la simulation consiste à refroidir doucement le système jusqu'à la température utilisée expérimentalement, à savoir 298K. En l'équilibrant correctement à chaque palier, nous obtenons ainsi la structure d'équilibre.

Nous pouvons alors superposer la structure équilibrée à 298K à la structure expérimentale (réalisée à 300K) et à la structure optimisée (cf figure 5.2) :

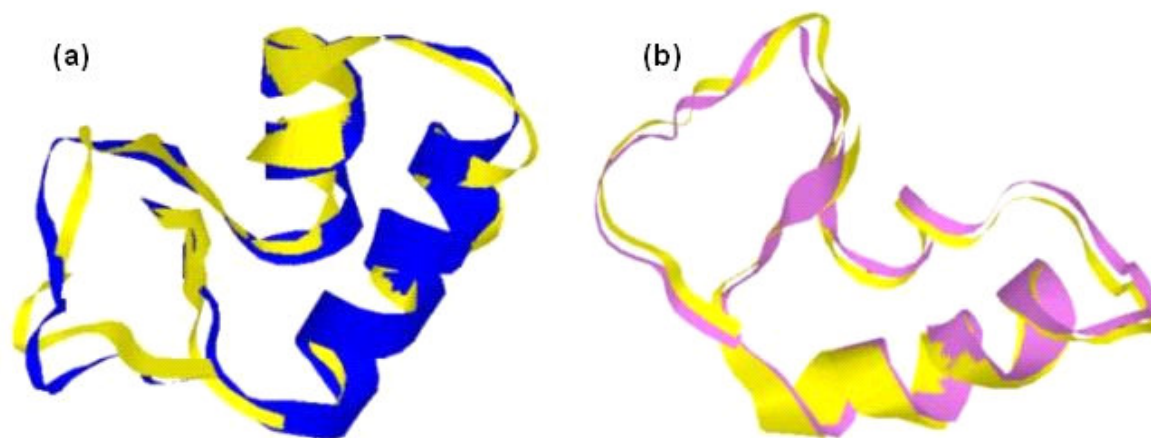


FIG. 5.2 – Recuit simulé de VA3 (représentation en ruban des squelettes) : (a) Superposition de la structure finale obtenue à 298K (en jaune) à la structure optimisée (en bleu). (b) Superposition de la structure finale obtenue à 298K (en jaune) à la structure expérimentale (en violet).

Ces superpositions (à gauche) nous donnent des informations sur les parties où l'entropie contribue le plus, correspondant alors aux parties les plus mobiles. Nous pouvons remarquer que les ponts disulfures maintiennent la position des hélices alpha. La comparaison (à droite) avec la structure expérimentale, hydratée, nous montre indirectement que l'effet du solvant (l'eau) sera plus important dans les régions mobiles des feuillets beta et de la boucle. De manière générale, ces différentes structures se superposent avec de très grandes similitudes. Les superpositions des structures à 298K aux structures expérimentales nous indiquent une grande similitude de structure, nous sommes capables de reproduire la réalité. Nous pouvons donc valider le champ de force utilisé.

Nous nous sommes intéressés à quatre structures de la protéine crambine afin d'étudier l'effet des mutations le long de la séquence primaire. Cependant le fichier référencé sous le code 1CXR nous paraissait incomplet. Il a donc fallu le reconstruire en corrigeant les labels des atomes, la position des hydrogènes et générer un nouveau fichier afin d'obtenir un fichier PDB correct. Ensuite, nous avons suivi le même protocole pour la simulation. Pour S22/I25, nous pouvons remarquer que le paramètre RMS (donné dans le tableau de l'annexe C) est toujours supérieur à 1Å, mais très légèrement, impliquant donc de petites différences dans la structure obtenue. Le ratio est compris dans les limites imposées pour les faibles températures uniquement. La reconstruction du fichier initial nous a donné un fichier PDB correct qui sera utilisé dans la suite des simulations.

Pour les 3 autres crambines, les critères sont respectés.

Le fichier PDB de la crambine P22/L25 contient plusieurs modèles. Afin d'avoir un ordre de grandeur sur le RMS obtenu expérimentalement entre deux structures, nous avons calculé le RMS entre différents modèles du fichier PDB. Par exemple, certains de ces résultats sont répertoriés dans le tableau suivant :

	modèles 1 et 3	modèles 2 et 4	modèles 2 et 8	modèles 3 et 7	modèles 5 et 6
RMS(\AA)	1.613699	0.973330	1.018200	1.171108	1.444182

TABLE 5.1 – Calculs du RMS entre différents modèles du fichier pdb expérimental de la crambine P22/L25.

Nous pouvons remarquer que les valeurs obtenues sont du même ordre de grandeur que le critère fixé dans la méthode de recuit simulé. Le RMS entre les différents modèles de la crambine P22/L25 est proche de 1\AA . Un RMS de 1\AA correspond donc à l'ordre de grandeur de la "distance" séparant deux conformations de la crambine.

La méthode de recuit simulé donne une structure stable et ainsi les différentes structures équilibrées qui nous serviront de point de départ pour les simulations suivantes.

5.2 Analyses des structures équilibrées

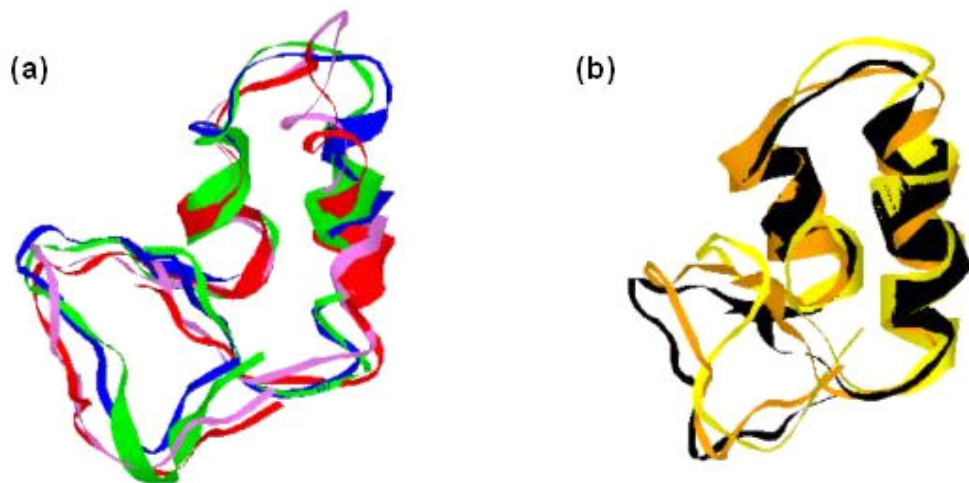


FIG. 5.3 – (a) Superposition des crambines : P22/L25 en bleu, S22/I25 en rouge, P22/I25 en vert et S22/L25 en violet. (b) Superposition des viscotoxines : VA2 en noir, VA3 en orange et VB en jaune.

En superposant les structures obtenues par la méthode de recuit simulé à l'aide du logiciel Rasmol [Sayle et Milner-White 95], on remarque une grande similitude de structure secondaire, malgré les mutations d'acides aminés entre ces différentes protéines décrites précédemment.

Sur la figure 5.3, nous pouvons déjà remarquer que la crambine S22/I25, en rouge, et S22/L25, en violet, sont très proches. Ces deux crambines ont le même résidu en 22 et une mutation en 25. De la même façon, P22/L25, en bleu, et P22/I25, en vert, se superposent avec plus de similitudes, tout en sachant que ces protéines présentent également le même résidu en 22 et une mutation en 25. On peut alors supposer que le résidu situé en 22 "contrôle" la structure dans l'espace. De plus nous pouvons supposer que ce 22^{ième} résidu joue un rôle clé dans la flexibilité de la protéine puisque, d'après la Fig. 4.3, il est situé dans la boucle

reliant les deux hélices alpha (et donc dans le motif HTH), une zone très mobile comme nous le montrerons plus tard.

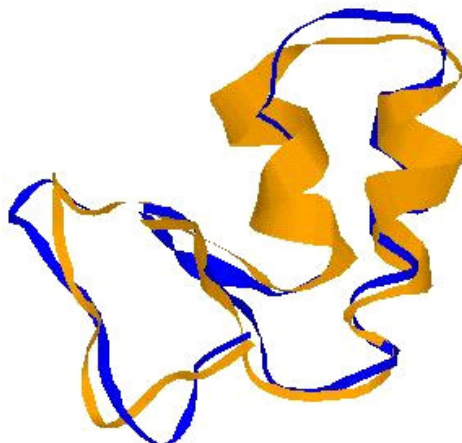


FIG. 5.4 – Superposition de la viscotoxine A3 (en orange) et de la crambine P22/L25 (en bleu).

Sur la figure 5.4, nous avons superposé une crambine (*P22/L25*) à une viscotoxine (*A3*) afin de montrer les grandes similitudes entre ces thionines malgré leur différence d'activité biologique. De manière générale, aucune modification dans la structure secondaire n'intervient entre ces protéines, ce qui est en accord avec les résultats de la littérature [Romagnoli *et al.* 00].

5.3 Comparaisons avec les techniques de RMN et de diffraction X

Les fichiers RMN sont composés de plusieurs modèles, ils ont été obtenus à $300K$. Nous avons calculé les fluctuations entre chacun de ces modèles pour la crambine P22/L25 et comparé à nos calculs de dynamique moléculaire à $300K$. La Fig. 5.5 représente le MSD des atomes lourds calculé avec Charmm à gauche et des fichiers RMN à droite. Nous constatons que le MSD obtenu par RMN est 20 fois plus faible que celui calculé par dynamique moléculaire. Des différences interviennent entre ces deux graphes mais l'étude détaillée de la localisation des atomes mobiles montrent que les mêmes atomes ont un MSD élevé, même si leur ordre de grandeur est modifié. Par exemple, le CH_3 du résidu THR1 a un MSD d'environ 1.8\AA^2 avec Charmm (soit environ $1/3$ du MSD_{max}) et il a un MSD de 0.1\AA^2 par RMN (soit également $1/3$ du MSD_{max}). La RMN met en évidence la très forte mobilité de la boucle.

Les fichiers DRX déposés sur la Protein Data Bank sont réalisés à $150K$, nous avons calculé à partir des B facteurs le RMS des carbones alpha de la crambine P22/L25 (fichier PDB 1CNR) et comparé à nos calculs à cette même température (cf Fig. 5.6). Les résultats obtenus avec Charmm (à gauche) donnent des valeurs du MSD 50 fois plus faibles que celles de DRX (à droite). A $150K$, nous sommes en-dessous de la température de transition dynamique, c'est pourquoi les valeurs calculées par dynamique moléculaire sont aussi faibles.

La différence entre l'amplitude des fluctuations calculées et celles mesurées en DRX pourrait provenir du désordre statique, ignoré dans nos calculs.

La localisation des sites mobiles est globalement la même dans nos simulations et dans les expériences. Le carbone alpha du résidu 37 a un MSD de 6.8\AA^2 en DRX, soit une valeur proche de la mobilité maximale ($=7.0\text{\AA}^2$), et il a un MSD de 0.09\AA^2 avec Charmm, également proche du MSD_{max} ($= 0.10\text{\AA}^2$).

Des différences existent donc sur l'étude de la flexibilité d'une protéine suivant la technique employée. L'analyse détaillée permet de constater des modifications mais de manière générale, les sites de grandes mobilités restent les mêmes.

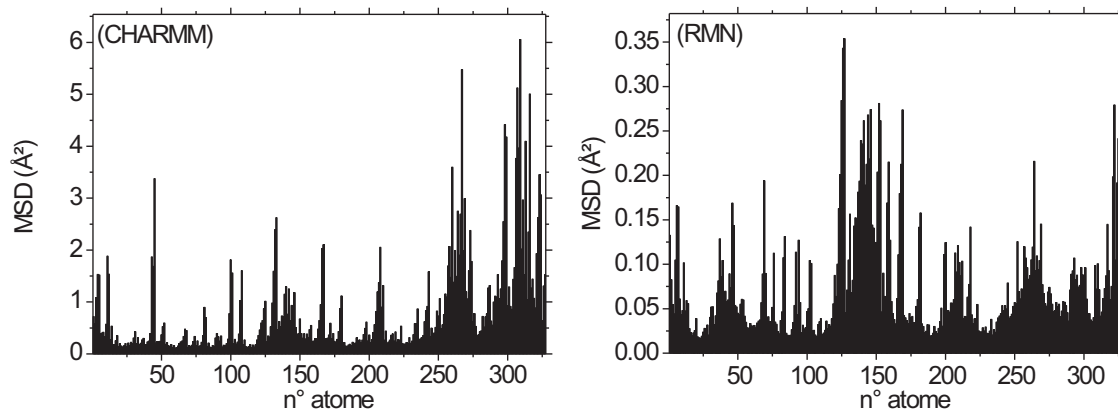


FIG. 5.5 – MSD des atomes lourds calculés pour P22/L25 hydratée à 300K par dynamique moléculaire à gauche et à partir des modèles RMN à droite.

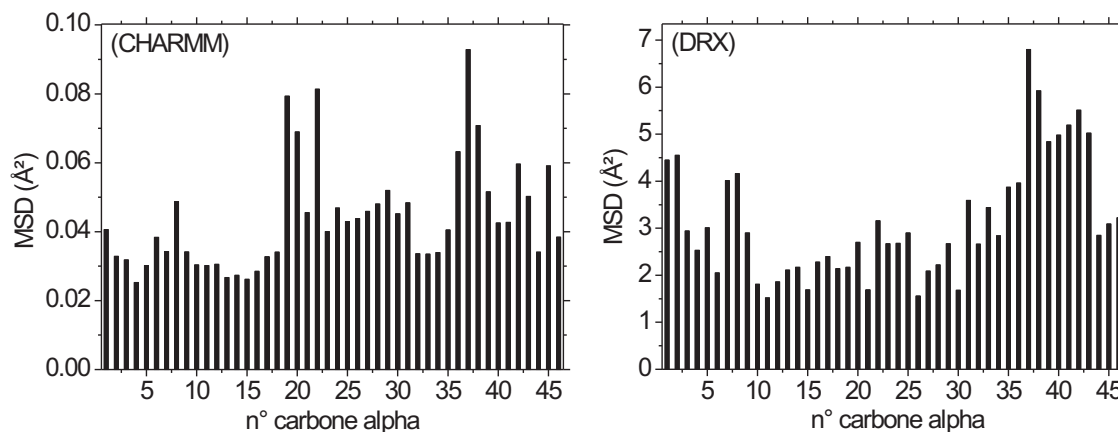


FIG. 5.6 – MSD des carbones alpha calculés pour P22/L25 hydratée à 150K par dynamique moléculaire à gauche et à partir des structures DRX à droite.

DEUXIÈME PARTIE

**Dynamique conformationnelle et
transition dynamique**

Introduction

1.1 La transition dynamique des protéines

Les protéines sont des polymères constitués d'un enchaînement d'acides aminés liés par des liaisons covalentes (cette séquence d'acides aminés est appelée structure primaire comme nous avons pu le voir précédemment). La séquence des acides aminés est codée par le génome afin d'assurer une fonction biologique. Chaque protéine possède un état thermodynamiquement stable, appelé « état replié » ou « état natif », où elle est fonctionnelle. Une protéine qui n'est pas dans son état replié est dite dénaturée et ne peut plus exercer sa fonction. La protéine assurera donc sa fonction en adoptant, le plus souvent spontanément, une structure organisée spécifique en milieux aqueux (sa structure native).

L'étude expérimentale des structures tertiaires des protéines débute en 1960 et représente un intense champ de recherches [Kendrew *et al.* 60]. Au 1^{er} Mars 2008, on dénombrait plus de 49295 structures connues dans les bases de données PDB. Décrire la structure d'une protéine est cependant insuffisant pour décrire complètement sa fonction : il faut comprendre comment la protéine fluctue autour de son minimum d'énergie libre et comment ses conformations peuvent être perturbées. Des changements de conformations ou des modifications de dynamique contribuent à des pathologies comme la maladie d'Alzheimer. La flexibilité et la dynamique d'une protéine sont des caractéristiques essentielles pour sa fonction. Pour que la protéine adopte sa structure active, des réarrangements dans l'espace sont nécessaires. Par exemple, lors de l'utilisation d'une enzyme, la liaison d'un substrat modifie la conformation de l'enzyme de sorte que la protéine et le substrat puissent s'ajuster au mieux [Rigdena *et al.* 06]. C'est également le cas pour les protéines à transfert de charge dont l'activité dépend de leurs fluctuations structurales. La flexibilité des protéines peut être quantifiée en observant l'évolution de ses fluctuations structurales en fonction de la température. Ces fluctuations peuvent être mesurées notamment par DRX, par RMN et par diffusion incohérente de neutrons (cf chapitre 3, partie I).

La dynamique des protéines est complexe, elle dépend à la fois de la macromolécule et du solvant : les forces qui maintiennent la structure native et celles des liaisons hydrogène entre les molécules d'eau sont en effet du même ordre de grandeur. Le moyen le plus simple de sonder le comportement dynamique d'un tel système est de contrôler son énergie cinétique en faisant varier sa température. C'est ainsi que Fritz Parak et ses collègues du département de physique de l'Université de Münschen et de l'Institut de Physique-Chimie de Moscou observèrent pour la première fois en 1981 la transition de phase dynamique des protéines [Parak *et al.* 81].

Leur expérience consista à mesurer les déplacements de l'atome de Fer de la metmyoglobine en utilisant la spectroscopie de Mössbauer. Ils s'attendaient à ce que le déplacement quadratique moyen de l'atome augmente de façon linéaire avec la température. Mais ils ont observé ce comportement linéaire seulement à basses températures et une augmentation non-linéaire avec la température au-delà d'environ 200-220K [figure 1.1]. Cette transition, dite transition dynamique, ne correspondait à aucun modèle étudié. Leurs résultats confirment des travaux réalisés plus tôt sur la dépendance en température de la myoglobine par Frauenfelder et ses coauteurs en 1975 [Austin *et al.* 75]. Depuis ces premiers travaux, la transition dynamique a été observée par diffusion de neutrons [Doster *et al.* 89, Gabel et Bellissent-Funel 07], par RMN (Résonance Magnétique Nucléaire) [Lee et Wand 01], et par diffraction X [Zanotti *et al.* 07, Rasmussen *et al.* 92, Teeter *et al.* 01] pour plusieurs protéines (la myoglobine, la C-phycoyanine, la calmoduline, la ribonucléase, la crambine,

le lysosyme et la bactériorhodopsine respectivement [Zanotti *et al.* 07, Miyazaki *et al.* 00, Ferrand *et al.* 93]).

La transition dynamique des protéines a également été décrite comme une transition de type verre pour souligner certaines similitudes entre les changements de comportement dynamique des protéines et les changements de viscosité et autres propriétés dynamiques des liquides lorsqu'ils deviennent vitreux [Bicout et Zaccai 01]. Cette dynamique est déterminée par l'énergie libre de la protéine, une fonction décrivant comment l'énergie de la protéine varie lorsque la structure et/ou la température changent [Zanotti *et al.* 07]. L'énergie libre possède plusieurs minima [Frauenfelder et McMahon 98] comme celle des verres, représentant chacun des sous états conformationnels. Les fluctuations des positions des atomes des protéines en solution provoquent des sauts entre ces différents minima et ont pour conséquences des changements de conformation qui peuvent ainsi faire disparaître la fonction biologique [Ferrand *et al.* 93]. Cette transition marque ainsi le passage d'un régime harmonique à un régime anharmonique avec une composante diffusionnelle [Ringe et Petsko 03].

Schématiquement, la transition peut être décrite par un modèle en double puits : à basses températures, la protéine est piégée dans un minimum d'énergie libre et le mouvement est harmonique ; à partir d'une certaine température, la protéine peut franchir une barrière effective d'énergie (nous serons alors dans le cas d'un mouvement anharmonique) et changer de conformation, les sauts entre minima d'énergie s'ajoutent aux fluctuations structurales autour de ces minima.

La transition a été décrite par des modèles simplifiés, que nous développerons dans un prochain chapitre. Ces modèles simples permettent de définir des paramètres de la transition (différences d'énergie libre et/ou constantes effectives) mais n'en décrivent pas l'origine microscopique.

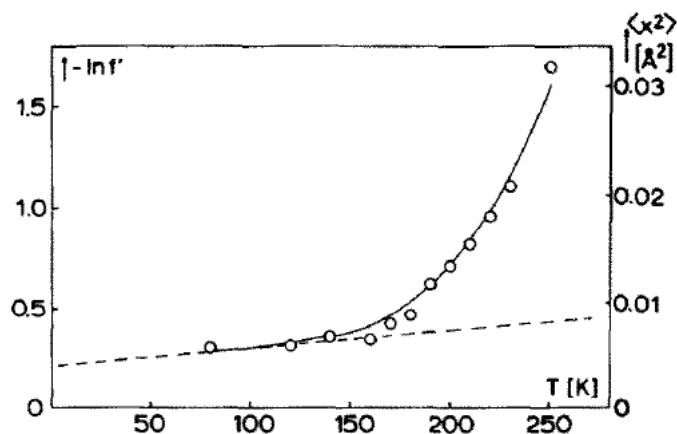


FIG. 1.1 – Courbe typique représentant la transition dynamique, obtenue dans les premiers travaux de Parak [Parak *et al.* 80] où $\langle x^2 \rangle$ est le déplacement quadratique moyen de l'atome de Fer du groupe heme de la metmyoglobine.

L'origine microscopique de cette transition divise actuellement les experts. Plusieurs observations et hypothèses ont été reportées dans la littérature [Kumar *et al.* 06] :

1) Les travaux en RMN sur la calmoduline de Wand indiquèrent que l'eau ne jouait pas de rôle clé dans la transition dynamique, tout en mettant en avant la contribution des méthyles [Lee et Wand 01].

2) D'autre part, plusieurs études, sur l'ARN en diffusion de neutrons [Caliskan *et al.* 06], sur la ribonucléase A à l'aide de simulations de dynamique moléculaire [Curtis *et al.* 04] ou encore avec l'étude expérimentale et théorique de la xynalase [Reat *et al.* 00], prouvent que la transition est causée par le solvant mais le mécanisme conduisant à cette transition reste une question ouverte.

3) La protéine est esclave de l'eau et les fluctuations de l'eau augmentant avec la température entraînent la transition dynamique de la protéine [Weik *et al.* 01].

4) Des travaux en neutrons et spectroscopie de Mössbauer sur différentes protéines démontrèrent également que l'eau est un plastifiant qui diminue la viscosité des chaînes latérales [Doster et Settles 99].

5) La mesure expérimentale de l'écart quadratique moyen par neutrons et spectroscopie Mössbauer de la myoglobine par Fenimore démontre que la transition est couplée à l'hydratation [Fenimore *et al.* 04, Fenimore *et al.* 05].

6) La transition est liée à une transition eau liquide-eau, comme le montrent des simulations de dynamique moléculaire du lysosyme et de l'ADN [Kumar *et al.* 06].

7) Le réseau de liaisons hydrogène est étroitement lié à la transition dynamique puisque les travaux de Teeter sur la protéine crambine montre que le réseau change de topologie (i.e, le nombre de cycles à 5 molécules d'eau et de cycles à 7 molécules d'eau varie avec l'augmentation de la température) [Teeter et Whitlow 87]. La réorganisation du réseau de liaisons hydrogène et la transition eau liquide-eau liquide sont étroitement reliées. L'eau à la surface des protéines est une eau "interfaciale" dont le réseau de liaisons hydrogène est modifié par rapport à l'eau de volume. Les travaux de Zanotti et de ses collaborateurs en 2007 montrent que l'eau interfaciale (à l'interface de silicate hydrophile) a une transition de phase vitreuse à 165K et une transition d'eau liquide de faible densité à de l'eau liquide de haute densité vers 240K. Vers 220K, la transition eau liquide-eau liquide s'accompagne d'une augmentation significative des fluctuations rotationnelles de l'eau à la surface des silicates. Cette évolution des fluctuations structurales de l'eau avec la température suit celle observée pour les mouvements internes de la protéine suggérant un fort couplage entre l'eau d'hydratation et les mouvements de la protéine [Zanotti *et al.* 07, Zanotti *et al.* 99].

8) Plusieurs études montrent que la transition dynamique n'existent pas pour des protéines déshydratées, comme par exemple l'étude avec Charmm du lysozyme [Roh *et al.* 05] ou encore lors des travaux de Fenimore développés précédemment.

Pour résumer, voici la liste des protéines pour lesquelles la transition dynamique a été étudiée, que ce soit expérimentalement ou théoriquement : la ribonucléase (124 résidus) [Rasmussen *et al.* 92], la myoglobine (153 résidus) [Doster *et al.* 89], le lysozyme (129 résidus) [Zanotti *et al.* 07], la crambine (46 résidus) [Teeter *et al.* 01], la calmoduline (148 résidus) [Lee et Wand 01] et la bactériorhodopsine [Ferrand *et al.* 93]. Un phénomène similaire a été aussi observé pour l'ADN [Kumar *et al.* 06] et l'ARN [Caliskan *et al.* 06]. Ces observations indiquent un rôle certain de l'eau mais les mécanismes microscopiques doivent être probablement différents pour les acides nucléiques et les protéines.

La transition dynamique de petites protéines globulaires telle que la crambine a été peu étudiée. La transition dynamique de toxines fortement chargées telles que les viscotoxines étudiées dans notre thèse est abordée ici pour la première fois.

Les questions centrales abordées dans cette partie sont :

- Quel est le rôle de l'eau et plus généralement du solvant dans la transition dynamique des protéines et plus particulièrement de petites protéines globulaires ? Comment caractériser ce rôle de l'eau ?

- Les courbes de transition dynamique (cf Fig. 1.1) nous donnent une information globale sur la flexibilité de la protéine. Mais comment la transition se produit-elle le long de la séquence ? Les fluctuations sont-elles distribuées de façon homogène ou inhomogène ? La transition affecte-t-elle tous les résidus et les atomes de la même manière ? Une façon d'étudier ce problème est de mesurer les fluctuations de manière locale (expérimentalement par DRX, cette technique présente tout de même un inconvénient puisque la précision sur les B-facteurs n'est pas toujours très bonne [Tilton *et al.* 92], ou par RMN [Lee *et al.* 00], et théoriquement par simulations de dynamique moléculaire [Curtis *et al.* 04]).

- Une des autres motivations importantes de cette partie de notre travail est de comprendre comment les mutations ponctuelles peuvent engendrer des modifications de dynamique. Plusieurs paragraphes sont consacrés à l'étude de la mobilité locale en fonction de la température, nous examinons notamment le rôle des groupes méthyles, l'évolution et l'apparition d'un certain désordre dynamique et la manière dont la mobilité varie le long de la séquence afin de répondre aux questions suivantes : Peut-on utiliser les méthyles comme sonde ? Quel est l'effet d'une mutation sur la transition dynamique ? Expérimentalement, les méthyles sont d'ores et déjà utilisés comme des sondes des changements de dynamique d'une molécule. Les travaux en RMN de Lee sur la calmoduline en 2000 montrent que l'analyse de la dynamique des chaînes latérales, tout particulièrement des méthyles, permet de donner des informations sur les modifications de la dynamique globale d'une molécule [Lee *et al.* 00]. Ceci a été également confirmé en 2003 par des travaux en RMN de Millet sur l'effet des mutations, où les mouvements des méthyles sondent la dynamique globale [Millet *et al.* 03]. D'un point de vue expérimental, étudier la dynamique des méthyles par RMN permet de détecter les changements de dynamique de la protéine.

Comme nous l'avons vu précédemment, les protéines étudiées appartiennent à la famille des thionines. Nous avons choisi cette famille puisqu'il existe des données expérimentales sur les crambines (ceci nous permettra ainsi de confirmer notre étude théorique) mais ce choix est également motivé par des intérêts thérapeutiques. En effet, les viscotoxines, appartenant à cette famille, présentent une toxicité vis-à-vis des membranes cellulaires et donc des applications potentielles dans le traitement du cancer et comme antibiotiques [Romagnoli *et al.* 00]. L'iodination de la tyrosine 13 conduit à une perte totale de la toxicité des viscotoxines [Stec *et al.* 04], d'où notre intérêt dans l'étude de modifications chimiques de ces protéines par l'iodination. Cet aspect des relations dynamique-cytotoxicité sera abordé dans la troisième partie.

1.2 Hydratation des protéines

La relation importante entre l'environnement moléculaire d'une protéine et sa dynamique interne a été soulignée par plusieurs expériences, comme nous l'avons vu au paragraphe précédent. En effet, lors de la transition dynamique des protéines, la modification des interactions eau-protéine ont des conséquences importantes dans ce phénomène. Il s'agit alors d'évaluer quelle est la part due à la protéine elle-même et celle due aux interactions eau-protéine [Gregory 94].

Des tests ont ainsi été réalisés sur le taux d'hydratation et plusieurs conclusions ont été tirées :

1) Un taux minimal de 0.4 gramme d'eau par gramme de protéine est nécessaire pour déclencher la transition dynamique des protéines [Orecchini et Paciaroni 04]. Expérimentalement, cette étude est possible avec des cristaux, où le taux pourra varier entre 0 et 1 [Paciaroni *et al.* 02].

2) Pour un taux $h = 0$, la transition dynamique semble disparaître à la fois pour les protéines [Parak 86], l'ADN [Kumar *et al.* 06] et l'ARN [Caliskan *et al.* 06]. Le taux d'hydratation est également important pour la fonction biologique. En effet, les enzymes commencent à montrer une activité catalytique pour $h \simeq 0.12 - 0.2$ [Paulaitis *et al.* 92]. Les travaux de Rupley sur le lysozyme montre que l'activité enzymatique augmente avec le taux d'hydratation [Rupley *et al.* 83].

3) La température de dénaturation de la molécule dépend également du taux d'hydratation comme le montrent les travaux de Fujita sur la lysozyme (cf Fig. 1.2), h modifie fortement la stabilité des protéines.

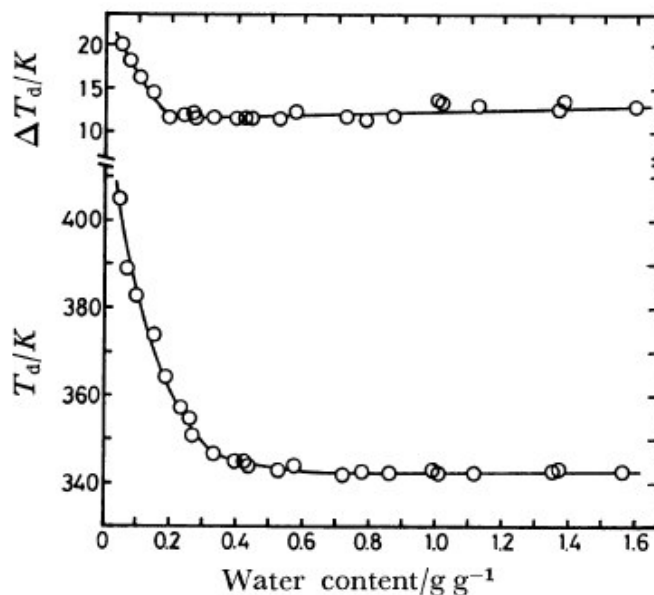


FIG. 1.2 – Variation de la température de dénaturation T_d et de la largeur des pics de la température de dénaturation ΔT_d du lysozyme en fonction du taux d'hydratation [Fujita et Noda 78].

4) Pour confirmer le rôle de l'eau dans la transition dynamique, Alexander L. Tournier et son équipe réalisèrent en 2005 des simulations par dynamique moléculaire de la myoglobine où la température du solvant était fixée à une température inférieure à celle de la protéine. Dans cette construction théorique, la transition dynamique de la protéine disparaît car les changements conformationnels de la protéine ne sont plus activés par le solvant [Tournier *et al.* 03].

5) Les propriétés du solvant diffèrent suivant qu'il s'agit de l'eau d'hydratation de la protéine, (en contact avec la molécule) ou des couches d'eau plus éloignées dans le "volume" [Gabel et Bellissent-Funel 07]. Des tests ont prouvé que ces propriétés du "volume" existaient pour $h > 1.4$. Mais certains tissus et autres organismes présentent ces propriétés du "volume" pour un taux d'hydratation inférieur [Prestrelski *et al.* 93].

6) Le taux d'hydratation modifie également la température de transition dynamique : plus le taux h sera élevé, plus la température de transition diminuera [Gregory et Chai 93, Pissis 89].

Notre objectif est d'étudier le rôle de l'eau pour des protéines en solution ($h \simeq 4$ et $h = 7.5$) et des protéines totalement déshydratées représentées par des protéines dans le vide ($h = 0$) ainsi que pour des protéines cristallisées ($h = 0.9$). Les temps de simulations correspondent aux ordres de grandeur des fenêtres temporelles en neutrons (cf chapitre 3, partie I).

Dynamique des protéines déshydratées

2.1 Modèles simplifiés de la transition dynamique

2.1.1 Modèle de Doster

L'étude de la transition dynamique passe par la compréhension des mouvements internes de la protéine. Ils peuvent être formellement décomposés en une partie vibrationnelle et une partie « diffusive » [Moritsugu et Smith 06, Doster *et al.* 89].

Les mouvements « diffusifs » sont des mouvements internes activés thermiquement qui nécessitent le passage d'un puits d'énergie libre à un autre (mouvements inter-minima).

Les mouvements « vibrationnels » sont ceux pour lesquels les atomes de la protéine fluctuent autour de leurs positions d'équilibre (mouvements intra-minima).

Afin d'interpréter les mesures des fluctuations structurales de la myoglobine et l'existence d'une transition dynamique, Doster proposa un modèle très simplifié en double puits asymétrique (cf. Fig. 2.1 de l'énergie libre de la protéine dans l'ensemble NPT, une discussion détaillée de la surface d'énergie libre des protéines sera présentée dans la section 3.5 de cette partie).

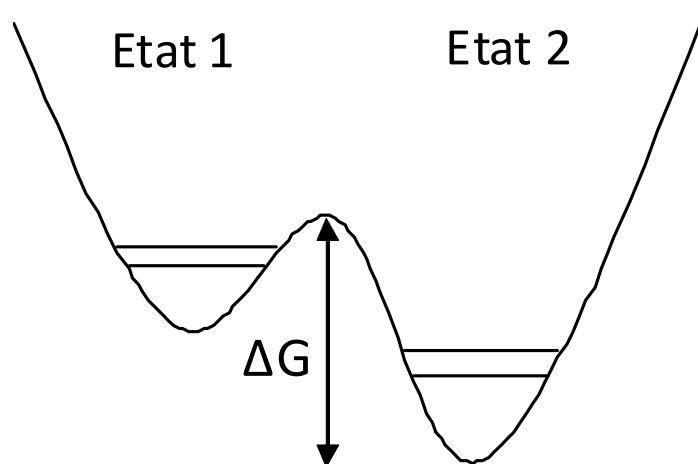


FIG. 2.1 – Barrière de potentiel ΔG séparant deux puits de potentiels. Dans l'ensemble NPT, $\Delta G = \Delta H - T\Delta S$, où $\Delta H = \Delta U + P\Delta V$. Dans l'ensemble NVT, l'énergie libre est notée F , telle que $\Delta F = \Delta U - T\Delta S$.

Dans ce modèle, [Doster *et al.* 89], la variance moyenne des déplacements atomiques (notée $\frac{MSD}{N}$, cf section 2.2 de la partie I) peut s'écrire :

$$\frac{MSD}{N} = \sigma_{vib}^2 + P_1 P_2 \frac{d^2}{3}, \quad (2.1)$$

où le premier terme σ_{vib}^2 représente la contribution vibrationnelle aux fluctuations des positions des atomes dans un des minima de la figure 2.1 et où le second terme est un modèle des mouvements diffusifs, activés entre les deux minima de l'énergie libre. Dans l'Eq. 2.1, P_1 et P_2 sont les probabilités de trouver l'atome dans l'état fondamental et dans l'état excité, respectivement (cf Fig. 2.1). La quantité d est une distance effective entre les deux minima.

Le terme vibrationnel peut se décomposer en une partie harmonique et une partie anharmonique :

$$\sigma_{vib}^2 = \sigma_{vib}^2(harm) + \sigma_{vib}^2(anharm). \quad (2.2)$$

Dans l'approximation de la mécanique classique, le terme harmonique est proportionnel à la température (théorème d'équipartition de l'énergie [Landau et Lifshitz 84]) :

$$\sigma_{vib}^2(harm) = 3 \frac{k_B T}{K_R}, \quad (2.3)$$

où K_R est la constante de force effective, k_B est la constante de Boltzmann et T est la température.

La contribution $\sigma_{vib}^2(anharm)$ est supposée beaucoup plus petite que la contribution anharmonique des mouvements diffusifs et peut être négligée en première approximation.

Dans le modèle de Doster, Eq. 2.1 et 2.3, la transition dynamique est entièrement due à l'activation de mouvements diffusifs, qui dépendent des probabilités P_1 et P_2 .

Selon la loi de Boltzmann :

$$\frac{P_2}{P_1} = \exp\left(-\frac{\Delta G}{RT}\right) = \frac{p_{12}}{p_{21}}, \quad (2.4)$$

où $\Delta G = \Delta H - T\Delta S$ est la hauteur de barrière (cf Fig. 2.1) avec ΔH la différence d'enthalpie et ΔS la variation d'entropie. Dans la dernière égalité à droite, p_{12} (p_{21}) représente la probabilité de transition de l'état 1 (2) vers l'état 2 (1) [Arrhenius 89].

Le modèle de Doster du $\frac{MSD}{N}$ d'une protéine s'écrit finalement :

$$\frac{MSD}{N} = 3 \frac{k_B T}{K_R} + \frac{\exp\left(-\frac{\Delta H}{RT} + \frac{\Delta S}{R}\right)}{\left[1 + \exp\left(-\frac{\Delta H}{RT} + \frac{\Delta S}{R}\right)\right]^2} \cdot \frac{d^2}{3}. \quad (2.5)$$

Dans l'article original de Doster, le modèle a été affiné uniquement sur les déplacements quadratiques des atomes d'hydrogène, les seuls mesurés en diffusion incohérente de neutrons (cf chapitre 3, partie I). Dans notre travail, nous avons appliqué ce modèle à toutes les protéines étudiées (pour $h = 4$) mais pour représenter la transition dynamique de l'ensemble des atomes.

Plusieurs méthodes d'affinement des paramètres ont été utilisées. Puisque dans nos calculs de dynamique moléculaire nous avons observé un comportement linéaire du $\frac{MSD}{N}$ en-dessous de $150K$, nous avons contraint la constante K_R (Eq. 2.3) à sa valeur ajustée sur les fluctuations à basses températures ($T < 150K$) en négligeant les mouvements activés (second terme de l'équation 2.1). Nous présentons les résultats de cet ajustement dans ce chapitre. Cependant d'autres méthodes d'ajustement ont été réalisées et un ajustement de tous les paramètres avec les données non-pondérées de la viscotoxine A3 est donné en annexe J. Il faut noter qu'un ajustement de l'ensemble des paramètres sans contrainte ne donne pas systématiquement un résultat physique acceptable.

L'ajustement des paramètres du modèle de Doster sur nos calculs de dynamique moléculaire est représenté sur la Fig. 2.2 pour la viscotoxine VA3.

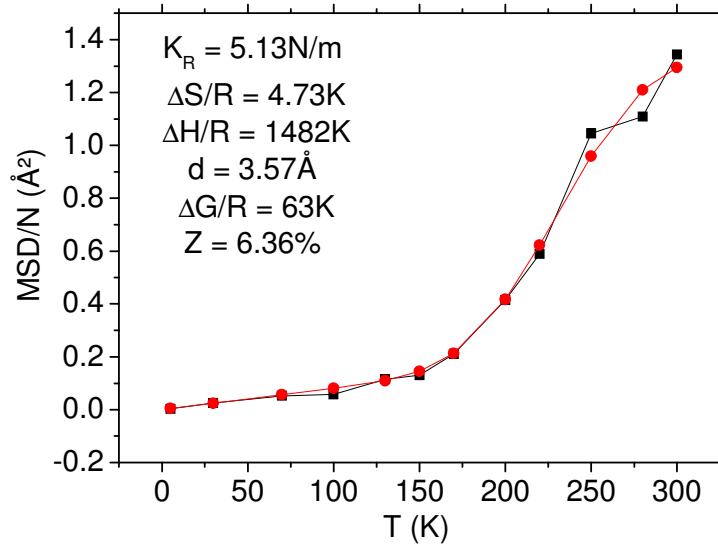


FIG. 2.2 – Ajustement des paramètres du modèle de Doster pour la viscotoxine VA3 hydratée. La courbe noire est l'écart quadratique moyen calculé par nos simulations de dynamique moléculaire et la courbe rouge est le modèle de Doster.

Un accord satisfaisant avec les résultats de dynamique moléculaire est obtenu dans ce cas ($Z = 3.36\%$, où Z est le pourcentage d'erreur entre la courbe de dynamique moléculaire, courbe noire, et le modèle, courbe rouge), ainsi que pour toutes les crambines et viscotoxines étudiées dans notre thèse (cf annexe J). La différence d'énergie libre ΔG a été calculée à 300 K .

Dans le tableau suivant, nous avons résumé les résultats des ajustements pour toutes les protéines étudiées.

	P22/L25	S22/I25	P22/I25	S22/L25	VA3	VA2	VB
$K_R (N/m)$	5.22	4.19	4.74	4.19	5.13	4.44	4.62
$\frac{\Delta G}{R} (K)$	17	133	912	147	63	992	109
$\Delta G (kcal/mol)$	0.034	0.264	1.812	0.292	0.125	1.971	0.217
$Z (\%)$	7.74	13.15	6.94	3.61	6.36	15.93	7.32

TAB. 2.1 – Résultats des ajustements du modèle de Doster pour les thionines hydratées.

Tout d'abord, les résultats obtenus sont en accord avec nos simulations car le pourcentage d'erreur entre les deux courbes reste faible. Nous constatons que les constantes effectives sont du même ordre de grandeur pour toutes les protéines. Dans la partie harmonique, peu de différences interviennent. VA2 et P22/I25 ont les valeurs de barrière les plus élevées et P22/L25, S22/I25 les plus faibles.

2.1.2 Modèle harmonique à deux états

Dans la suite de notre étude, l'état associé aux états conformationnels de basses températures sera appelé l'état R(= rigide), et celui associé aux états conformationnels de hautes températures l'état F (= flexible). Physiquement, l'état R représente un état où la protéine est "piégée" dans un minima (la protéine est bloquée dans un puits de potentiel). Dans l'état F, la flexibilité de la protéine permet l'ouverture de canaux diffusifs, la protéine pourra ainsi explorer plusieurs états conformationnels (la protéine passera du puits de gauche à celui de droite et vice versa, cf Fig. 2.1).

En nous inspirant du modèle de Doster [Doster *et al.* 89] et de Zaccai [Zaccai 00], nous définissons deux constantes de force effectives, une à basses températures (état R, pour $T < T_d$), l'autre à hautes températures (état F, pour $T > T_d$), reliées à l'écart quadratique moyen par l'Eq. 2.6 :

$$\begin{aligned} \frac{1}{K_R} &= \frac{1}{3} \left[\frac{\partial \langle \bar{X}^2 \rangle}{\partial (k_B T)} \right] \text{ pour } T < T_d, \\ \frac{1}{K_F} &= \frac{1}{3} \left[\frac{\partial \langle \bar{X}^2 \rangle}{\partial (k_B T)} \right] \text{ pour } T > T_d, \end{aligned} \quad (2.6)$$

où $\langle \bar{X}^2 \rangle$ est l'écart quadratique moyen, k_B la constante de Boltzmann, K_R la constante effective de l'état R et K_F la constante effective de l'état F.

Dans notre modèle (Eq. 2.6), chaque atome est dans un puits de potentiel identique harmonique de constante effective K (modèle d'Einstein). L'augmentation de la température provoque un élargissement du puits.

Cet élargissement se traduit par une augmentation d'entropie. En effet, la fonction de partition Z d'un oscillateur harmonique quantique isotrope à 3D est :

$$Z = Z_x Z_y Z_z = \frac{e^{-\frac{3}{2}\beta\hbar\omega}}{(1 - e^{-\beta\hbar\omega})^3}, \quad (2.7)$$

où Z_x, Z_y, Z_z représentent les fonctions de partition suivant x, y et z respectivement, \hbar est la constante de Planck, $\beta = \frac{1}{k_B T}$ et ω la fréquence telle que $\omega = \sqrt{\frac{K}{m}}$, avec K la constante de force et m la masse.

L'entropie de l'oscillateur est exactement :

$$S = 3k \left[\left(\frac{1}{u} \right) + \left(\frac{1}{u} \right) \left(\frac{1}{e^{\frac{1}{u}} - 1} \right) - \ln \left(1 - e^{-\frac{1}{u}} \right) \right], \quad (2.8)$$

où $u = \frac{kT}{\hbar\omega}$.

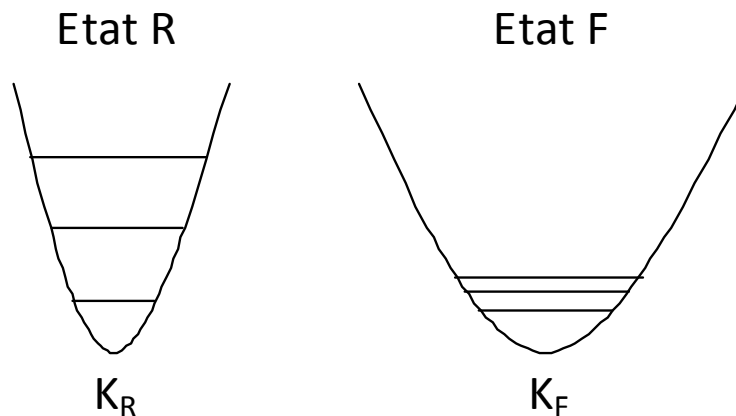


FIG. 2.3 – Variation du puits de potentiel entre l'état R, à basses températures, et l'état F, à hautes températures.

A hautes températures ($u \gg 1$), l'expression se simplifie :

$$S \simeq 3k \left(1 + \ln \frac{kT}{\hbar\omega} \right) = 3k (1 + \ln u). \quad (2.9)$$

L'équation 2.9 est l'entropie d'un oscillateur classique. Cette formule est appropriée à nos calculs (classiques) de dynamique moléculaire. De façon remarquable, la différence d'entropie entre les états R et F ne dépend que du rapport des constantes de force :

$$S_F - S_R \simeq k_B \ln \left(\frac{K_R}{K_F} \right)^{\frac{3}{2}}, \quad (2.10)$$

où S_F représente l'entropie pour l'état F, S_R l'entropie pour l'état R.

En calculant le rapport entre les deux constantes de force, nous aurons ainsi une estimation de l'entropie créée par la transition dynamique. Il faut bien entendu garder à l'esprit que la flexibilité de l'état F est due certainement à des mouvements diffusifs. Notre modèle, comme les autres modèles simplifiés, ne donne pas de description microscopique du phénomène, leur objectif est de quantifier le comportement global de différentes protéines. Il ne faut donc pas sur-interpréter les valeurs ainsi obtenues. Notons cependant que nous pouvons donner une interprétation microscopique rigoureuse de la constante K_R comme nous le verrons plus loin (dans la section 2.2.2).

Les constantes de force effectives des protéines ainsi obtenues sont généralement de l'ordre de $0.1 - 3N/m$.

2.2 Flexibilité conformationnelle de protéines déshydratées

2.2.1 Résultats pour la transition dans le vide des crambines et des viscotoxines

2.2.1.1 Paramètres tirés des modèles

La courbe 2.4 représente l'évolution de l'écart quadratique moyen avec la température de la crambine P22/L25 dans le vide pour une trajectoire de 1ns.

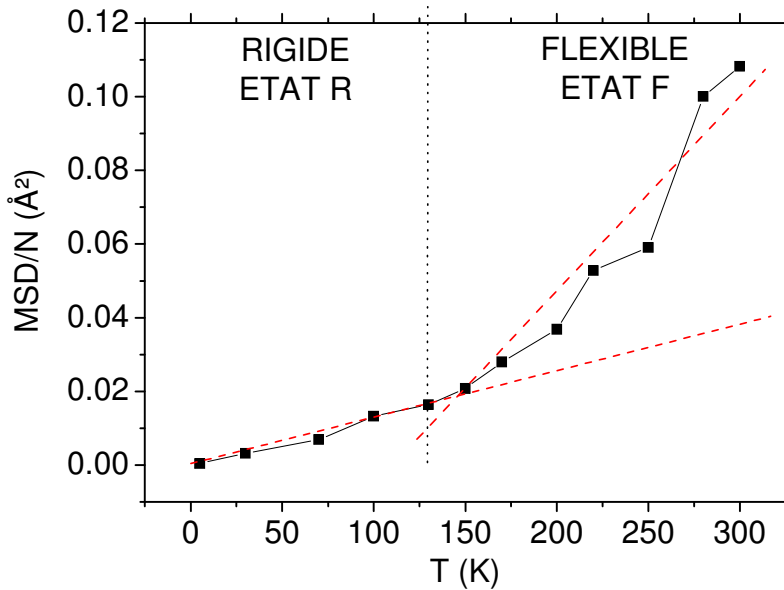


FIG. 2.4 – Evolution de l'écart quadratique moyen de P22/L25 dans le vide, sur une trajectoire de 1ns, les lignes rouges en traits interrompus correspondent aux régressions linéaires dans les états R et F.

A partir de cette courbe, nous pouvons tirer les paramètres effectifs décrits précédemment. En effet, nous pouvons tout d'abord visualiser une rupture de pente pour $T \simeq 130 - 150K$. Cette rupture de pente définit les deux états, l'état R à basses températures pour la protéine rigide et l'état F à hautes températures pour la protéine flexible. Puis, en traçant les régressions linéaires à la courbe pour ces deux domaines, nous pouvons calculer les deux constantes de force équivalentes K_R et K_F de la façon suivante et d'après l'équation 2.6 :

$$\frac{MSD}{N} = \underbrace{3 \frac{k_B}{K}}_{\text{pente de la régression linéaire (Å}^2\text{)}} \times T, \quad (2.11)$$

où k_B représente la constante de Boltzman, N le nombre d'atomes et K la constante effective.

L'état R est défini pour $0 \leq T \leq 100K$ et l'état F est défini pour $130 \leq T \leq 300K$. A partir de l'équation 2.10, nous avons évalué la différence d'entropie entre ces deux états. Nous obtenons ainsi pour la crambine P22/L25 :

	$K_R(N/m)$	$K_F(N/m)$	$\Delta S_{FR}(kcal/mol)$
P22/L25	34.6	7.5	4.5

TAB. 2.2 – Paramètres effectifs de la crambine P22/L25 dans le vide pour une trajectoire de 1ns.

De la même façon, nous avons représenté l'écart quadratique moyen des autres thionines, dans le vide et pour une trajectoire de 1ns (Fig. 2.5).

	$K_R(N/m)$	$K_F(N/m)$	$\Delta S_{FR}(kcal/mol)$
P22/L25	34.6	7.5	4.5
S22/I25	35.1	7.8	4.5
P22/I25	32.4	7.4	4.4
S22/L25	38.0	9.4	4.1
VA3	32.5	12.6	2.8
VA2	43.6	8.7	4.8
VB	47.2	15.5	3.3

TAB. 2.3 – Paramètres effectifs des thionines dans le vide.

Nous pouvons constater qu'à basses températures, les constantes de force des thionines sont du même ordre de grandeur, alors qu'à hautes températures, les différences sont plus prononcées. Nous pouvons déjà remarquer qu'elles n'ont pas la même flexibilité : P22/I25 est la plus flexible et la viscotoxine B la plus rigide dans l'état F. De plus, les courbes représentant la variation du MSD (cf Fig. 2.4 et 2.5) diffèrent d'une protéine à une autre, ceci nous indique qu'une seule mutation peut provoquer un comportement global totalement différent vis-à-vis de l'élévation de la température.

Cependant, il est difficile de parler d'une "transition dynamique" pour ces protéines déshydratées. En effet, la variation de l'écart quadratique moyen ne présente pas une rupture de pente très prononcée. Mais un phénomène existe bel et bien dans le vide, il est responsable du passage de l'état R à l'état F. Nous étudierons ce phénomène en détails dans la suite de notre analyse.

Nous pouvons également remarquer la plus grande augmentation d'entropie (et donc de désordre dans la protéine) pour VA2 et la plus faible pour VA3. Ces paramètres nous indiquent d'ores et déjà des différences de comportements dynamiques dans le vide au sein de cette famille de protéines.

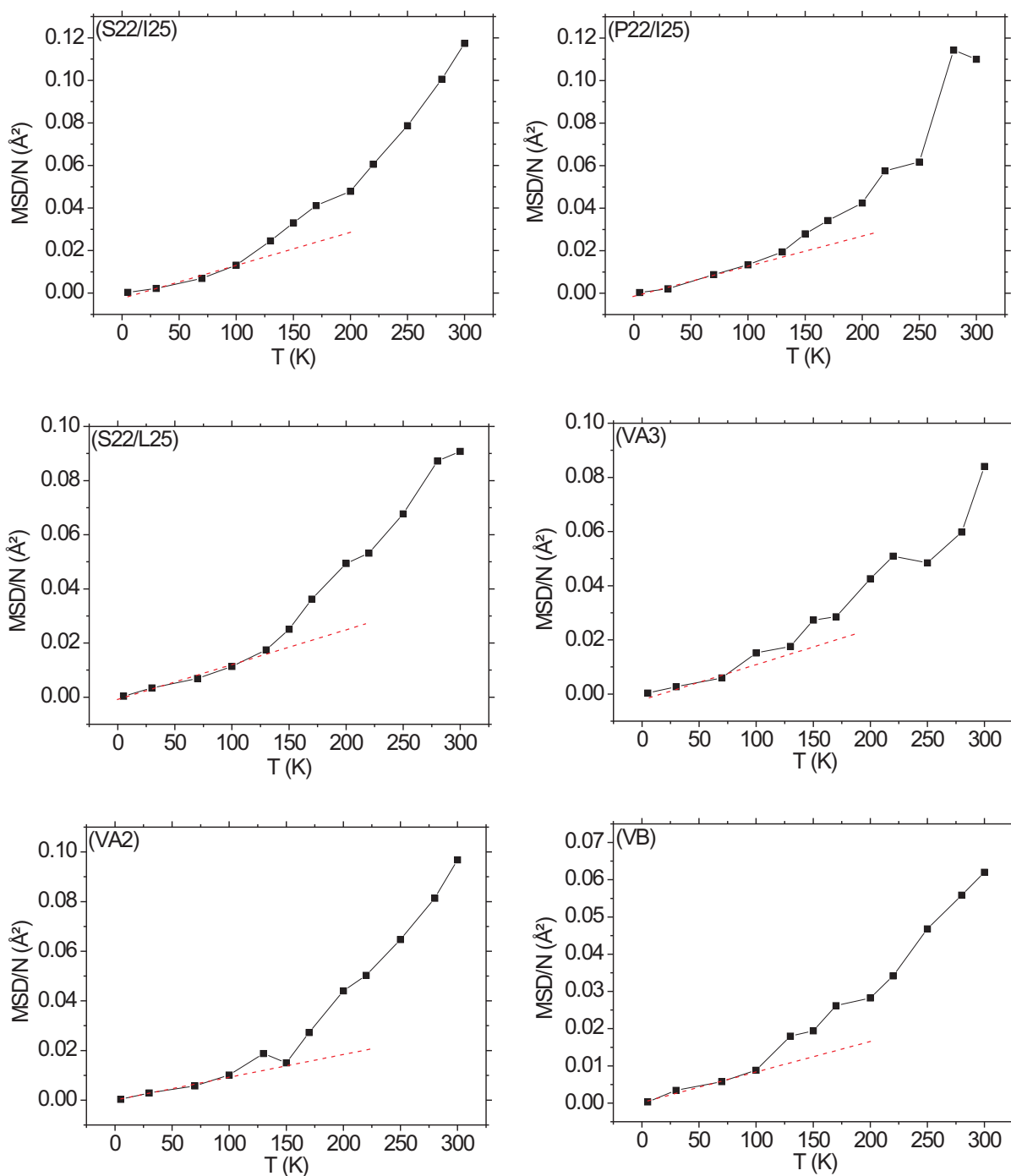


FIG. 2.5 – Evolution de l'écart quadratique moyen des thionines dans le vide pour une trajectoire de 1ns. Les lignes rouges en traits interrompus représentent la régression linéaire de la courbe pour l'état R, nécessaire pour le calcul des constantes effectives.

Puis nous nous sommes intéressés aux paramètres effectifs pour un type d'atome donné, à savoir d'une part tous les hydrogènes de la protéine, d'autre part tous les atomes qui ne sont pas des hydrogènes. Nous avons alors tracé l'évolution de l'écart quadratique moyen (cf Fig. 2.6 et 2.7) pour ces différents types d'atomes et de la même façon, nous avons pu calculer les constantes de force.

Ces figures nous indiquent également que les atomes lourds ne subissent aucun changement de dynamique : leur mobilité augmente linéairement avec la température, il est donc impossible de définir les deux états R et F. C'est pourquoi nous avons tracé une seule régression linéaire sur la totalité de la courbe, de coefficient de corrélation R^2 , et défini une seule constante de force K_R .

	$K_R(N/m)$	$K_F(N/m)$	$\Delta S_{FR}(kcal/mol)$
P22/L25	21.4	4.1	4.9
S22/I25	21.3	4.2	4.8
P22/I25	20.6	4.2	4.7
S22/L25	24.2	5.2	4.5
VA3	20.9	7.1	3.2
VA2	28.4	4.8	5.3
VB	29.9	8.8	3.6

TAB. 2.4 – Paramètres effectifs des thionines dans le vide pour les atomes d'hydrogène.

	$K_R(N/m)$	R^2
P22/L25	59.2	0.892
S22/I25	59.2	0.964
P22/I25	46.0	0.861
S22/L25	59.2	0.945
VA3	59.2	0.925
VA2	59.2	0.907
VB	82.8	0.936

TAB. 2.5 – Paramètres effectifs des thionines dans le vide pour tous les atomes lourds.

L'analyse de ces résultats met en évidence la forte mobilité des hydrogènes. En effet, les constantes effectives des hydrogènes sont inférieures à celles des autres atomes (i.e. plus la constante effective sera faible, plus les atomes seront mobiles). De plus, en comparant la variation d'entropie de la protéine globale et celle des différents types d'atomes, nous pouvons constater que cette variation est quasiment due à l'entropie créée par les hydrogènes. Les atomes lourds ont seulement une contribution harmonique (et donc faible) dans la mobilité des thionines dans le vide. 5 protéines sur les 7 étudiées ont la même constante effective pour ce type d'atomes. Les atomes lourds de la viscotoxine B sont les moins mobiles.

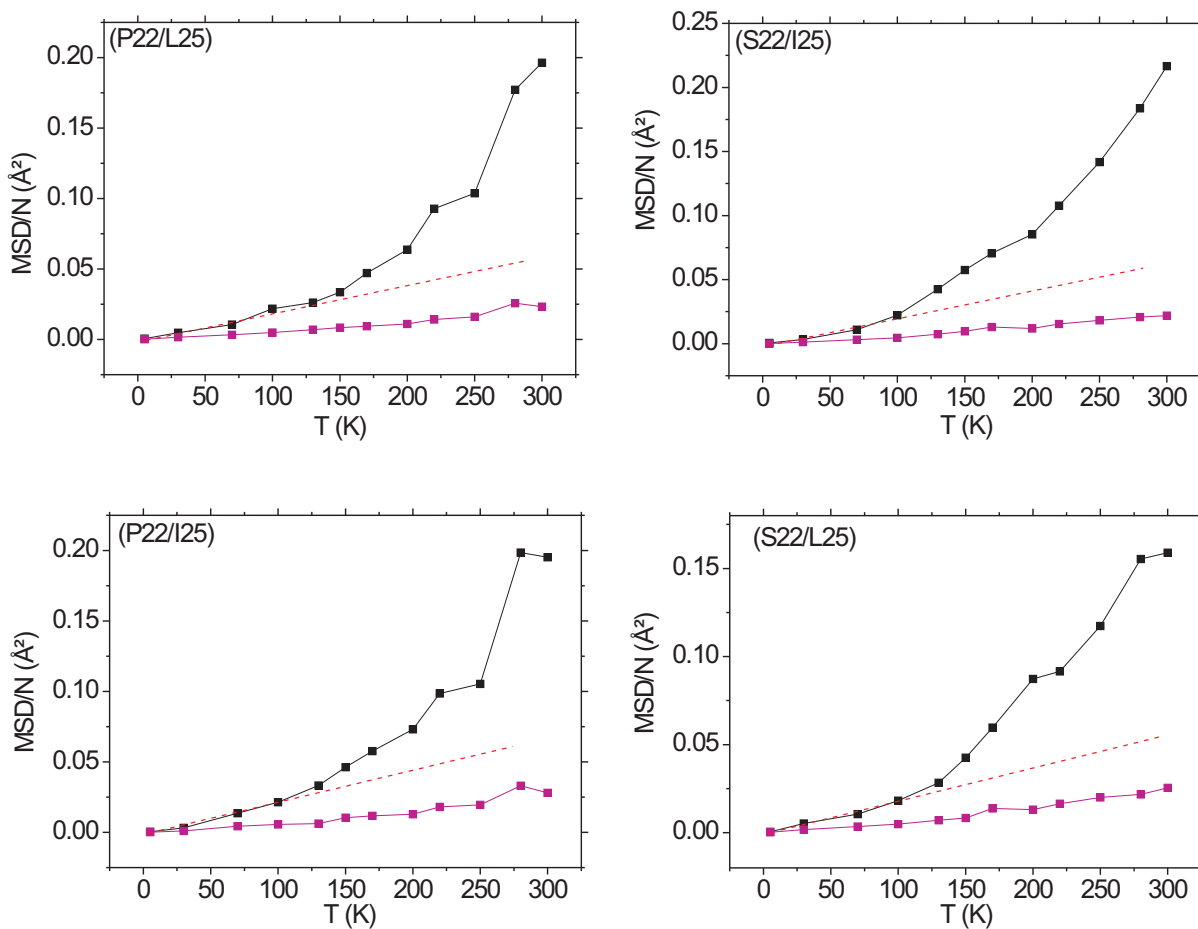


FIG. 2.6 – Evolution de l'écart quadratique moyen des différents types d'atomes pour les crambins dans le vide sur une trajectoire de 1ns. La courbe rose est l'écart quadratique moyen de tous les atomes qui ne sont pas des hydrogènes, la courbe noire est la même variation pour les hydrogènes et la ligne rouge en traits interrompus est la régression linéaire de la courbe dans l'état R.

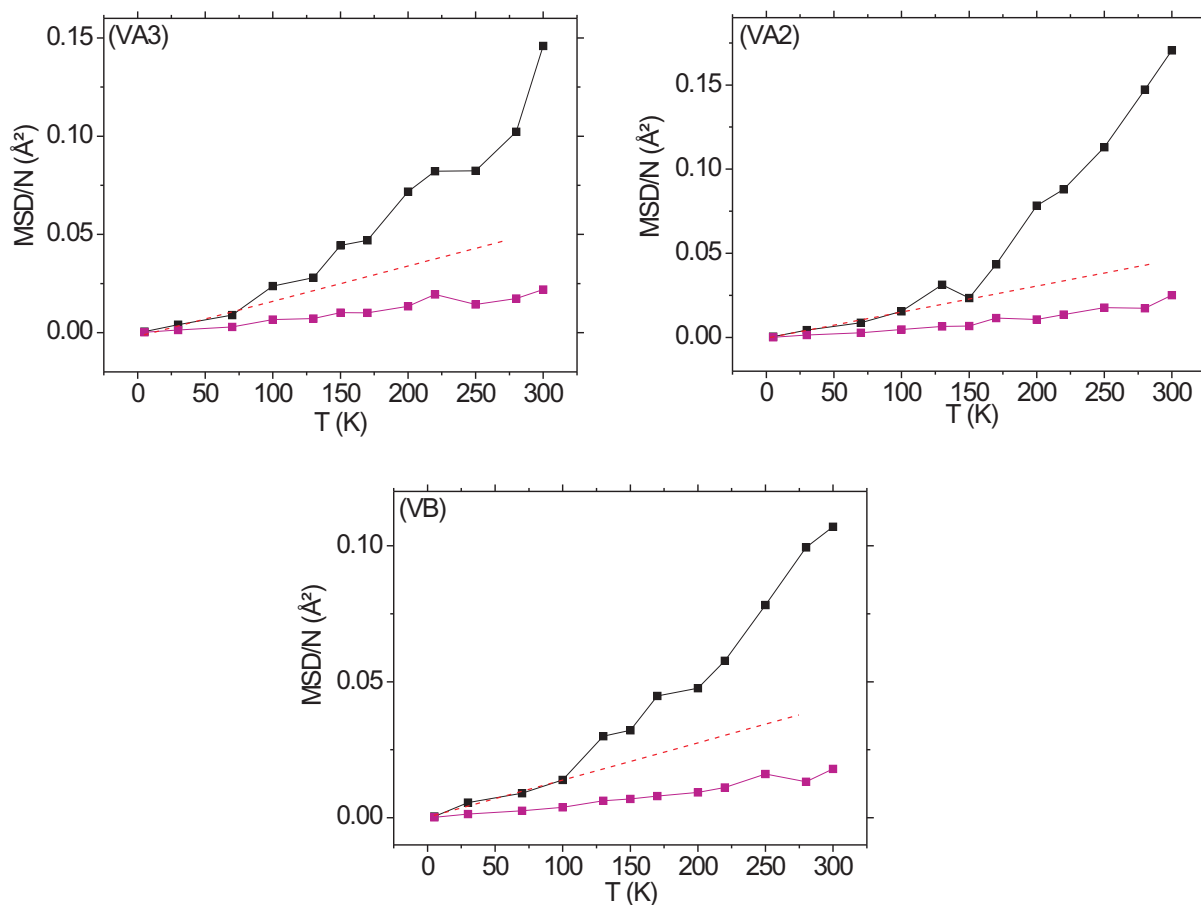


FIG. 2.7 – Evolution de l'écart quadratique moyen des différents types d'atomes pour les viscotoxines dans le vide sur une trajectoire de 1ns. La courbe rose représente l'écart quadratique moyen de tous les atomes qui ne sont pas des hydrogènes, la courbe noire représente la variation pour les hydrogènes seulement et la ligne rouge en traits interrompus est la régression linéaire de la courbe dans l'état R.

2.2.1.2 Comparaison qualitative avec l'expérience

Les données expérimentales sur les protéines déshydratées ne sont pas disponibles. Afin d'obtenir un premier ordre de grandeur des constantes de force, nous avons appliqué notre méthode de calcul à la courbe de la transition dynamique (cf Fig. 3.2, p.39) de la **myoglobine hydratée** obtenue par Parak en 2003 [Parak 03], nous obtenons les valeurs du tableau 2.6 (les constantes effectives obtenues par neutrons sont calculées seulement pour les atomes d'hydrogène).

Dans le cas de la myoglobine, les valeurs obtenues pour les atomes d'hydrogène dans les états R et F sont du même ordre de grandeur que celles obtenues pour les petites protéines globulaires **déshydratées**. On observe également une différence importante entre les paramètres effectifs mesurés par spectroscopie Mössbauer (atomes lourds et environnement du

groupe heme) et ceux mesurés par diffusion incohérente de neutrons (atomes d'hydrogène) : dans l'état R, les constantes des atomes lourds sont plus élevées que celles des hydrogènes. Cette différence est similaire à celle trouvée dans nos calculs pour les atomes "lourds" (cf tableau 2.5) et celle des atomes légers (cf tableau 2.4). Par contre, nous observons que les constantes de force des atomes lourds de la myoglobine dans l'état F sont du même ordre de grandeur que ceux des atomes légers dans l'état F. Ce qui n'est pas le cas dans notre étude des protéines **déshydratées**.

Myoglobine	$K_R(N/m)$	$K_F(N/m)$	$\Delta S_{FR}(kcal/mol)$
Mössbauer	25.9	3.0	3.9
Neutrons	12.4	3.6	5.8

TAB. 2.6 – Paramètres effectifs de la myoglobine hydratée.

2.2.1.3 Séparation des contributions en type d'atomes, Rôle des méthyles

Nous avons déjà pu constater que la flexibilité de la protéine dans le vide sera fortement provoquée par celles des hydrogènes. Ceci se vérifie également par la Fig. 2.8, où sont représentés les écarts quadratiques moyens de la protéine P22/L25, de ses atomes d'hydrogène et de tous les autres atomes la constituant. Nous pouvons ainsi remarquer que les fluctuations des hydrogènes dominent la flexibilité de la protéine. (Le même type de courbes a été obtenu pour les autres thionines, cf annexe D).

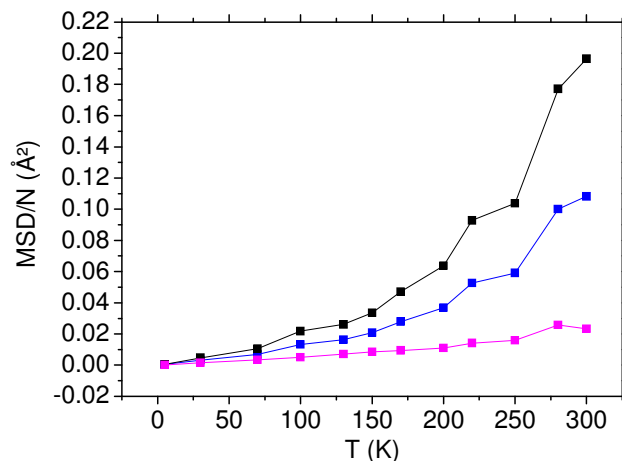


FIG. 2.8 – Evolution de l'écart quadratique moyen de la crambine P22/L25 (courbe bleue), de ses atomes d'hydrogène (courbe noire) et de tous ses autres atomes (courbe rose), dans le vide et sur une trajectoire de 1ns.

2.2.1.4 Dépendance en temps des simulations

Comme nous avons pu le voir dans un précédent chapitre (chapitre 1, partie I), les résultats obtenus dépendent des temps de simulations. En effet, les molécules se comportent comme des systèmes évoluant avec le temps. Nous avons ainsi simulé les protéines sur différents temps : 50ps, 100ps, 200ps et 1ns. Puis nous avons tracé l'évolution de l'écart quadratique moyen pour ces différentes trajectoires. La figure 2.9 représente les résultats pour la crambine P22/L25, mais les mêmes courbes ont été réalisées pour les sept thionines, (cf annexe E).

Nous pouvons remarquer sur la figure 2.9 que plus le temps de simulation est long, plus la protéine aura le temps de parcourir ses sous-états conformationnels. Ceci peut s'expliquer par le fait que la protéine a besoin d'un temps relativement long pour que toutes ses fluctuations structurales internes soient activées avec la température. Tous les mouvements des atomes ne débutent pas aux mêmes instants. De plus, jusqu'à 150K, les 4 courbes se superposent. Dans la partie harmonique, les temps de simulation n'impliquent donc aucune différence. Les autres thionines ont les mêmes comportements.

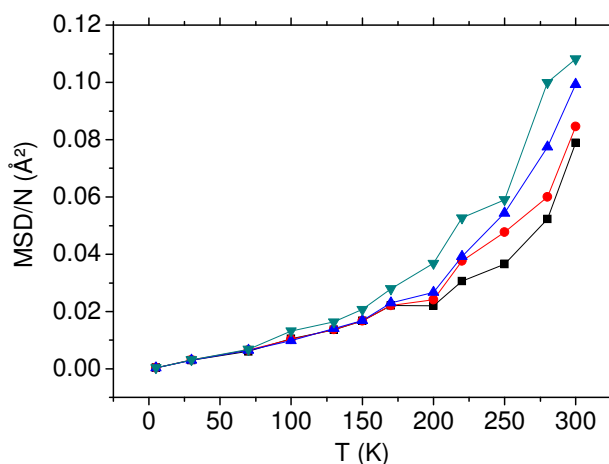


FIG. 2.9 – Dépendance en temps de simulations de l'écart quadratique moyen de la crambine P22/L25 dans le vide, la courbe noire représente une trajectoire de 50ps, la courbe rouge de 100ps, la courbe bleue de 200ps et la courbe verte de 1ns.

Nous avons calculé les constantes effectives dans les états R et F, en fonction du temps de simulation. Les résultats sont donnés dans le tableau 2.7. Nous pouvons remarquer que dans l'état R, les constantes sont du même ordre de grandeur pour des temps de production compris entre 50 et 200ps. Dans l'état F, les constantes de force diminuent lorsque le temps de simulation augmente (ie, la flexibilité augmente avec le temps de production). On peut supposer que les mouvements anharmoniques (état F) sont activés, ils correspondent à des sauts entre sous-états séparés par des barrières de potentiel. Plus la température augmente, plus la probabilité de passage est grande selon la loi d'Arrhenius. A une température donnée, la probabilité d'observer un saut dans une trajectoire de longueur finie augmente avec le temps de simulation.

	$K_R(N/m)$	$K_F(N/m)$	$\Delta S_{FR}(kcal/mol)$
50ps	41.9	12.1	3.7
100ps	41.5	10.4	4.1
200ps	42.8	8.6	4.7
1ns	34.6	7.5	4.5

TAB. 2.7 – Constantes effectives de la crambine P22/L25 dans le vide en fonction du temps de simulation.

2.2.1.5 Rayons de Giration

Le rayon de giration mesure la taille effective d'un polymère. D'après les équations 2.23 et 2.24 de la partie I, nous avons pu voir que ce rayon est relié aux B-facteurs, et donc au RMS (avec l'équation 1.7).

Nous avons alors calculé les variations des rayons de giration des protéines dans le vide à l'aide d'un programme fortran (donné en annexe T), qui analyse les variations du rayon de giration le long d'une trajectoire pour une température donnée. Plusieurs calculs ont ainsi été réalisés : tout d'abord, le calcul de ce rayon de giration pour toute la protéine, puis pour tous les atomes la constituant qui ne sont pas des hydrogènes (et pondérés par leur masse) et enfin les mêmes calculs pour le rayon de giration moyen, R_{gmean} (noté \bar{R}_g^2 dans l'Eq. 2.23).

Crambines Sur la Fig. 2.10, nous avons représenté la variation du rayon de giration avec l'augmentation de la température pour les crambines sur une trajectoire de 1ns. A gauche, nous avons tracé le rayon de giration total, alors qu'à droite il s'agit de R_{gmean} . Sur la Fig. 2.11, nous avons représenté de la même façon R_{gyr} et R_{gmean} , mais pour les protéines où les hydrogènes ne sont pas pris en compte dans le calcul.

De ces courbes nous pouvons conclure que les variations sont très faibles, les rayons de giration n'augmentent pas plus de 0.02\AA avec la température. Comme on doit s'y attendre, la structure est plus compacte (R_{gmean} plus petit) que la structure complète lorsque l'on ne tient pas compte des atomes d'hydrogène dans le calcul, quelque soit la protéine étudiée. De plus, il n'y a que très peu de différences entre R_{gyr} et R_{gmean} .

Ces résultats montrent également que les rayons de giration des crambines P22/L25 et P22/I25 ont des valeurs très proches (courbes bleue et noire), de même pour les crambines S22/I25 et S22/L25 (courbes rouge et verte). En analysant les séquences de ces protéines, nous pouvons remarquer que les protéines ayant le même résidu situé en 22^{ième} position ont des comportements semblables pour le rayon de giration. Les mutations sur le 25 n'ont que peu d'influences sur ce rayon. Ceci nous permet donc de conclure que le résidu 22 a certainement un rôle clé dans la flexibilité de la protéine, comme nous avons déjà pu le voir dans l'étude de la structure secondaire (cf chapitre 5, partie I).

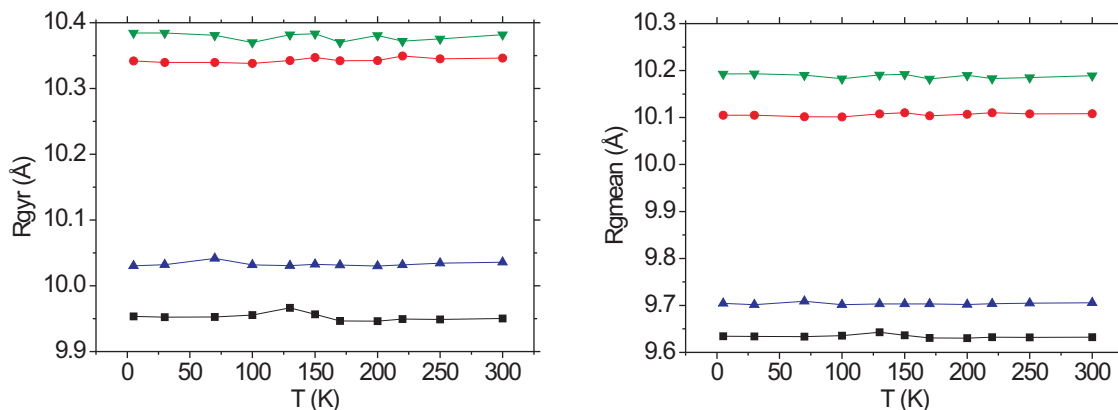


FIG. 2.10 – Rayon de giration des crambines dans le vide, sur une trajectoire de 1ns : à gauche : R_{gyr} (Eq. 2.23), à droite : R_{gmean} (Eq. 2.24). Les triangles verts correspondent à S22/L25, les ronds rouges à S22/I25, les triangles bleus à P22/I25 et les carrés noirs à P22/L25.

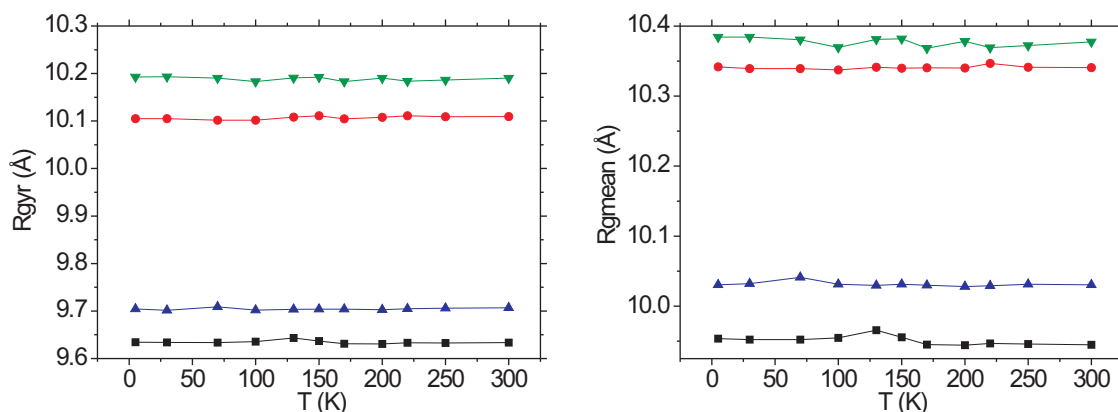


FIG. 2.11 – Rayon de giration des crambines dans le vide, en éliminant tous les atomes d'hydrogène, sur une trajectoire de 1ns : à gauche : R_{gyr} (Eq. 2.23), à droite : R_{gmean} (Eq. 2.24). Les triangles verts correspondent à S22/L25, les ronds rouges à S22/I25, les triangles bleus à P22/I25 et les carrés noirs à P22/L25.

Viscotoxines Dans le cas des viscotoxines (Fig. 2.12 et 2.13), aucune différence de comportement n'apparaît entre R_{gyr} et R_{gmean} . De plus, l'évolution du rayon de giration reste identique avec ou sans les hydrogènes. On observe comme attendu une diminution du rayon de giration moyen (R_{gmean}) si on ne tient pas compte des atomes d'hydrogène. Le rayon de VA3 diminue le plus et celui de VA2 le moins. C'est pourquoi le rayon de giration de VA2 est plus proche de VA3 avec les hydrogènes et de VB sans les hydrogènes.

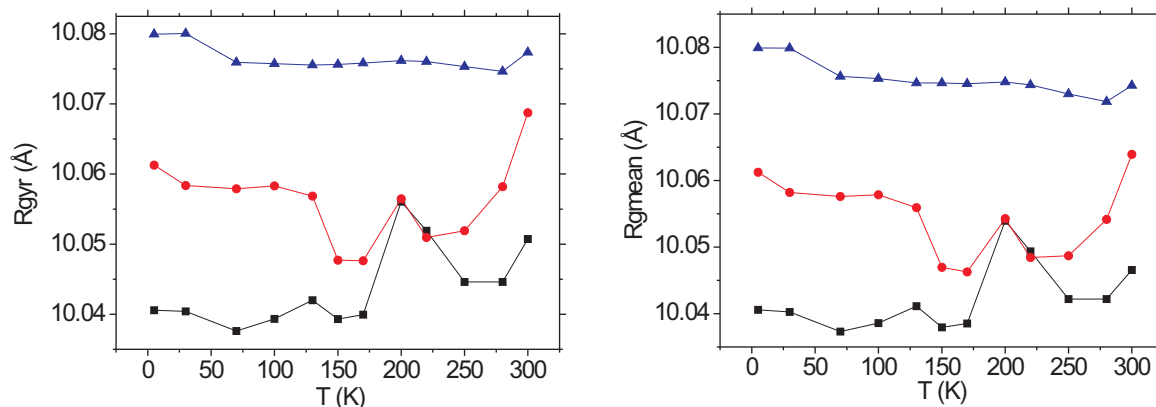


FIG. 2.12 – Rayon de giration de la viscotoxine A3 (carrés noirs), A2 (ronds rouges) et B (triangles bleus) dans le vide, sur une trajectoire de 1ns : à gauche : R_{gyr} , à droite : R_{gmean} .

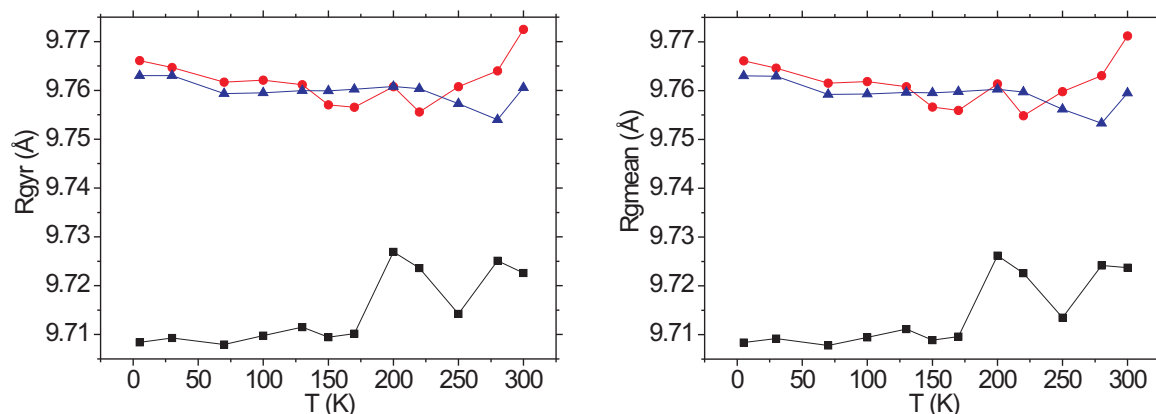


FIG. 2.13 – Rayon de giration de la viscotoxine A3 (carrés noirs), A2 (ronds rouges) et B (triangles bleus) dans le vide, en éliminant tous les atomes d'hydrogène, sur une trajectoire de 1ns : à gauche : R_{gyr} , à droite : R_{gmean} .

D'autre part, en comparant la variation de l'écart quadratique moyen de VA3 par rapport à son rayon de giration (cf Fig. 2.14), nous pouvons remarquer qu'à la température de transition, son rayon de giration augmente brusquement. Sa flexibilité, due à l'augmentation de la température, implique une augmentation de son rayon de giration.

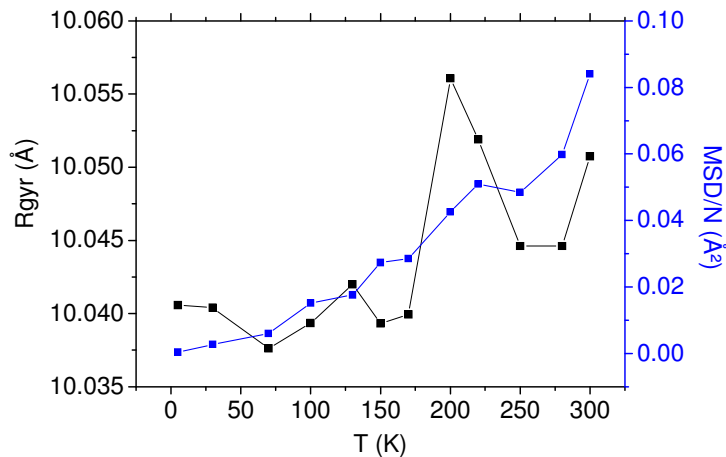


FIG. 2.14 – Comparaison entre l'évolution de l'écart quadratique moyen (carrés bleus) de la viscotoxine A3 par rapport à son rayon de giration (carrés noirs), dans le vide et sur une trajectoire de 1ns.

Quelle est la signification physique d'un pic dans la courbe $R_{gyr}(T)$? Pour répondre à cette question, nous définissons un rayon "instantané" calculé au temps t :

$$R_{gyr}^2(t) = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N w_i w_j d_{ij}^2(t). \quad (2.12)$$

Le rayon de giration est :

$$\langle R_{gyr}^2 \rangle = \frac{1}{t_{simul}} \int_0^{t_{simul}} dt R_{gyr}^2(t), \quad (2.13)$$

où t_{simul} est le temps de simulation. Admettons l'hypothèse ergodique (sa validité est approximative dans nos simulations numériques) :

$$\langle R_{gyr}^2 \rangle = \sum_{k=1}^n p_k R_{gyr}^2(k), \quad (2.14)$$

où n est l'ensemble des micro-états possibles de la protéine et $R_{gyr}^2(k)$ la valeur du rayon de giration dans l'état k qui est observée avec la probabilité p_k :

$$p_k = \frac{e^{-\frac{E_k}{k_B T}}}{Z}, \quad (2.15)$$

où E_k est l'énergie du micro-état k et Z la fonction de partition canonique de la protéine.

A partir de l'Eq. 2.14, on déduit :

$$\frac{\partial}{\partial T} \langle R_{gyr}^2 \rangle = \sum_{k=1}^n \left(\frac{\partial p_k}{\partial T} \right) R_{gyr}^2(k). \quad (2.16)$$

En utilisant l'Eq. 2.15, on a :

$$\frac{\partial}{\partial T} \langle R_{gyr}^2 \rangle = \frac{1}{k_B T^2} [\langle R_{gyr}^2 E \rangle - \langle R_{gyr}^2 \rangle \langle E \rangle], \quad (2.17)$$

qui est notre résultat principal. On conclut immédiatement qu'un pic dans la courbe $R_{gyr}(T)$ correspond à la condition :

$$\left[\frac{\partial}{\partial T} \langle R_{gyr}^2 \rangle \right]_{T=T_c} = 0. \quad (2.18)$$

Soit

$$\langle R_{gyr}^2 E \rangle = \langle R_{gyr}^2 \rangle \langle E \rangle, \quad (2.19)$$

c'est à dire que pour cette température critique T_c le rayon de giration et l'énergie de la protéine sont des variables statistiquement indépendantes. Les fluctuations des distances interatomiques (Eq. 2.12) sont indépendantes de l'énergie, les distances (certaines, celles qui fluctuent le plus) peuvent varier de manière appréciable sans que l'énergie ne soit affectée. Il s'agit d'une sorte d'"instabilité structurale". Le fait que nous observons un maximum implique que :

$$\left[\frac{\partial}{\partial T} \langle R_{gyr}^2 \rangle \right]_{T=T_c} < 0. \quad (2.20)$$

En dérivant l'Eq. 2.17, on trouve aisément que :

$$\left[\frac{\partial^2}{\partial T^2} \langle R_{gyr}^2 \rangle \right]_{T=T_c} = \frac{1}{k_B^2 T^4} [\langle R_{gyr}^2 E^2 \rangle - \langle R_{gyr}^2 \rangle \langle E^2 \rangle]. \quad (2.21)$$

Par conséquence, on déduit :

$$\langle R_{gyr}^2 E^2 \rangle < \langle R_{gyr}^2 \rangle \langle E^2 \rangle. \quad (2.22)$$

Les variations de distances interatomiques (R_{gyr}^2) sont corrélées au carré de l'énergie de la protéine.

Une expression équivalente à l'Eq. 2.21 est :

$$\left[\frac{\partial^2}{\partial T^2} \langle R_{gyr}^2 \rangle \right]_{T=T_c} = \frac{1}{k_B^2 T^4} [\langle R_{gyr}^2 E^2 \rangle - \langle E \rangle^2 \langle R_{gyr}^2 \rangle] - \langle R_{gyr}^2 \rangle C_v, \quad (2.23)$$

où C_v est la capacité calorifique à volume constant de la protéine :

$$C_v \equiv \left(\frac{\partial E}{\partial T} \right)_v = \frac{1}{k_B T^2} [\langle E^2 \rangle - (\langle E \rangle)^2]. \quad (2.24)$$

On en déduit une propriété intéressante :

$$C_v(T_c) > \frac{1}{k_B^2 T^4} \left[\frac{\langle R_{gyr}^2 E^2 \rangle}{\langle R_{gyr}^2 \rangle} - (\langle E \rangle)^2 \right]. \quad (2.25)$$

Le comportement de la capacité calorifique au voisinage des températures de transition est une propriété importante définissant les changements de phase. Il serait intéressant d'étudier ces propriétés dans un travail futur.

2.2.1.6 Test sur la transition dynamique dans le vide jusqu'à 450K

En analysant les courbes représentatives de l'écart quadratique moyen des protéines dans le vide (Fig. 2.4 et 2.5), il est difficile de visualiser une rupture de pente nette indiquant la présence d'une transition dynamique comme dans l'étude de la myoglobine par Parak (cf Fig. 1.1). Nous nous sommes alors demandés si la transition dynamique des protéines déshydratées pouvait avoir lieu à des températures supérieures à 300K. Si la transition dynamique est une propriété intrinsèque à la protéine, il serait alors possible d'envisager qu'une protéine déshydratée nécessiterait une température supérieure aux protéines hydratées pour déclencher sa transition dynamique. C'est pourquoi nous avons simulé la crambine S22/I25 jusqu'à 450K, et tracé l'évolution de son écart quadratique moyen sur la Fig. 2.15.

Ce test nous permet de constater qu'aucune transition dynamique n'apparaît dans le comportement global de la crambine S22/I25 pour une température supérieure à 300K. Même si la mobilité de la protéine augmente avec la température, il est difficile d'estimer qu'une rupture de pente très prononcée intervient à 200K ou à une température supérieure. C'est pourquoi la présence de cette transition pour les protéines déshydratées est aussi controversée à l'heure actuelle.

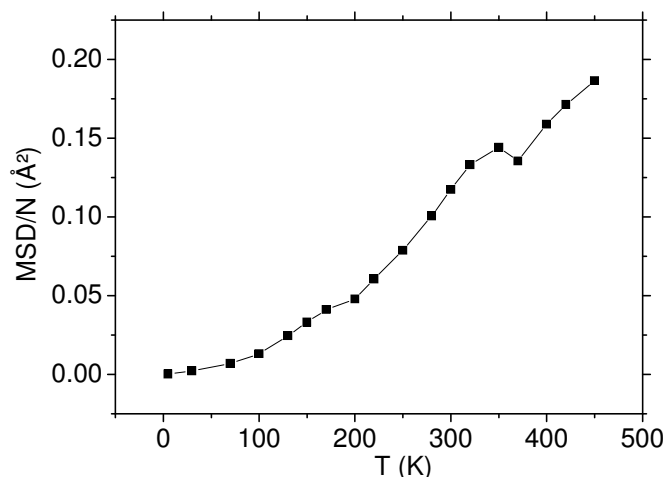


FIG. 2.15 – Evolution de l'écart quadratique moyen de la crambine S22/I25 dans le vide, sur une trajectoire de 1ns et jusqu'à 450K.

2.2.2 Désordre dynamique - inhomogénéité des fluctuations

2.2.2.1 Dispersion des fluctuations, définition

La dispersion spatiale des fluctuations nous informe sur l'évolution de la répartition de la mobilité le long de la séquence avec la température. Nous avons quantifié cette dispersion de la façon suivante :

$$\sigma^2 = \frac{\sum_{i=1}^N \left[MSD_i - \frac{MSD}{N} \right]^2}{N}, \quad (2.26)$$

où MSD_i représente le MSD de l'atome i , MSD représente la somme des MSD_i sur tous les atomes et N le nombre d'atomes.

Attardons nous sur la signification de ce nouveau concept (Eq. 2.26). Supposons que chaque atome soit dans un puits de potentiel harmonique de constante de force effective K_i . D'après le théorème d'équipartition, nous aurons à l'équilibre :

$$MSD_i = \frac{3k_B T}{K_i}. \quad (2.27)$$

Le MSD de la protéine sera la moyenne des MSD_i , soit :

$$MSD = 3k_B T \left(\sum_{i=1}^N \frac{1}{K_i} \right). \quad (2.28)$$

La constante de force effective évoquée plus haut (Eq. 2.6 de la partie II) est exactement dans ce modèle :

$$K_R = \frac{N}{\left(\sum_{i=1}^N \frac{1}{K_i} \right)}. \quad (2.29)$$

Dans l'approximation harmonique, la constante effective est dominée par les atomes possédant le plus petit K_i (les plus mobiles) comme nous l'avons d'ailleurs observé plus haut où les atomes d'hydrogène dominaient les fluctuations structurales des protéines. L'Eq. 2.29 donne le sens physique précis de la constante K_R introduite plus haut. Son application à l'état F est discutable dans la mesure où celui-ci n'est pas harmonique.

Que vaut le nouveau paramètre que nous avons introduit (Eq. 2.26) dans l'approximation harmonique (Eq. 2.27) ?

On déduit immédiatement :

$$\frac{\sigma_{harm}^2}{9k_B^2 T^2} = \frac{\sum_{i=1}^N \left[\frac{1}{K_i} - \frac{1}{K_R} \right]^2}{N} \equiv \sigma_K^2. \quad (2.30)$$

L'Eq. 2.30 est une propriété fondamentale : dans l'approximation harmonique $\frac{\sigma^2}{T^2}$ est une constante indépendante de la température. Sa valeur mesure la dispersion σ_K^2 des constantes de force par rapport à la constante moyenne, c'est un paramètre physique clairement identifié. Il complète le modèle d'Einstein que nous avons développé : K_R définit l'environnement

moyen d'un atome au sens de l'Eq. 2.29 et l'hétérogénéité des environnements dans la protéine est définie par $\frac{\sigma^2}{T^2}$ (Eq. 2.30). L'étude de la dispersion (renormalisée), $\frac{\sigma^2}{T^2}(T)$, nous renseigne directement sur l'hétérogénéité dynamique et sa dépendance en température sur l'activation de mouvements anharmoniques. Nous sommes à notre connaissance les premiers à étudier ce paramètre.

Nous avons tracé l'évolution de $\frac{\sigma^2}{T^2}$ en fonction de la température et nous l'avons comparée à celle de $\frac{MSD}{NT}$ pour la crambine P22/L25 (les mêmes courbes ont été tracées pour les autres thionines, cf annexe F).

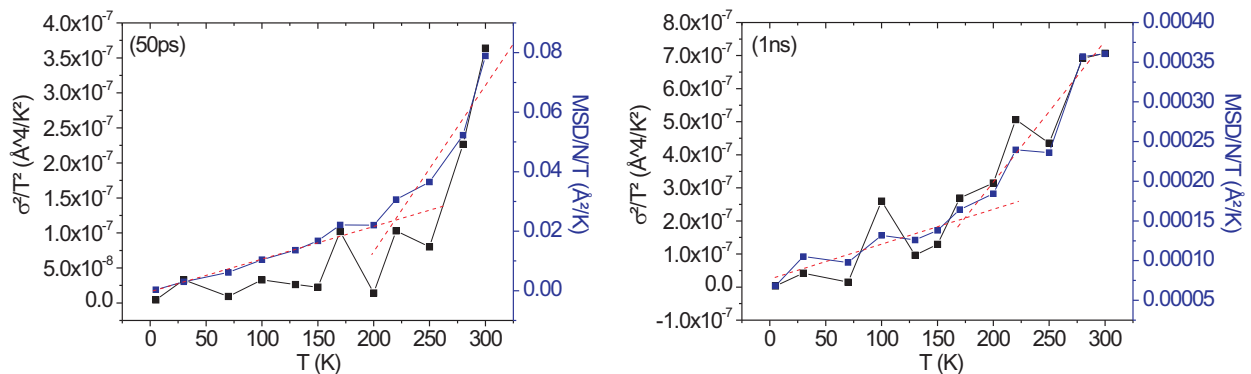


FIG. 2.16 – Variation de $\frac{\sigma^2}{T^2}$ (carrés noirs) et de $\frac{MSD}{NT}$ (carrés bleus) en fonction de la température pour la crambine P22/L25 dans le vide, sur une trajectoire de 50ps à gauche et sur une trajectoire de 1ns à droite. Les lignes rouges en traits interrompus correspondent aux régressions linéaires des états R et F.

La Fig. 2.16 nous indique qu'à la fois la flexibilité moyenne de la protéine $\frac{MSD}{N}$ et la dispersion $\frac{\sigma^2}{T^2}$ varient, et de la même manière, avec la température, contrairement à l'hypothèse harmonique. On observe clairement en particulier dans la courbe $\frac{\sigma^2}{T^2}$ l'activation de mouvements anharmoniques autour de 100 – 150K. Nous verrons dans la suite de notre étude que cette contribution est due majoritairement aux hydrogènes des groupes méthyles.

Nous pouvons en conclure que l'augmentation de la flexibilité de la protéine s'accompagne également d'une augmentation de l'hétérogénéité dynamique : plus la température sera élevée, plus il y aura de différences entre les fluctuations structurales des zones mobiles et des zones plus rigides de la protéine. En comparant ce calcul sur deux trajectoires, l'une de 50ps et l'autre de 1ns, nous pouvons constater qu'augmenter le temps de production revient à abaisser la température d'activation des mouvements anharmoniques des hydrogènes.

2.2.2.2 Courbes générales de la mobilité des thionines

La figure 2.17 représente la variation de la flexibilité des crambines d'une part, et des viscotoxines d'autre part. Ces courbes nous montrent tout d'abord que la différences entre les crambines sont minimales en ce qui concerne leur comportement global. Il est très difficile de déterminer quelle crambine sera la plus flexible. Pour les viscotoxines, les différences sont plus importantes. Alors que VA2 est la plus flexible (avec un $\frac{MSD}{N}_{\max} = 0.10\text{\AA}^2$), VB présente la mobilité la plus faible (avec un $\frac{MSD}{N}_{\max} = 0.06\text{\AA}^2$). Les ordres de grandeurs du $\frac{MSD}{N}$ sont les mêmes pour les crambines et les viscotoxines déshydratées. Nous pouvons conclure que si les mutations impliquent des différences dans la mobilité globale de la protéine, celles-ci restent faibles dans le vide et très probablement dans la barre d'erreur de nos simulations.

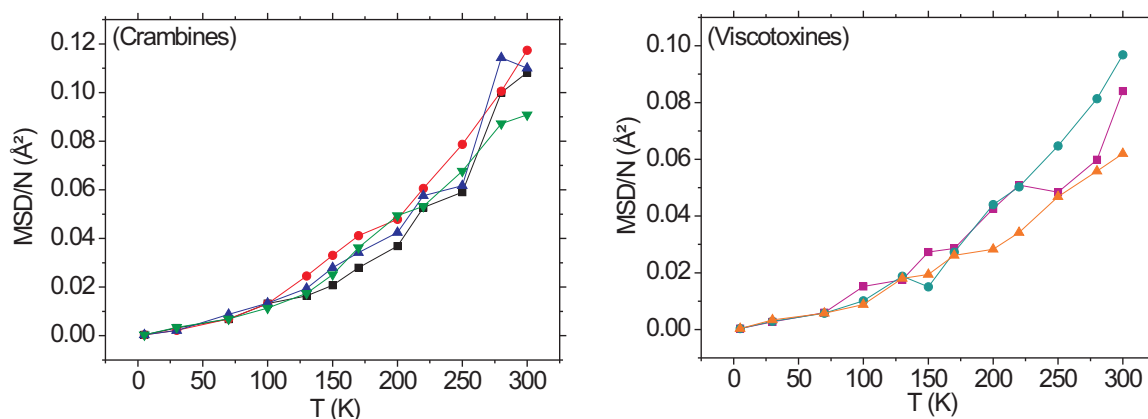


FIG. 2.17 – Comparaisons de la flexibilité globale des crambines (P22/L25 : carrés noirs, S22/I25 : ronds rouges, P22/I25 : triangles bleus, S22/L25 : triangles verts), à gauche, et des viscotoxines (VA3 : carrés roses, VA2 : ronds verts, VB : triangles oranges), à droite, dans le vide et sur une trajectoire de 1ns.

2.2.2.3 Superpositions des structures à 100K et 300K

Sur la Fig. 2.18, nous avons superposé les crambines (avec une représentation en rubans), à 100K à gauche et à 300K à droite. Cette superposition nous permet de constater les différences engendrées par l'augmentation de la température. Nous pouvons constater que les zones les plus sensibles à l'élévation de la température se situent surtout dans la boucle reliant les deux hélices et dans la partie où se situent les feuilletts beta. Une fois encore, nous constatons que les deux crambines avec une proline en 22ième résidu (P22/L25 en bleu et P22/I25 en vert) se superposent avec beaucoup de similitudes, de même pour les deux crambines avec une serine en 22ième résidu (S22/I25 en rouge et S22/L25 en violet).

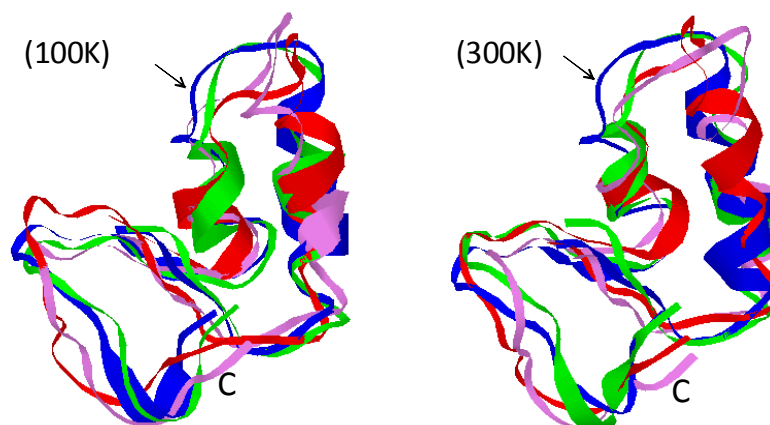


FIG. 2.18 – Superpositions des crambines équilibrées dans le vide, à 100K à gauche, et à 300K à droite : P22/L25 en bleu, S22/I25 en rouge, P22/I25 en vert et S22/L25 en violet. C indique l'extrémité de la protéine correspondante au résidu 46. La flèche localise l'acide aminé n°22.

Sur la Fig. 2.19, nous présentons la même représentation mais pour les viscotoxines.

Nous remarquons que les mêmes zones sont sensibles à l'élévation de la température : la boucle et les feuilletts beta. De manière générale, nous pouvons également remarquer que toutes ces protéines se superposent avec beaucoup de similitudes dans les régions en hélices. Les mutations apportent des différences dans les boucles et les terminaisons C et N.

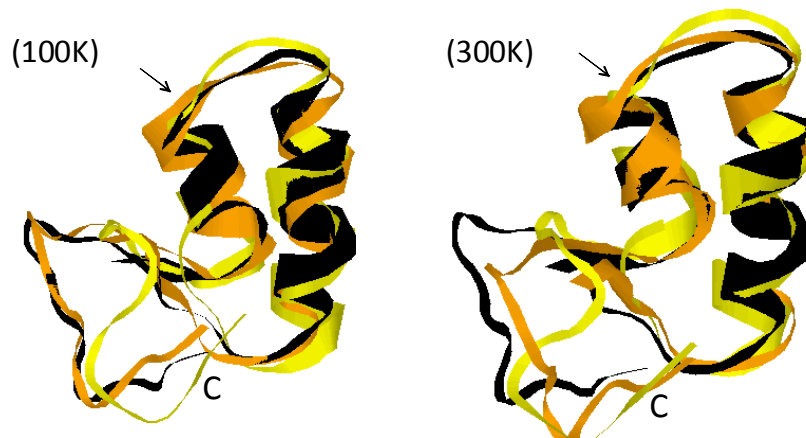


FIG. 2.19 – Superpositions des viscotoxines équilibrées dans le vide, à 100K à gauche, et à 300K à droite : VA3 en orange, VA2 en noir et VB en jaune. C indique l'extrémité de la protéine correspondante au résidu 46. La flèche localise l'acide aminé n°22.

Nous constatons sur la Fig. 2.19 que les structures de VA3 et VA2 se superposent à basses températures (100K). C'est le cas également à 300K sauf dans les régions des terminaisons N et C. On observe que la boucle dans la région C-terminale de VB est significativement différente de celle de VA3 et VA2, à toutes les températures.

2.2.2.4 Identification des atomes mobiles

Nous avons pu voir précédemment que la mobilité globale des protéines ne montrait pas une rupture de pente nette à la transition dynamique malgré une augmentation de sa flexibilité (Fig. 2.4 et 2.5). Nous avons alors regardé comment l'augmentation de la flexibilité de la protéine se traduisait le long de sa séquence et de ses atomes. Nous avons représenté le MSD de chaque atome aux différentes températures étudiées (cf Fig. 2.20), puis identifié à quels types d'atomes les MSD les plus élevés correspondaient.

Nous pouvons tout d'abord remarquer que ces résultats confirment les résultats précédents sur la dispersion : l'augmentation de la température implique une augmentation de zones mobiles, (représentées par des "pics" de mobilité). Ces pics correspondent majoritairement aux hydrogènes des groupes CH_2/CH_3 mais aussi aux groupes CH des cycles aromatiques.

En identifiant à quels types d'atomes correspondent ces pics, nous nous sommes rendu compte que la mobilité de la protéine était due principalement aux groupements méthyles. Les mêmes résultats ont été trouvés pour les autres thionines (les courbes équivalentes sont présentées dans l'annexe G). Notre étude est en accord avec les résultats de Lee et Wand (cités au chapitre 1, partie II, RMN pour des protéines hydratées [Lee et Wand 01]) : il existe un phénomène d'activation des méthyles dans le cas des protéines déshydratées qui est responsable d'une augmentation apparente de la flexibilité de la protéine avec la température.

Cependant, les pics de mobilité ne sont pas les mêmes d'une protéine à une autre. Les mutations impliquent une répartition de la mobilité totalement différente le long de la séquence. Par exemple, pour les crambines, le groupe CH_3 du 1^{er} résidu (atomes 11, 12, 13 et 14) est considéré comme mobile sauf pour S22/L25. De la même façon, le CH_3 du résidu 35 (atomes 493, 494, 495 et 496) n'a une forte mobilité que dans cette crambine et pas dans les autres. Pour les viscotoxines, VA2 montre une grande mobilité du CH_3 du résidu 29 (atomes 412, 413, 414 et 415) qui n'existe pas pour VA3 et VB alors qu'il s'agit du même résidu pour les 3 viscotoxines, la leucine.

Comme nous l'avons vu précédemment, les résultats obtenus dépendent du temps de simulation. Ceci se confirme également dans l'étude de la mobilité des atomes à 300K qui diffèrent pour des trajectoires de 50, 100, 200ps et 1ns (cf étude de P22/L25 dans le vide sur la Fig. 2.21, les mêmes résultats pour les autres thionines sont présentés en annexe H).

Par exemple, si nous regardons l'évolution de la mobilité du groupe CH_3 du résidu 15 (atomes 214, 215, 216 et 217), nous pouvons remarquer qu'il a une mobilité de 0.52\AA^2 à 50ps, 0.87\AA^2 à 100ps, 1.01\AA^2 à 200ps et 0.69\AA^2 à 1ns. L'augmentation n'est pas systématique avec le temps de production, tout dépendra de l'instant où le mouvement, que subit le groupe méthyle, sera activé. C'est ainsi que le groupe CH_3 du résidu 15 a une mobilité supérieure sur 200ps à celle sur 1ns.

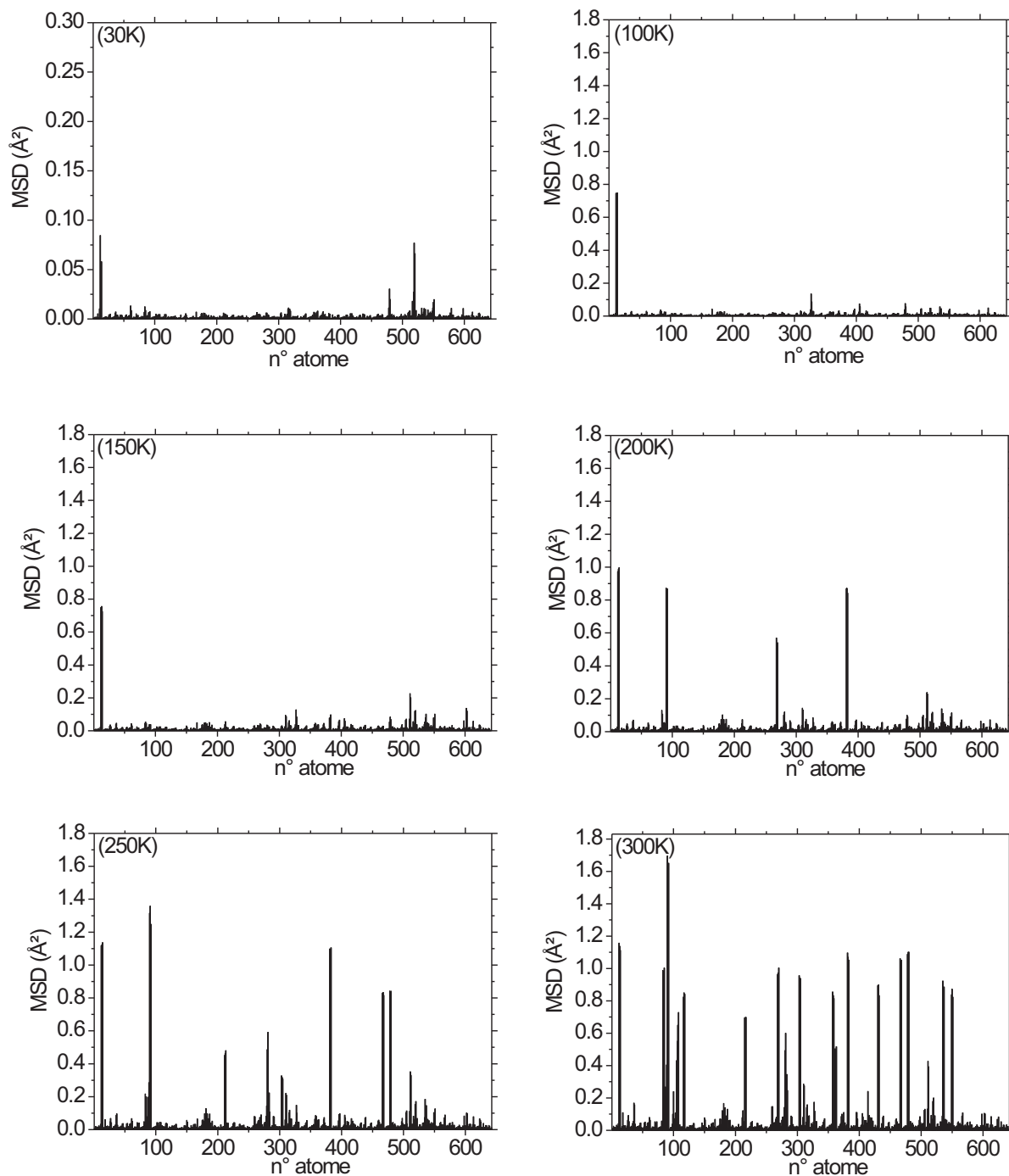


FIG. 2.20 – Variation de la mobilité des atomes avec la température de la crambine P22/L25 dans le vide sur une trajectoire de 1ns.

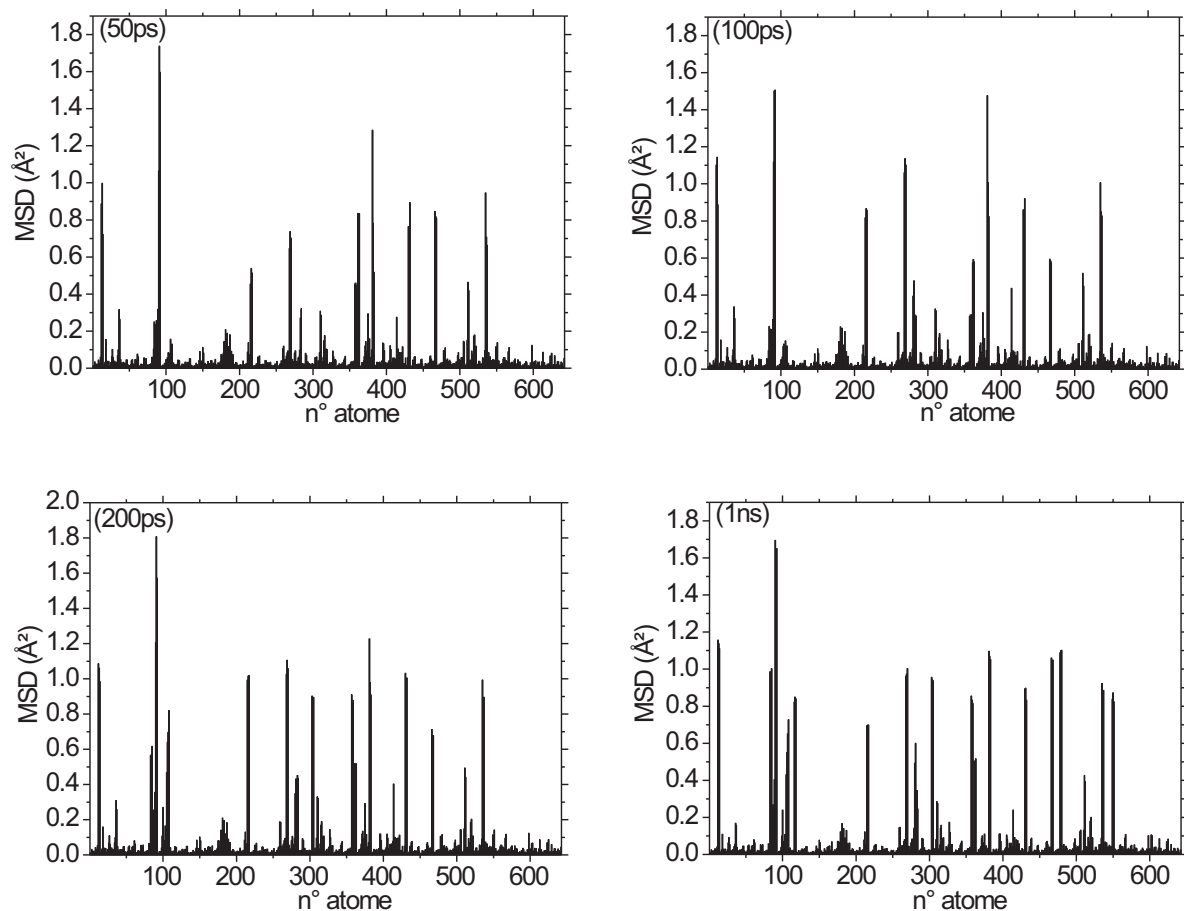


FIG. 2.21 – Variation de la mobilité des atomes avec le temps de simulation à 300K pour la crambine P22/L25 dans le vide sur une trajectoire de 1ns.

2.2.2.5 Mobilité des résidus

Afin de comparer la mobilité des résidus, nous avons défini un critère de sélection pour lequel les résidus seront considérés comme les plus mobiles, à savoir : tous les résidus dont leur $\frac{MSD}{T}$ sera supérieur au $\frac{MSD_{max}}{2T}$ seront considérés comme les plus mobiles (où MSD_{max} représente le MSD du résidu présentant la plus grande mobilité pour une protéine donnée). Dans le tableau suivant, nous répertorions les résidus les plus mobiles pour les différentes thionines.

La flexibilité de la crambine P22/I25 est due majoritairement au résidu ILE25 avec un $\frac{MSD}{T}=0.042\text{Å}^2/K$, puis la mobilité de ses autres résidus chute à $0.020\text{Å}^2/K$. VA3 possède de nombreux résidus très mobiles. Ces résidus correspondent à des fluctuations anharmoniques dont l'amplitude augmente plus vite que linéairement avec la température. Les résidus hydrophobes sont représentés en jaune, nous pouvons remarquer que la plupart des résidus mobiles pour les protéines déshydratées dans le tableau 2.8 sont des résidus hydrophobes (environ un quart de la séquence des crambines est composée de résidus hydrophobes, contre un sixième

Protéine	Résidus les plus mobiles							
P22/L25	THR1	ILE7	LEU25	ALA27	ILE34			
$\frac{MSD}{T} (\text{Å}^2/K)$	0.012	0.030	0.014	0.011	0.012			
S22/I25	ILE7	VAL8	VAL15	LEU18	THR30	ILE33	ILE35	THR39
$\frac{MSD}{T} (\text{Å}^2/K)$	0.021	0.019	0.017	0.027	0.013	0.023	0.019	0.012
P22/I25	ILE25	TYR29	THR30	ILE33	ILE34			
$\frac{MSD}{T} (\text{Å}^2/K)$	0.042	0.014	0.011	0.013	0.019			
S22/L25	ILE7	VAL15	LEU25	ILE33	PRO41			
$\frac{MSD}{T} (\text{Å}^2/K)$	0.036	0.023	0.013	0.014	0.019			
VA3	THR7	ILE12	LEU18	THR19	ALA21	LEU29	ILE35	
$\frac{MSD}{T} (\text{Å}^2/K)$	0.023	0.011	0.021	0.018	0.013	0.013	0.014	
VA2	PHE18	VAL25	LEU29	ILE34	ALA37			
$\frac{MSD}{T} (\text{Å}^2/K)$	0.020	0.022	0.012	0.023	0.014			
VB	ILE12	THR15	ARG25	LEU29	ILE35	PRO41		
$\frac{MSD}{T} (\text{Å}^2/K)$	0.012	0.011	0.022	0.011	0.013	0.012		

TAB. 2.8 – Résidus considérés comme mobiles pour les thionines déshydratées sur une trajectoire de 1ns. Les résidus hydrophobes sont représentés en jaune. La valeur du MSD/T correspond au MSD du résidu à $T=300K$ divisé par $300K$.

pour les viscotoxines, ce type de résidu n'est donc pas majoritaire dans les séquences).

Une étude par le logiciel CSU réalisée par Eran Eyal en 2003 [Eyal *et al.* 03] classa les résidus suivant leur mobilité. Cette approche prend en compte les atomes voisins dans l'espace et les distances interatomiques. Les structures proviennent de fichiers PDB de la Protein Data Bank, obtenues en diffraction X avec une résolution minimale de 2Å.

D'après lui, un type d'acide aminé donné présente une mobilité donnée. Par exemple, les acides aminés actifs sont très peu mobiles afin qu'ils gardent une configuration type, nécessaire à la fonction biologique. Ses travaux amenèrent les conclusions suivantes :

- Sont considérés comme mobiles : les résidus polaires (Glu, Gln, Arg, Lys), les résidus chargés, la serine et la valine.

- Sont considérés comme rigides : les acides aminés aromatiques (Tyr, Trp, Phe, His) et la cystéine (très peu mobile car "fixée" par les ponts disulfures).

Cette classification n'est pas concluante dans le vide. Elle sera plus en accord avec les protéines hydratées dans le chapitre suivant.

Sur la figure 2.22, nous avons représenté la mobilité des résidus par des histogrammes pour la crambine P22/L25 (ces mêmes courbes pour les autres thionines sont présentées en annexe I), où est calculé le nombre de résidus suivant leur valeur du $\frac{MSD}{T}$. Nous pouvons ainsi remarquer l'augmentation de la mobilité des résidus avec la température et de son hétérogénéité. Dès 100/130K, le nombre de résidus très mobiles va augmenter, prouvant ainsi l'activation de mouvements due à l'élévation de température. Dans le vide, ce phénomène est provoqué par l'activation des mouvements des méthyles. C'est ce que nous allons montrer au paragraphe suivant.

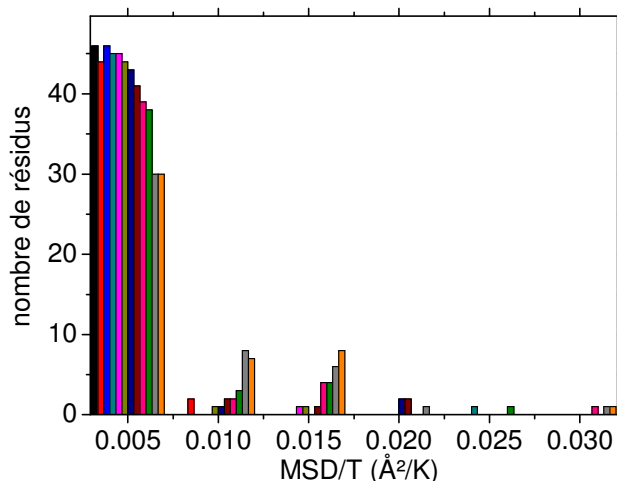


FIG. 2.22 – Variation du nombre de résidus suivant leur valeur du $\frac{MSD}{T}$ pour la crambine P22/L25 dans le vide, sur une trajectoire de 1ns. Les histogrammes correspondent aux températures suivantes : noir = 5K, rouge = 30K, bleu = 70K, bleu-vert = 100K, fushia = 130K, kaki = 150K, bleu marine = 170K, marron = 200K, rose = 220K, vert = 250K, gris = 280K, orange = 300K.

Sur la Fig. 2.23, nous représentons la variation du MSD pour chaque acide aminé en fonction de la température pour la crambine P22/L25 dans le vide. Ces courbes pour les autres thionines sont données en annexe I.

Cette représentation met clairement en évidence les mouvements activés thermiquement dès 100/130K. Il existe en effet trois types de comportements pour la variation du MSD : soit le résidu est harmonique, il aura alors une augmentation linéaire de sa mobilité avec la température, soit il subit une transition mise en évidence par une rupture de pente nette ou enfin, soit il a des "pics" de mobilité à différentes températures, dus à l'activation des mouvements des atomes.

Pour les protéines dans le vide, nous retrouvons le plus souvent ce dernier type de comportement, avec des "sauts" dans la variation de la mobilité des résidus avec la température. Par exemple, le résidu ILE7 de la crambine P22/L25 a un MSD de 6.09Å^2 à 220K, 5.18Å^2 à 250K et 9.01Å^2 à 300K. Tous les résidus avec un $MSD_{\max} < 0.50\text{Å}^2$ font partie des acides aminés qui ne présentent aucun mouvement, et donc aucune mobilité, créé par l'augmentation de la température.

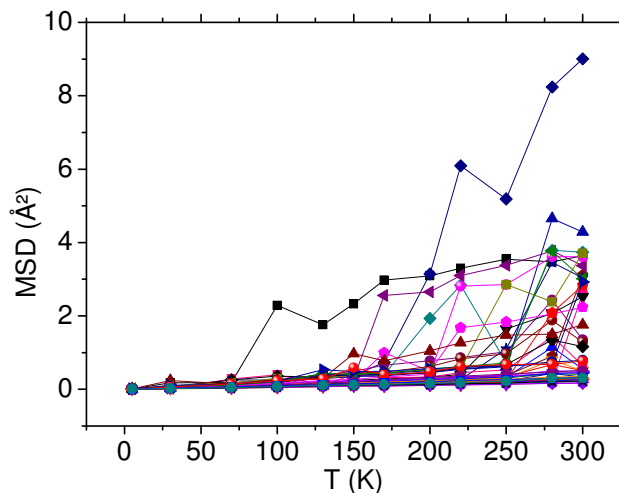


FIG. 2.23 – Variation du MSD de chaque acide aminé avec la température pour la crambine P22/L25 dans le vide, sur une trajectoire de 1ns.

2.2.2.6 Analyses des types de mouvements

Le comportement *global* de la protéine montre une augmentation de la flexibilité avec la température, mais sans une rupture de pente aussi marquée que dans les expériences ou dans nos calculs pour des protéines hydratées (cf chapitre suivant). L'étape suivante a été d'identifier les types de mouvements impliqués dans le passage de l'état R à l'état F.

Pour les sept thionines, nous avons analysé les mouvements que subissaient les atomes présentant une grande mobilité (représentée par des "pics" de mobilité sur la Fig. 2.20 pour P22/L25 et en annexe G pour les autres thionines) au moyen d'un logiciel de visualisation (VMD).

Deux types de mouvements seulement ressortent de cette étude : des mouvements de basculement et des mouvements de rotation (Fig. 2.24), ils impliquent des atomes d'hydrogène. L'augmentation de la flexibilité des protéines déshydratées est donc due à l'activation de ces mouvements impliquant des atomes d'hydrogène. Ce résultat confirme une étude réalisée par dynamique moléculaire par l'équipe de Curtis en 2004 sur la ribonucléase. Leur analyse démontra l'existence de ces mouvements de rotation et de basculement pour les méthyles [Curtis *et al.* 04].

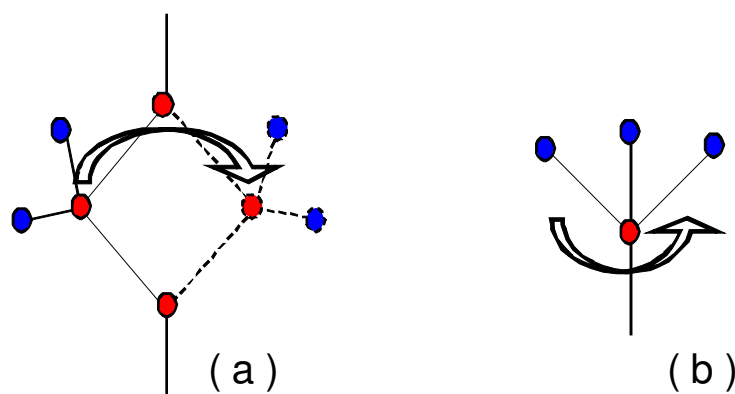


FIG. 2.24 – Mouvements possibles avec l'augmentation de la température, les sphères bleues représentent des hydrogènes, les sphères rouges des atomes de carbone : (a) basculement d'un côté à l'autre, (b) rotation des 3 hydrogènes autour du carbone.

Nous avons alors localisé par des sphères les hydrogènes présentant une grande mobilité, à différentes températures (cf Fig. 2.25). Cette étude a été réalisée par un programme fortran (cf annexe T), qui représente la mobilité d'un atome par une sphère. Le diamètre de la sphère sera proportionnel au $RMS (= \sqrt{MSD})$. Cette représentation nous donnera une indication sur une interaction possible entre plusieurs hydrogènes éloignés dans la séquence, mais proches dans l'espace, et la présence ou non de "couplages". De plus, les travaux de Bicout en 2001 montrent que les atomes des chaînes latérales ont une mobilité supérieure à ceux situés au coeur de la protéine [Bicout et Zaccai 01]. Nous pourrions ainsi, avec la figure 2.25, nous en assurer pour la famille des thionines.

En étudiant la figure 2.25, nous pouvons tout d'abord confirmer que l'augmentation de la température implique une augmentation de zones mobiles. Cette étude est faite pour la crambine P22/L25, mais les mêmes résultats ont été trouvés pour les autres thionines. Nous pouvons également remarquer que la plupart des hydrogènes mobiles se situent en surface de la protéine. En effet, les hydrogènes situés dans les chaînes latérales ont plus de "volume libre" permettant des mouvements de rotation et de basculement.

Sur la figure 2.26, nous avons représenté ces sphères pour les autres protéines, à 300K, sur une trajectoire de 1ns. Nous pouvons ainsi remarquer que les hydrogènes mobiles diffèrent d'une protéine à une autre. Une mutation va donc impliquer une distribution des mobilités totalement différente pour la protéine, et pas seulement au niveau de la mutation. Même si le comportement global de ces protéines montre peu de différences d'une protéine à une autre, l'étude détaillée de la mobilité nous permet de constater que les protéines se comportent de manière totalement différente. Certaines protéines ont de nombreux sites mobiles, comme pour la viscotoxine A2 et la crambine P22/I25, et d'autres beaucoup moins comme pour VB. Ceci confirme donc les résultats obtenus avec la figure 2.17, où VB présentait une flexibilité globale plus faible. Nous pouvons également conclure que l'augmentation de la flexibilité des protéines déshydratées est due majoritairement aux mouvements subis par les méthyles. Aucun autre type de mouvement pour les autres atomes n'existe. Il est difficile de déterminer la présence ou non d'une transition dynamique, mais un phénomène dynamique existe bien dans le vide, il consiste en l'activation des mouvements des méthyles et des groupes $C-H$ des cycles aromatiques dès $100 - 130K$.

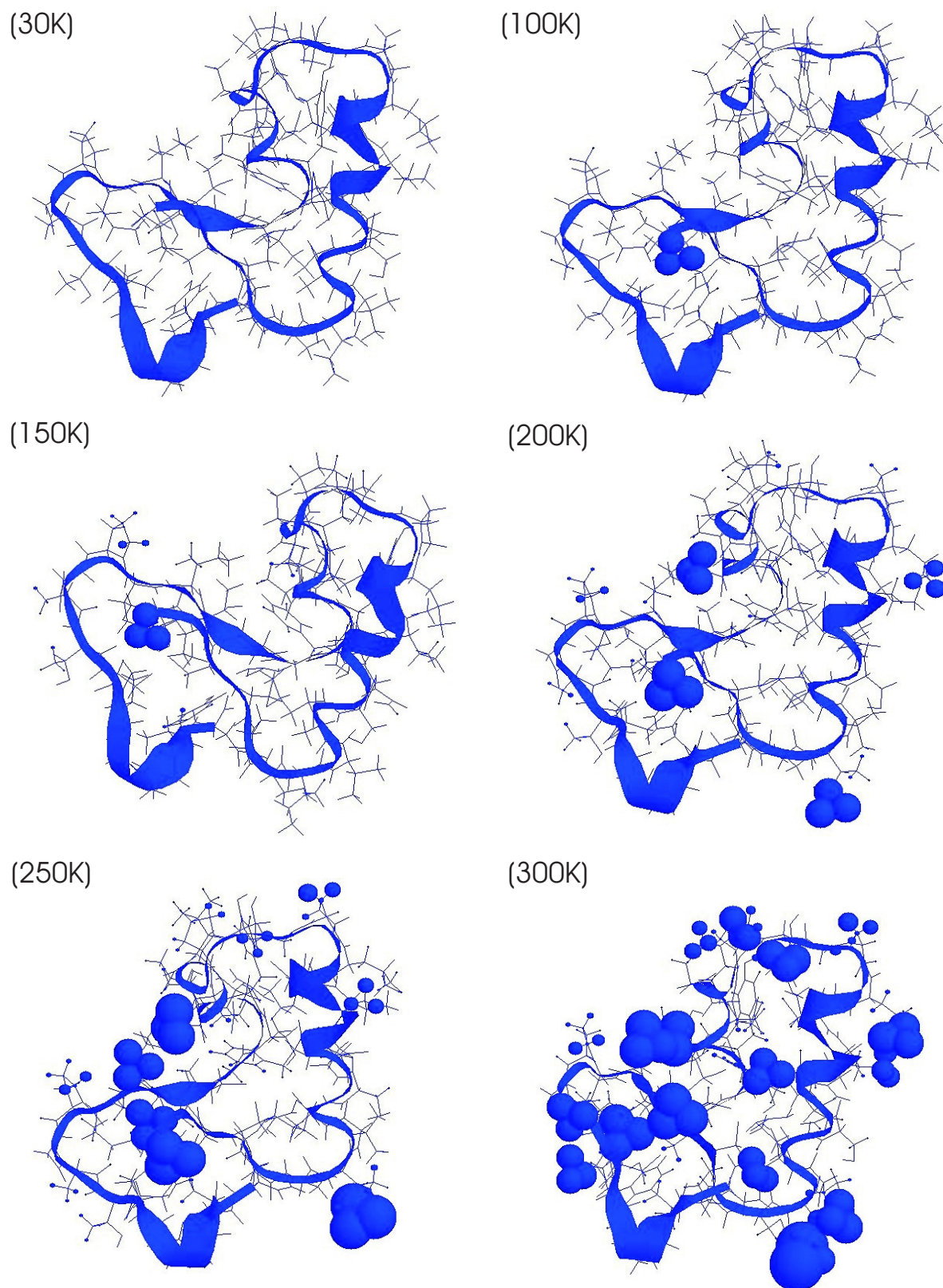


FIG. 2.25 – Localisation des hydrogènes mobiles par des sphères, de diamètre proportionnel à leur RMS, pour crambine P22/L25 déshydratée, sur une trajectoire de 1ns.

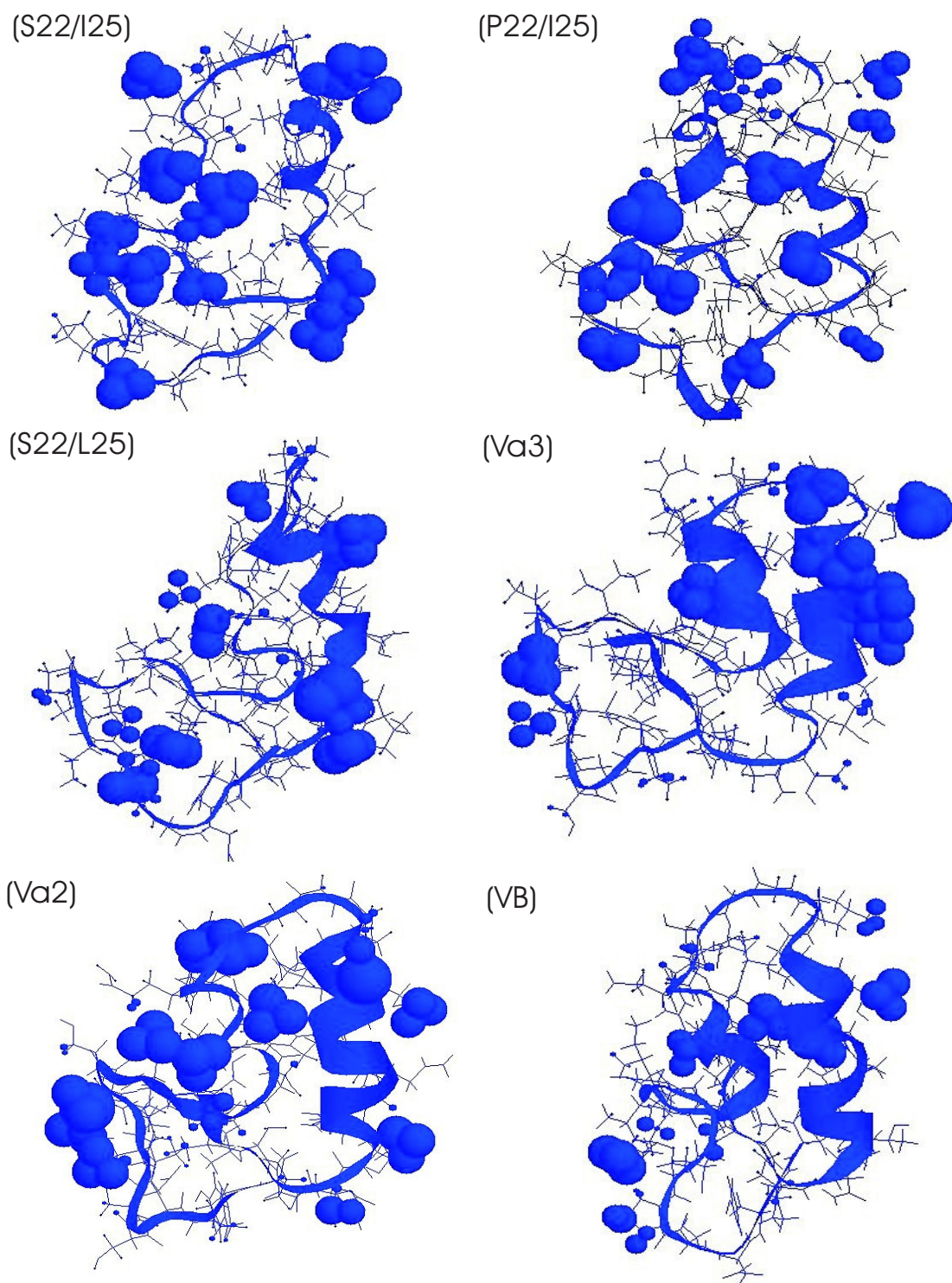


FIG. 2.26 – Localisation des hydrogènes mobiles par des sphères, de diamètre proportionnel à leur RMS, pour les thionines déshydratées à 300K, sur une trajectoire de 1ns.

2.2.2.7 Barrières d'énergie de rotation des groupes méthyles

Afin de comprendre l'inhomogénéité de la mobilité des atomes d'hydrogène visibles sur les figures 2.25 et 2.26, nous avons étudié la distribution des barrières d'énergie de rotation des groupements méthyles. Chaque groupe méthyle a une symétrie de rotation d'ordre 3 et on s'attend à observer 3 minima d'énergie, tous les 120 degrés, séparés par des barrières d'énergie de rotation.

Pour calculer les barrières d'énergie de rotation des méthyles, nous avons considéré plusieurs structures : la structure obtenue après minimisation et la structure moyenne calculée sur la trajectoire à différentes températures. Sur ces structures, nous avons appliqué une rotation de 1 degré sur le premier méthyle puis avec le champ de force Charmm, nous avons calculé l'énergie totale. Nous avons ensuite appliqué une nouvelle rotation de 1 degré, calculé l'énergie et ainsi de suite jusqu'à 360°. Cette étude a été faite sur tous les méthyles.

Un exemple caractéristique des résultats obtenus est représenté sur la figure 2.27 pour une structure de la crambine P22/L25 optimisée par la méthode des gradients conjugués.

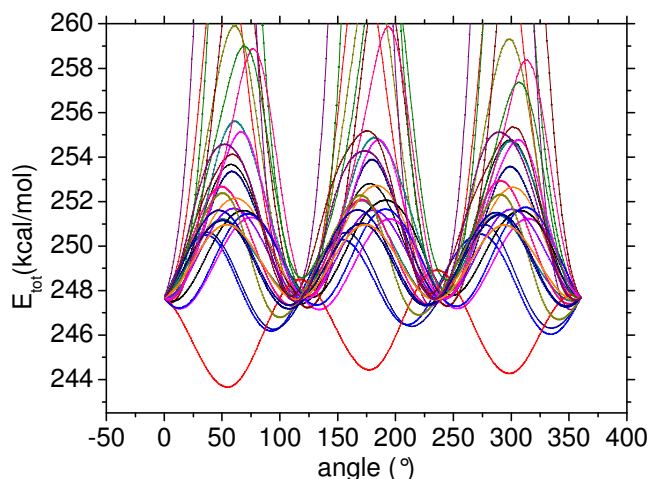


FIG. 2.27 – Barrières d'énergie de rotation des méthyles de la crambine P22/L25 optimisée.

La distribution des barrières d'énergie pour cette structure optimisée est représentée sur la figure 2.28.

Pour chaque méthyle, nous avons calculé la barrière d'énergie moyenne sur les trois barrières d'énergie de rotation. Le plus grand nombre de barrières est autour de $8 - 9 \text{ kcal/mol}$. Les calculs effectués sur les structures à différentes températures ne montrent pas de changements significatifs de barrières d'énergie, ce qui signifie que la structure évolue peu avec la température (il s'agit de l'énergie potentielle et non de l'énergie libre dans nos calculs). Ces résultats sont donnés en annexe K. Par conséquent, la distribution des barrières d'énergie évolue également peu avec la température.

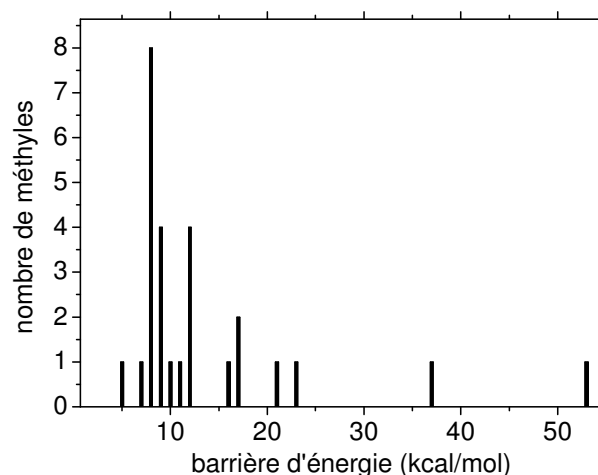


FIG. 2.28 – Distribution des barrières d'énergie pour la crambine P22/L25 optimisée.

Sur la Fig. 2.29, nous avons représenté l'évolution du MSD moyen des groupements méthyles (calculés comme la moyenne sur ses 3 atomes d'hydrogène) en fonction de la hauteur moyenne de la barrière d'énergie de rotation. On constate une très bonne corrélation entre la mobilité observée et la hauteur de barrière : la mobilité décroît significativement avec la hauteur de barrière comme le prévoit la loi exponentielle d'Arrhenius. Ce résultat montre aussi que la mobilité des méthyles est surtout dominée par l'énergie potentielle (utilisée dans nos calculs).

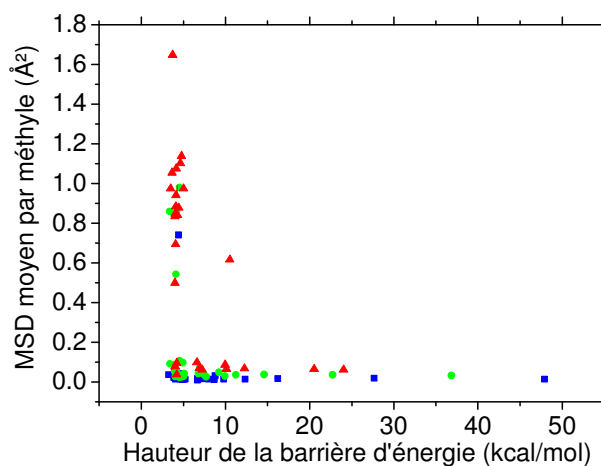


FIG. 2.29 – Evolution du MSD moyen des groupements méthyles (calculés comme la moyenne sur ses 3 atomes d'hydrogène) en fonction de la hauteur moyenne de la barrière d'énergie de rotation. Les carrés bleus correspondent à 100K, les ronds verts à 200K et les triangles rouges à 300K.

2.2.3 Origine des effets induits par les mutations

D'après toutes les études développées précédemment, nous pouvons conclure qu'une seule mutation va modifier toute la dynamique de la protéine. Le comportement global de la flexibilité paraît être semblable d'une crambine à une autre, mais l'étude détaillée de la distribution de cette mobilité nous a permis de constater qu'au contraire, chaque atome va réagir différemment. Une mutation n'implique donc pas des conséquences seulement dans son environnement mais tout au long de la séquence. De plus, le type de résidu impliqué dans la mutation a une très grande importance.

Nous avons pu voir que les deux crambines ayant une Proline en 22^{ième} position sont dynamiquement semblables, de même pour celles ayant une Serine en 22^{ième} position. La mutation du 25^{ième} résidu semble n'avoir que peu de conséquences, ceci peut s'expliquer par le fait que l'on remplace une Leucine par une Isoleucine, deux acides aminés très proches.

En ce qui concerne la dynamique des viscotoxines, VB semble se différencier de celle des deux autres viscotoxines. En étudiant le tableau 4.3, où sont répertoriées les séquences des protéines, nous avons pu constater que de nombreuses mutations existent entre les viscotoxines. Il est alors plus difficile d'évaluer les modifications dues à telle ou telle mutation.

Afin de comprendre comment une mutation modifie la dynamique de la protéine, nous avons étudié le réseau de liaisons hydrogène des thionines. Chaque liaison hydrogène est définie entre un atome d'oxygène-accepteur et un atome d'hydrogène-donneur, avec les critères suivants :

- Si $O - H < 2.4\text{\AA}$, il existe une liaison hydrogène,
- Si $O - H > 2.4\text{\AA}$, il n'y a pas de liaison.

Dans le tableau 2.9, nous répertorions les liaisons hydrogène des crambines déshydratées, à 300K. Les chiffres correspondent aux numéros des acides aminés dans la séquence (cf tableau 4.3).

P22/L25							
2-2	3 - 33	3 - 33	7 - 11	8 - 12	9 - 13	12 - 15	13 - 17
22 - 24	30-32	3 - 33	34-34	35-37	37-39	40-40	42-45
S22/I25							
4-4	7 - 10	9 - 12	13 - 17	14 - 18	18-20	23-26	24-26
25-27	30-27	32-27	33-33	35-37			
P22/I25							
4-44	4-44	4-6	6-6	6 - 10	9 - 6	10 - 13	13-11
13 - 16	13 - 17	20-17	27 - 24	32-30	39-37	39-39	
S22/L25							
3-33	4-6	7 - 10	9 - 12	10 - 13	16 - 13	19-21	24-26
25-27	32-27	34-34	35-35	46-44			

TAB. 2.9 – Réseaux des liaisons hydrogène des crambines déshydratées à 300K. Les liaisons hydrogène correspondants aux hélices sont en jaune, celles des feuillets en bleu.

D'après ce tableau, nous pouvons remarquer qu'une seule mutation provoque une modification du réseau global des liaisons hydrogène. De la même façon, nous avons étudié les liaisons hydrogène des viscotoxines dans le tableau 2.10.

VA3						
3 – 33	3 – 33	4-4	7 – 10	7 – 11	11 – 13	11 – 14
13 – 16	41-44	13 – 17	14 – 18	16 – 21	26 – 30	27 – 32
34 – 34	35-35	40-40				
VA2						
3 – 33	4-4	4-44	4-46	6-6	6 – 10	8 – 12
9 – 13	11 – 14	11 – 15	14 – 17	14 – 18	15 – 19	24 – 27
24 – 28	30 – 32	34-34	35-35	37-37	40-40	44-42
VB						
2-2	3-3	6 – 9	6 – 10	7 – 11	12 – 16	16 – 18
16 – 19	22 – 26	24 – 27	25 – 28	27 – 30	36-36	39-39
40-42	42-42	44-46				

TAB. 2.10 – Réseaux des liaisons hydrogène des viscotoxines déshydratées à 300K. Les liaisons hydrogène correspondants aux hélices sont en jaune, celles des feuillets en bleu.

Les mêmes conclusions peuvent être tirées : les mutations induisent un réseau de liaisons hydrogène totalement différent d'une protéine à une autre. Cette modification a très certainement un impact sur la dynamique des protéines.

Le réseau de liaisons hydrogène est dépendant de la température. Dans le tableau suivant, nous répertorions les liaisons hydrogène de la crambine S22/I25 à 100K. En comparant avec le tableau 2.9, nous pouvons remarquer que l'augmentation de la température va provoquer des modifications dans le réseau : deux liaisons existent à 300K et pas à 100K : 14-18 et 23-26, et trois liaisons existent à 100K et pas à 300K : 10-12, 13-16 et 40-42. Des modifications existent également pour les autres thionines entre 100 et 300K.

Cependant, ces modifications n'impliquent pas des résidus considérés comme mobiles, d'après le tableau 2.8.

S22/I25						
4-4	7-10	9-12	10-12	13-16	13-17	18-20
24-26	25-27	27-30	27-32	33-33	35-37	40-42

TAB. 2.11 – Réseaux des liaisons hydrogène de la crambine S22/I25 dans le vide à 100K.

2.3 Conclusion

Nous pouvons conclure qu'il est difficile d'établir ou non l'existence d'une réelle transition dynamique dans le vide pour les thionines. Cependant, un phénomène existe bien puisque les protéines présentent les deux états R et F et donc une anharmonicité due entièrement aux atomes d'hydrogène. En effet, l'étude des constantes de force a montré que les atomes lourds ne contribuent pas à ces modifications de dynamique. Les méthyles sont majoritairement responsables de l'augmentation de la mobilité de la protéine. Les deux types de mouvements qu'ils subissent, des rotations et des basculements, procurent à la protéine une certaine flexibilité. Cette flexibilité est due à des mouvements diffusifs activés thermiquement autour de 150K.

Nous avons également remarqué qu'une seule mutation modifie la dynamique globale de la protéine. Les modifications de la dynamique dépendent du type de résidu muté. Dans notre étude, les protéines avec une proline en 22ième résidu ont un rayon de giration semblable. De même, les protéines avec une serine en 22ième position ont un rayon de giration évoluant de la même manière avec la température.

Le réseau de liaisons hydrogène est fortement perturbé par une seule mutation. Nous n'avons pas trouvé de corrélation simple entre ces modifications et la localisation des résidus mobiles. La rotation des groupes méthyles semble peu affectée, du moins dans le vide, par cette réorganisation du réseau.

Dynamique des protéines hydratées : rôle de l'eau

3.1 Résultats pour la transition dynamique des crambines et des viscotoxines

3.1.1 Dépendance par rapport aux conditions initiales

Nous allons tout d'abord étudier les protéines hydratées avec un taux $h = 4$. Comme nous l'avons vu dans le chapitre 4, partie I, nous avons utilisé 5 conditions initiales différentes afin d'évaluer la dépendance de la mobilité de la protéine hydratée à ces conditions.

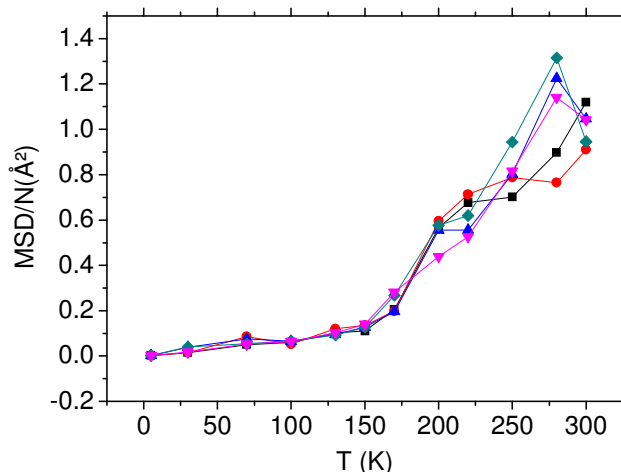


FIG. 3.1 – Variation de l'écart quadratique moyen de P22/L25 hydratée sur une trajectoire de 1ns. Les carrés noirs correspondent à la 1ère condition initiale, les ronds rouges à la 2nde, les triangles bleus à la 3ième, les losanges verts à la 4ième et les triangles roses à la 5ième.

Nous pouvons constater sur la Fig. 3.1 que l'écart quadratique moyen varie "légèrement" en fonction des conditions initiales utilisées, mais uniquement dans la partie anharmonique. En effet, pour $T < 150K$, les différences obtenues sont négligeables. C'est pourquoi, pour les autres thionines, les simulations avec 5 autres conditions initiales ont été réalisées pour $150 \leq T \leq 300K$ (cf annexe L). En utilisant 5 trajectoires différentes, nous pouvons calculer un comportement "moyen" du MSD de la protéine avec une moyenne sur ces 5 trajectoires et évaluer la reproductibilité de nos résultats.

3.1.2 Paramètres tirés des modèles

3.1.2.1 Constantes Effectives

Nous avons tout d'abord calculé les constantes de force de la crambine P22/L25 hydratée pour chacune des 5 conditions initiales représentées sur la Fig. 3.1 de la même façon qu'au chapitre précédent. La transition dynamique des protéines hydratées intervient vers 170/200K, nous avons alors défini l'état R pour $0 \leq T \leq 150K$ et l'état F pour $170 \leq T \leq 300K$.

	$K_R(N/m)$	$K_F(N/m)$	$\Delta S_{FR}(kcal/mol)$
n°1	5.7	0.7	6.3
n°2	4.8	0.9	4.9
n°3	5.2	0.6	6.5
n°4	5.4	0.6	6.4
n°5	5.1	0.6	6.3

TAB. 3.1 – Paramètres effectifs de la crambine P22/L25 hydratée pour les 5 conditions initiales.

Nous pouvons remarquer qu'en général, les différences obtenues dans les valeurs des constantes de force sont très faibles d'une condition initiale à une autre.

Dans la suite de notre étude, nous avons calculé les moyennes des écarts quadratiques moyens de chacune des thionines et en traçant les régressions linéaires aux deux états, R et F, nous avons calculé les paramètres effectifs, cf Fig. 3.2 et 3.3.

	$K_R(N/m)$	$K_F(N/m)$	$\Delta S_{FR}(kcal/mol)$	$\frac{MSD}{N}$ (Résidu 22) (\AA^2)
P22/L25	5.2	0.7	6.1	0.5
S22/I25	4.1	0.3	8.2	2.9
P22/I25	4.7	0.3	7.9	1.1
S22/L25	4.2	0.3	7.6	2.6
VA3	5.1	0.5	6.8	1.1
VA2	4.4	0.3	7.8	3.9
VB	4.6	0.3	8.3	2.1

TAB. 3.2 – Paramètres effectifs des thionines hydratées calculés sur la moyenne du MSD sur les 5 trajectoires générées et MSD du 22ième résidu.

Nous pouvons remarquer que la crambine P22/L25 est la plus rigide des thionines, et S22/I25 et VB sont les plus flexibles. Nous pouvons dès lors constater les modifications dues à l'hydratation. En effet, VB était la thionine la plus rigide dans le vide et elle devient une des plus flexibles une fois hydratée. Inversement, la crambine P22/L25 était la plus flexible dans le vide et devient la plus rigide dans l'eau.

En comparant avec le tableau 2.3, nous pouvons constater que l'hydratation revient à diviser les constantes effectives par un facteur 8 dans l'état R et par un facteur qui peut varier de 10 pour la crambine P22/L25 à 50 pour VB dans l'état F. L'hydratation implique une nette augmentation de la flexibilité de la protéine et elle aura des conséquences différentes d'une thionine à une autre. En comparant les séquences (chapitre 4, partie I), on observe que VA3 a aussi une proline en 22ième position, mais pas VA2 et VB (qui ont une serine). Ce résidu est probablement important pour la détermination de K_R et K_F puisque les valeurs des constantes de force entre VA3 et P22/L25 sont proches dans les états R et F. Le résidu Pro22 est situé dans la boucle et sert à bloquer les zones en hélices avec son cycle, il est donc impliqué dans le mode de plus grande amplitude (cf "Analyses des types de mouvements", partie 3.2.4). En comparant les valeurs du $\frac{MSD}{N}$ du résidu en 22ième position dans différentes protéines à 300K, nous constatons que le $\frac{MSD}{N}$ des PRO22 est plus faible que celui des SER22 ; par exemple, le $\frac{MSD}{N}$ est de 0.5\AA^2 pour PRO22 dans la crambine P22/L25 et vaut 2.9\AA^2 pour SER22 dans la protéine VA2.

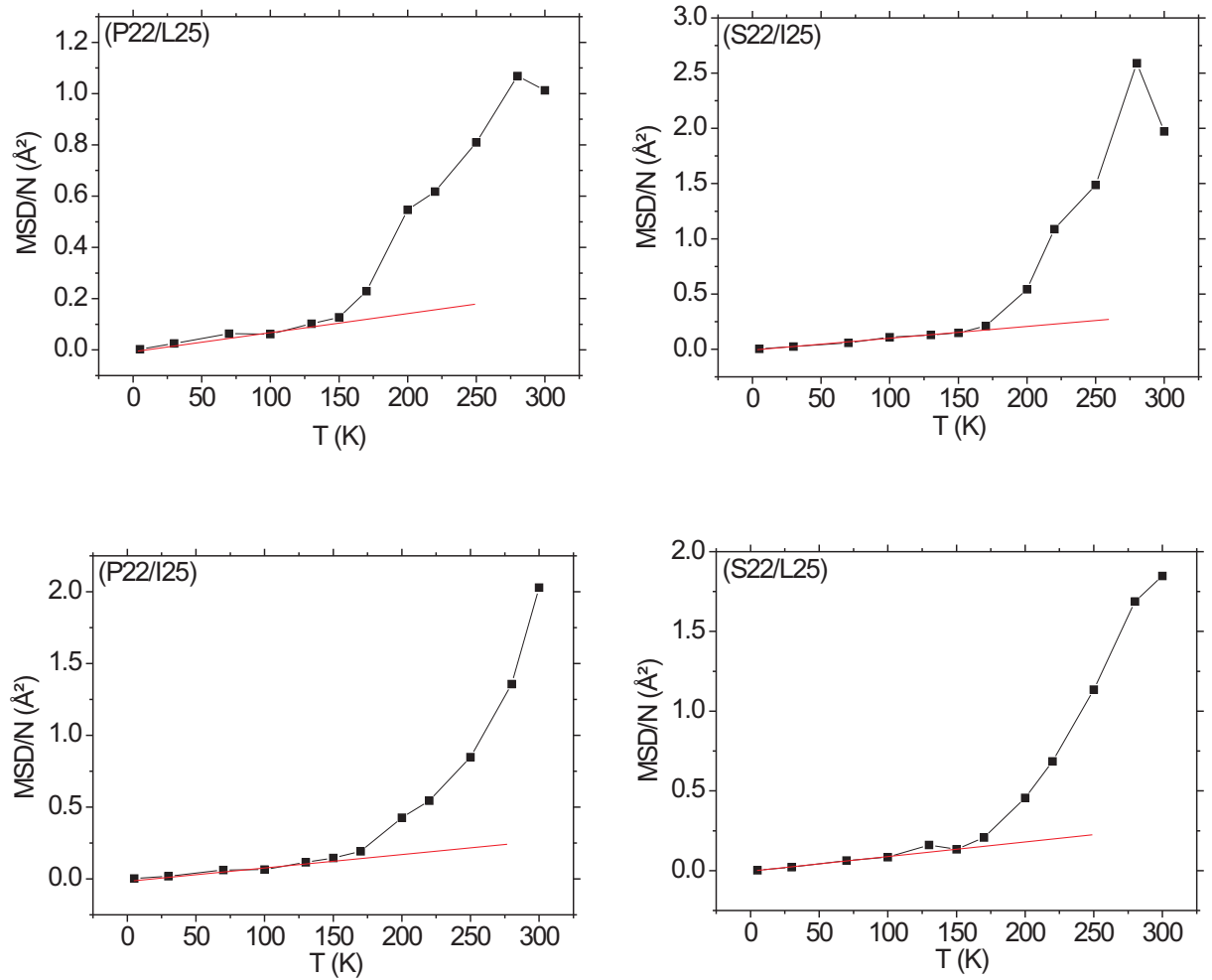


FIG. 3.2 – Evolution de la moyenne de l'écart quadratique moyen des crambines hydratées, calculée sur 5 conditions initiales. Les lignes rouges en traits interrompus correspondent aux régressions linéaires de la courbe pour l'état R.

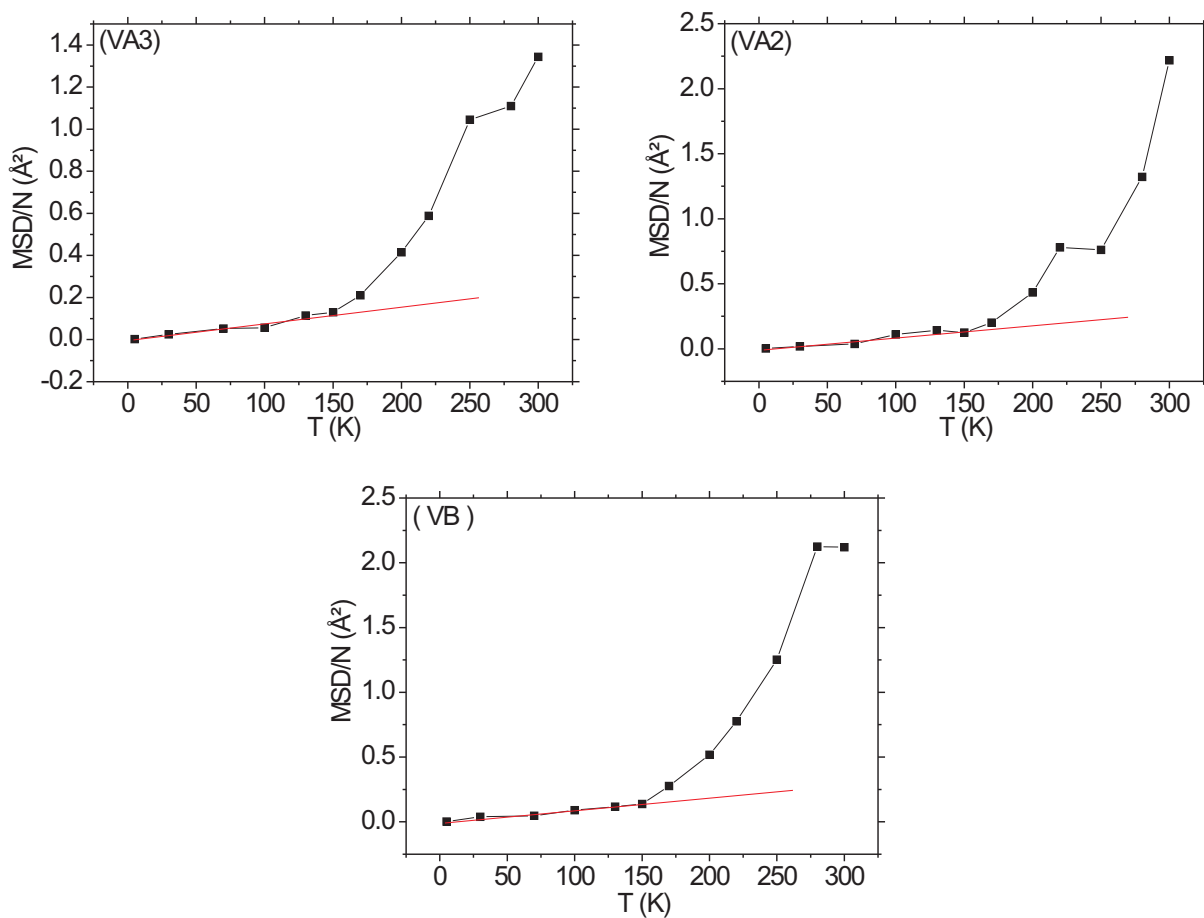


FIG. 3.3 – Evolution de la moyenne de l'écart quadratique moyen des viscotoxines hydratées, calculée sur 5 conditions initiales. Les lignes rouges en traits interrompus correspondent aux régressions linéaires de la courbe pour l'état R.

3.1.2.2 Comparaisons des courbes moyennes générales

D'après la Fig. 3.4, nous pouvons constater, tout d'abord pour les crambines, que P22/L25 a la flexibilité la plus faible (avec un $\frac{MSD}{N}$ max $\simeq 1.0\text{\AA}^2$), S22/I25 la plus grande (avec un $\frac{MSD}{N}$ max $\simeq 2.6\text{\AA}^2$ à $280K$) et les deux mélanges P22/I25 et S22/L25 une flexibilité intermédiaire. Nous pouvons ainsi remarquer qu'une seule mutation implique une dynamique globale différente également dans l'eau. Les ordres de grandeur du $\frac{MSD}{N}$ sont les mêmes entre les crambines et les viscotoxines. VA3 est la moins flexible des viscotoxines (avec un $\frac{MSD}{N}$ max $\simeq 1.4\text{\AA}^2$) et VB la plus flexible avec un $\frac{MSD}{N}$ max $\simeq 2.3\text{\AA}^2$). A $300K$, VA2 et VB ont des valeurs de leur flexibilité très semblables. Dans le cas des protéines hydratées, la transition dynamique est très prononcée, avec une rupture de pente très marquée pour $T_d \sim 170 - 200K$.

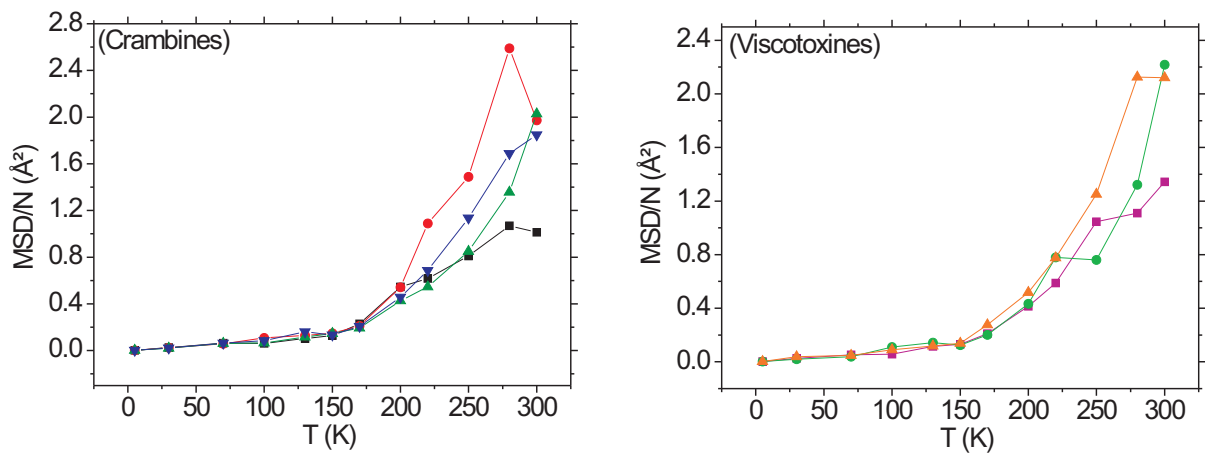


FIG. 3.4 – A gauche, Flexibilité globale des crambines hydratées (P22/L25 : carrés noirs, S22/I25 : ronds rouges, P22/I25 : triangles bleus, S22/L25 : triangles verts), et à droite des viscotoxines hydratées (VA3 : carrés roses, VA2 : ronds verts, VB : triangles oranges). Moyenne du MSD calculée sur 5 trajectoires de 1ns.

3.1.3 Séparation des contributions en type d'atomes

Nous avons calculé les moyennes des écarts quadratiques moyens des atomes d'hydrogène et de tous les autres types d'atomes, et nous avons calculé les constantes effectives des états R et F avec la méthode des interpolations linéaires, cf Fig. 3.5 et 3.6.

Les constantes effectives des hydrogènes sont données dans le tableau 3.3, et celles des autres types d'atomes dans le tableau 3.4. Les valeurs des constantes K_R et K_F pour P22/L25 et VA3 sont proches à la fois pour les atomes d'hydrogènes et pour les atomes lourds. Si on les compare aux tableaux 2.4 et 2.5 au chapitre précédent des mêmes protéines dans le vide, on constate que les constantes effectives de ces différents types d'atomes ont fortement diminué avec l'hydratation. La protéine la plus flexible est S22/I25, les plus rigides sont VA3 et P22/L25 quelque soit le type d'atome utilisé pour calculer les constantes de force effectives.

Les tableaux 3.3-3.5 démontrent que tous les types d'atomes, y compris les atomes du backbone, contribuent à la transition dynamique de la protéine. Dans le vide, seuls les hydrogènes jouaient un rôle dans la dynamique anharmonique de la protéine. Ce n'est plus le cas pour les protéines hydratées.

	$K_R(N/m)$	$K_F(N/m)$	$\Delta S_{FR}(kcal/mol)$
P22/L25	3.9	0.5	6.1
S22/I25	3.2	0.2	8.1
P22/I25	3.6	0.3	7.8
S22/L25	3.2	0.3	7.4
VA3	3.8	0.4	7.1
VA2	3.4	0.2	7.8
VB	3.5	0.2	8.4

TAB. 3.3 – Paramètres effectifs des thionines hydratées pour les atomes d'hydrogène, calculés sur la moyenne des 5 trajectoires.

	$K_R(N/m)$	$K_F(N/m)$	$\Delta S_{FR}(kcal/mol)$
P22/L25	7.5	0.9	6.1
S22/I25	5.9	0.3	8.8
P22/I25	6.8	0.4	8.3
S22/L25	6.1	0.4	7.9
VA3	7.8	0.7	7.2
VA2	6.1	0.4	7.9
VB	6.7	0.3	8.9

TAB. 3.4 – Paramètres effectifs des thionines hydratées pour tous les atomes qui ne sont pas des hydrogènes, calculés sur la moyenne des 5 trajectoires.

	$K_R(N/m)$	$K_F(N/m)$	$\Delta S_{FR}(kcal/mol)$
P22/L25	13.2	1.5	6.4
S22/I25	8.8	0.4	9.1
P22/I25	10.2	0.6	8.4
S22/L25	9.1	0.5	8.6
VA3	11.4	1.1	6.9
VA2	9.5	0.5	8.7
VB	10.2	0.5	8.9

TAB. 3.5 – Paramètres effectifs des thionines hydratées pour les carbones alpha, calculés sur la moyenne des 5 trajectoires.

Dans le cas des protéines hydratées, la contribution des hydrogènes au MSD global reste la plus importante. La courbe 3.7 confirme également ce résultat pour la crambine P22/L25 (mais le même type de courbes a été obtenu pour les autres thionines, cf annexe M) où sont représentés les écarts quadratiques moyens de la protéine P22/L25, de ses atomes d'hydrogène et de tous les autres atomes la constituant. Nous pouvons ainsi remarquer que les fluctuations des hydrogènes dominent la flexibilité de la protéine.

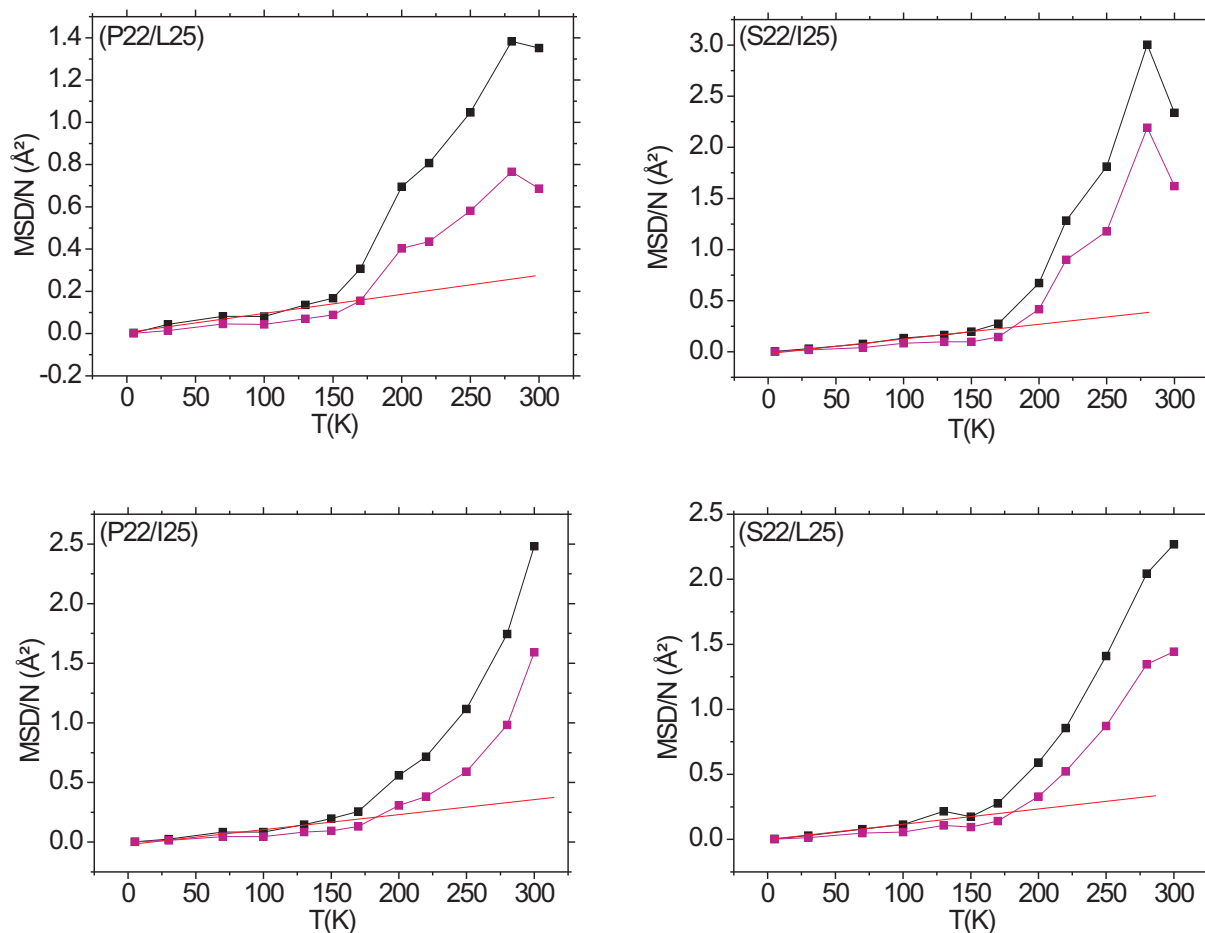


FIG. 3.5 – Evolution de l'écart quadratique moyen des différents types d'atomes pour les crambines hydratées, moyennes calculées sur 5 trajectoires de 1ns. La courbe rose représente l'écart quadratique moyen de tous les atomes qui ne sont pas des hydrogènes, la courbe noire représente cette même variation pour les hydrogènes seulement et la ligne rouge en traits interrompus est la régression linéaire de la courbe dans l'état R.

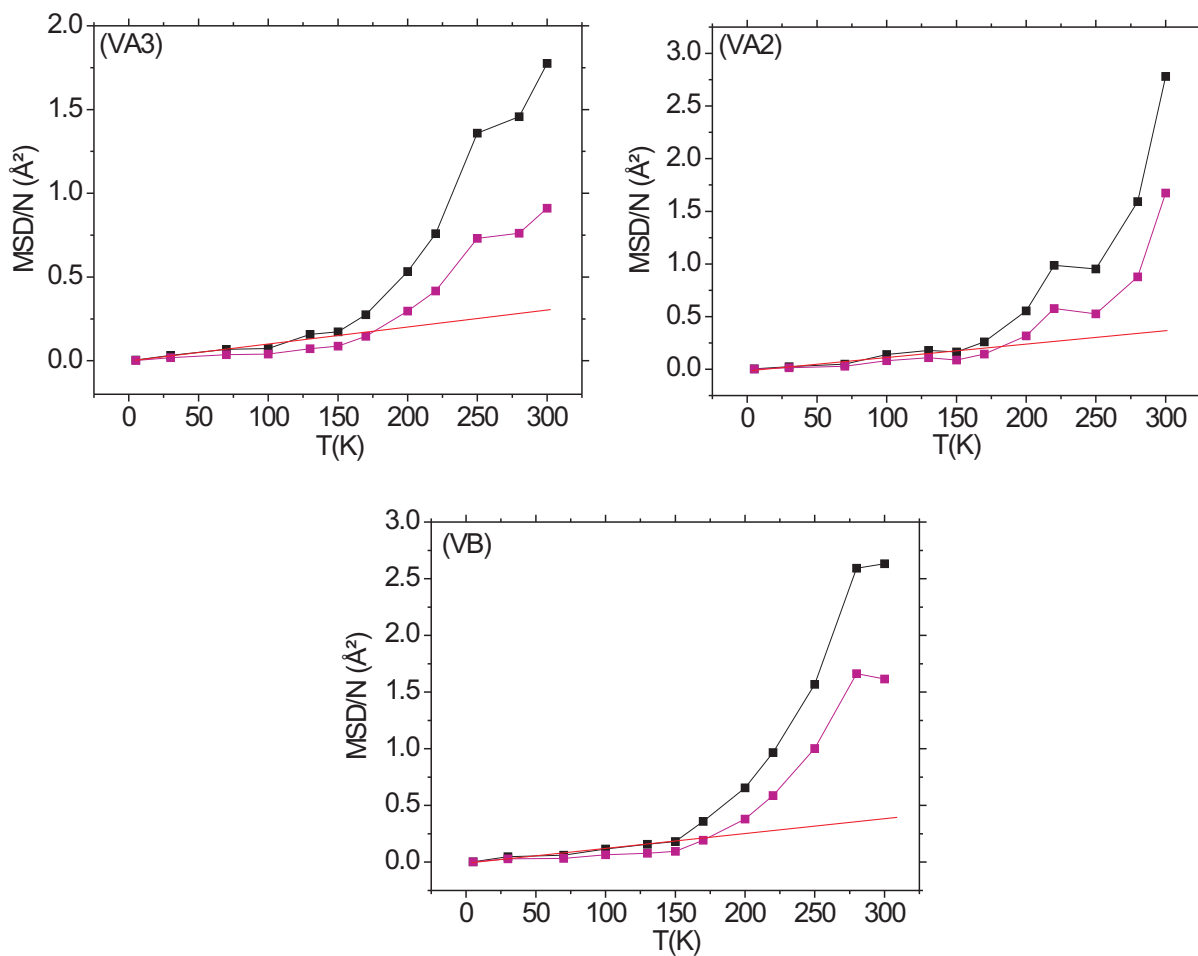


FIG. 3.6 – Evolution de l'écart quadratique moyen des différents types d'atomes pour les viscotoxines hydratées, moyennes calculées sur 5 trajectoires de 1ns. La courbe rose représente l'écart quadratique moyen de tous les atomes qui ne sont pas des hydrogènes, la courbe noire représente cette même variation pour les hydrogènes seulement et la ligne rouge en traits interrompus est la régression linéaire de la courbe dans l'état R.

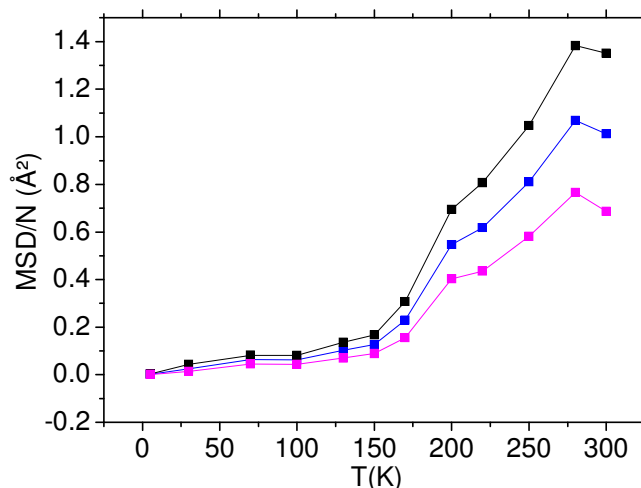


FIG. 3.7 – Evolution de l'écart quadratique moyen de la crambine P22/L25 (courbe bleue), de ses atomes d'hydrogène (courbe noire) et de tous ses autres atomes (courbe rose), hydratée et sur la moyenne des 5 trajectoires de 1ns.

3.1.4 Dépendance en temps de simulations

De la même façon que dans le chapitre 2, partie II, nous avons représenté les variations de l'écart quadratique moyen de P22/L25 hydratée pour les 4 temps de production étudiés : 50ps, 100ps, 200ps et 1ns. Les résultats sont représentés sur la Fig. 3.8 pour la condition initiale n°1. Nous pouvons remarquer que la flexibilité augmente avec le temps de production, comme dans le vide. Plus la trajectoire sera longue, plus la simulation pourra prendre en compte les différents types de mouvements que subit la protéine. Les mêmes résultats ont été réalisés pour les autres thionines (cf annexe N).

Nous avons également calculé les constantes effectives pour ces différents temps de simulations en traçant les régressions linéaires dans l'état R et F. Les conclusions sont les mêmes que pour les protéines dans le vide : les constantes effectives dans la partie harmonique sont indépendantes du temps de simulation, alors que dans la partie anharmonique, un temps de simulation plus long implique une flexibilité plus grande, et donc une constante effective plus petite.

	$K_R(N/m)$	$K_F(N/m)$	$\Delta S_{FR}(kcal/mol)$
50ps	7.2	1.4	4.9
100ps	6.9	1.9	3.9
200ps	6.7	0.8	6.5
1ns	5.7	0.7	6.3

TAB. 3.6 – Constantes effectives de la crambine P22/L25 hydratée en fonction du temps de simulation.

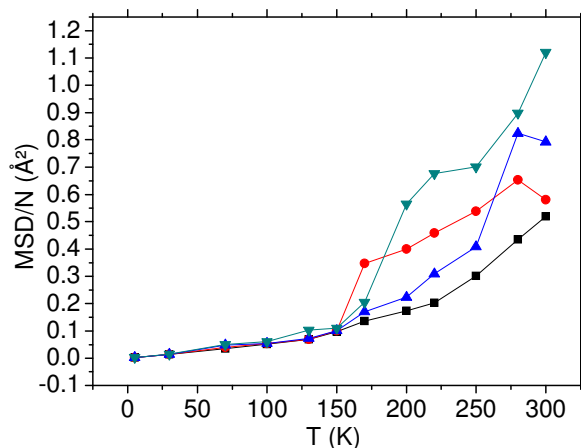


FIG. 3.8 – Dépendance en temps de simulations de l'écart quadratique moyen de la crambine P22/L25 hydratée, la courbe noire représente une trajectoire de 50ps, la courbe rouge de 100ps, la courbe bleue de 200ps et la courbe verte de 1ns.

3.1.5 Comparaisons vide et hydratation

Sur la Fig. 3.9, nous avons comparé les variations de l'écart quadratique moyen de deux thionines : P22/L25 à gauche et VA3 à droite, simulées dans le vide et dans une boîte d'eau. Cette courbe nous montre l'augmentation de la flexibilité globale de la protéine due à l'hydratation mais aussi l'apparition d'une rupture de pente plus nette à la température de transition. (Les mêmes conclusions ont été tirées pour les autres thionines).

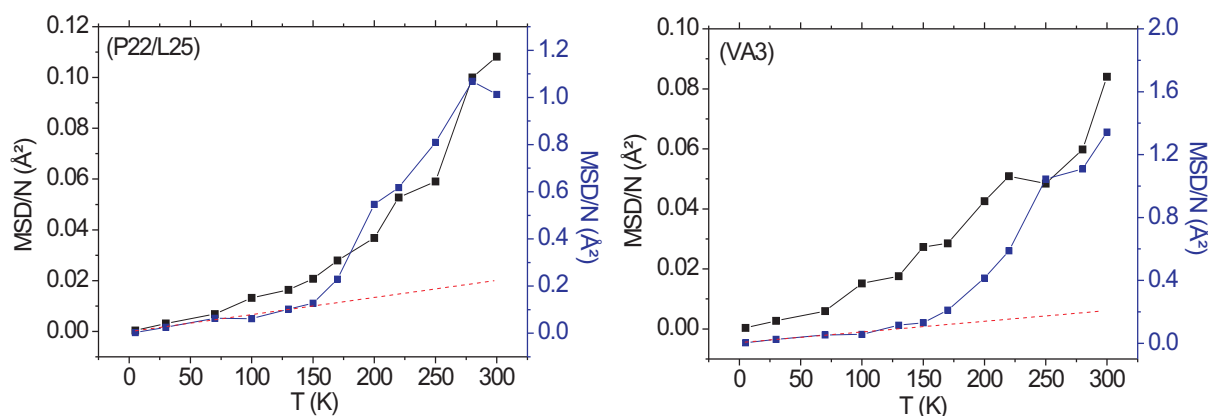


FIG. 3.9 – Comparaisons de la variation du MSD des protéines simulées dans le vide (courbes noires) et dans une boîte d'eau (courbes bleues) pour P22/L25 à gauche et VA3 à droite. Les lignes rouges en traits interrompus correspondent aux régressions linéaires de la courbe pour l'état R des protéines hydratées.

3.1.6 Superpositions des structures à 100K et 300K

De la même façon que pour les protéines dans le vide, nous avons superposé les crambines sur la Fig. 3.11 et les viscotoxines sur la Fig. 3.12, à 100K à gauche et à 300K à droite. Nous pouvons tirer les mêmes conclusions que pour les protéines déshydratées : les zones les plus sensibles à l'élévation de la température se situent dans la boucle reliant les deux hélices et dans les feuillets beta. Les mutations ne modifient pas la structure typique en L de la protéine et pour les crambines, le 22ième résidu joue toujours un rôle primordial dans la protéine puisque les boucles des crambines P22/L25 (en bleu) et P22/I25 (en vert) se superposent avec beaucoup de similitudes, de même pour les crambines S22/I25 (en rouge) et S22/L25 (en violet) à 100K. En comparant à la Fig. 2.19, nous constatons que VA3 et VA2 ont une boucle à 100K et 300K plus proche dans le vide. Dans l'eau, VB se distingue nettement.

3.2 Analyse des types de mouvements et inhomogénéité des fluctuations

3.2.1 Dispersion des fluctuations

De la même façon qu'au chapitre 2, partie II, nous avons étudié la dispersion de la mobilité des protéines, calculée avec l'équation 2.26. Nous avons alors tracé l'évolution de $\frac{\sigma^2}{T^2}$ en fonction de la température et nous l'avons comparée à celle du $\frac{MSD}{NT}$ pour la crambine P22/L25 sur la Fig. 3.10. Les calculs ont été réalisés sur la 1ère condition initiale (les courbes pour les autres thionines sont en annexe O).

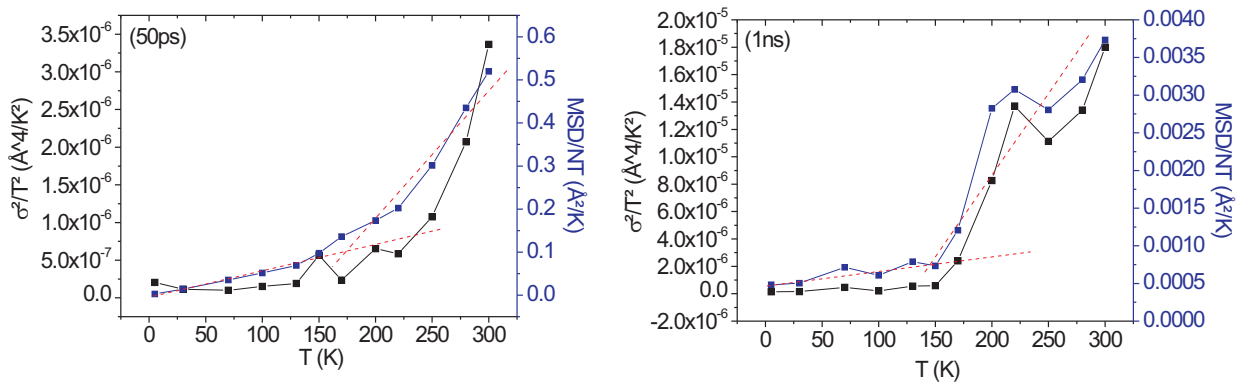


FIG. 3.10 – Variation de $\frac{\sigma^2}{T^2}$ (carrés noirs) et de $\frac{MSD}{NT}$ (carrés bleus) en fonction de la température pour la crambine P22/L25 hydratée, sur une trajectoire de 50ps à gauche et sur une trajectoire de 1ns à droite. Les lignes rouges en traits interrompus correspondent aux régressions linéaires des états R et F.

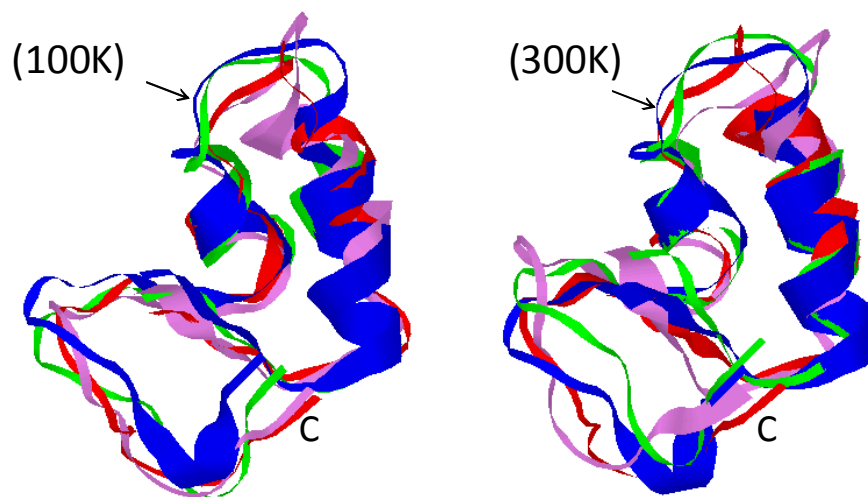


FIG. 3.11 – Superpositions des crambines équilibrées et hydratées, à 100K à gauche et à 300K à droite : P22/L25 en bleu, S22/I25 en rouge, P22/I25 en vert et S22/L25 en violet. C indique l'extrémité de la protéine correspondante au résidu 46. La flèche localise l'acide aminé n°22.

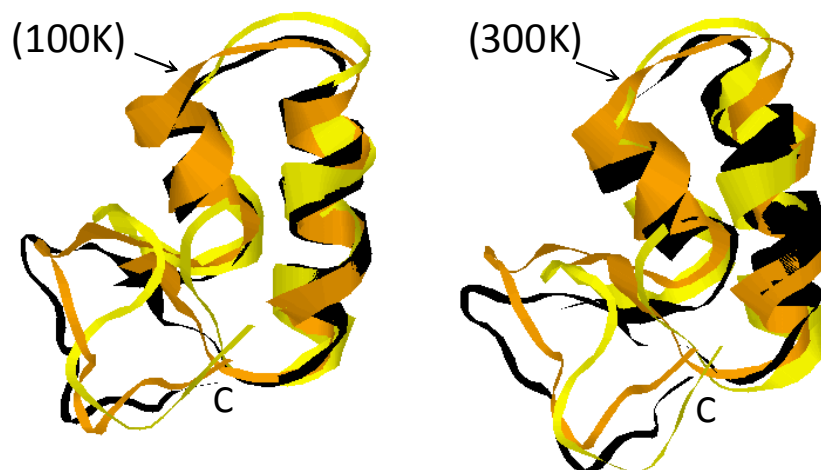


FIG. 3.12 – Superpositions des viscotoxines équilibrées et hydratées, à 100K à gauche et à 300K à droite : VA3 en orange, VA2 en noir et VB en jaune. C indique l'extrémité de la protéine correspondante au résidu 46. La flèche localise l'acide aminé n°22.

De la même façon que dans le vide, la dispersion suit la même évolution que l'écart quadratique moyen. L'élévation de la température implique une augmentation de zones mobiles pour les protéines hydratées. Nous pouvons remarquer que l'étude de la dispersion met bien en évidence la transition dynamique dans l'eau avec une rupture de pente très marquée, (pour P22/L25 à $\sim 150K$). Avec une trajectoire de 50ps, cette transition démarre plus tard, pour $T > 200K$.

3.2.2 Identifications des atomes mobiles

Nous avons pu voir précédemment que la flexibilité des protéines simulées dans le vide était due exclusivement aux groupements méthyles. En analysant la répartition de la mobilité le long de la séquence, nous pouvons voir quels types d'atomes contribuent à la flexibilité des protéines hydratées. Sur la Fig. 3.13, nous avons représenté le MSD de chaque atome de la crambine P22/L25 hydratée pour différentes températures, puis nous avons identifié à quels atomes correspondent les "pics" de mobilité obtenus.

Sur la Fig. 3.13, nous pouvons tout d'abord confirmer que, tout comme dans le vide, l'augmentation de la température implique une augmentation de zones mobiles. En identifiant les "pics" de mobilité obtenus, nous nous sommes rendus compte que les groupes méthyles n'étaient plus les seuls types d'atomes capables de subir des mouvements. Bien qu'ils contribuent toujours à la flexibilité de la protéine, d'autres groupements tels que NH_2 , NH_3 ou OH ont une grande mobilité pour les protéines hydratées. (Les mêmes conclusions sont tirées pour les autres thionines, présentées en annexe P). L'ajout d'un solvant a ainsi activé les mouvements de tous les atomes constituant la protéine ; tous les atomes sont susceptibles de bouger. Les protéines se différencient une fois encore par leurs atomes mobiles.

En effet, d'une protéine à une autre, les atomes mobiles ne sont pas les mêmes. Par exemple, le groupement NH_2 (atomes n°634, 635 et 636) du résidu ASN46 de la crambine S22/I25 a un MSD de 13\AA^2 alors que dans les deux mélanges de crambines (P22/I25 et S22/L25), le MSD de ce groupement est de 6\AA^2 et pour P22/L25 il vaut 3\AA^2 . De la même façon pour les viscotoxines, nous pouvons voir, par exemple dans le résidu LYS1, que le groupement NH_3 (atomes n°19, 20, 21 et 22) a un MSD de 2\AA^2 pour VB, 6\AA^2 pour VA3 et 18\AA^2 pour VA2. Les mutations impliquent donc des différences de zones mobiles tout le long de la séquence.

Avec la Fig. 3.14, nous montrons que les résultats dépendent du temps de simulation également pour les protéines simulées dans une boîte d'eau. Sur cette figure où nous représentons les résultats de la crambine P22/L25, les pics obtenus diffèrent d'une trajectoire à une autre (4 trajectoires ont été générées : 50ps, 100ps, 200ps et 1ns, les courbes pour les autres thionines sont en annexe Q). Par exemple, le CH_3 du résidu ALA38 (atomes n°534, 535, 536 et 537) a un MSD de 1\AA^2 à 50ps, 3\AA^2 à 100ps, 6\AA^2 à 200ps et 8\AA^2 à 1ns. Tous les mouvements ne sont pas activés aux mêmes instants.

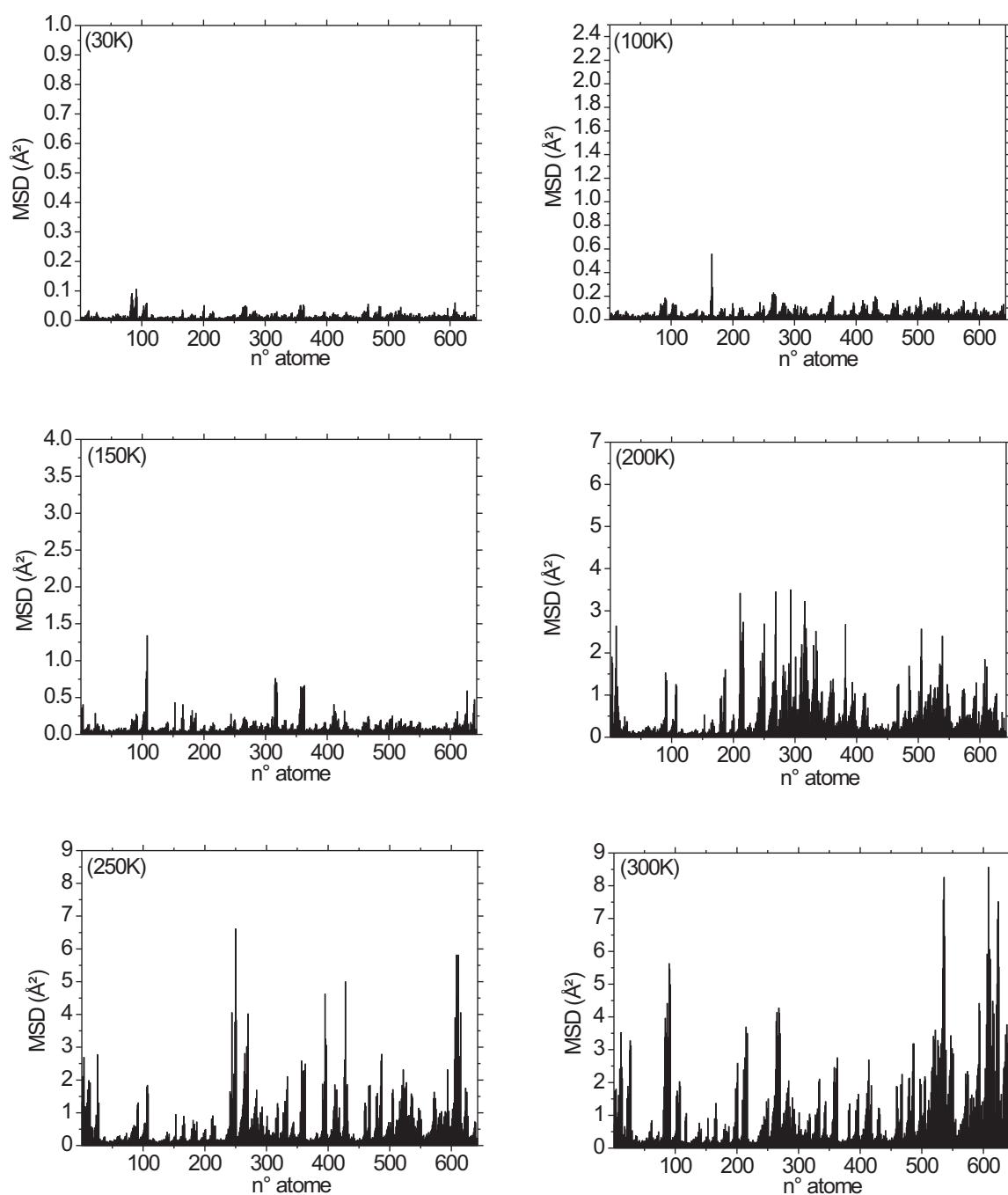


FIG. 3.13 – Variation de la mobilité des atomes avec la température de la crambine P22/L25 hydratée sur une trajectoire de 1ns.

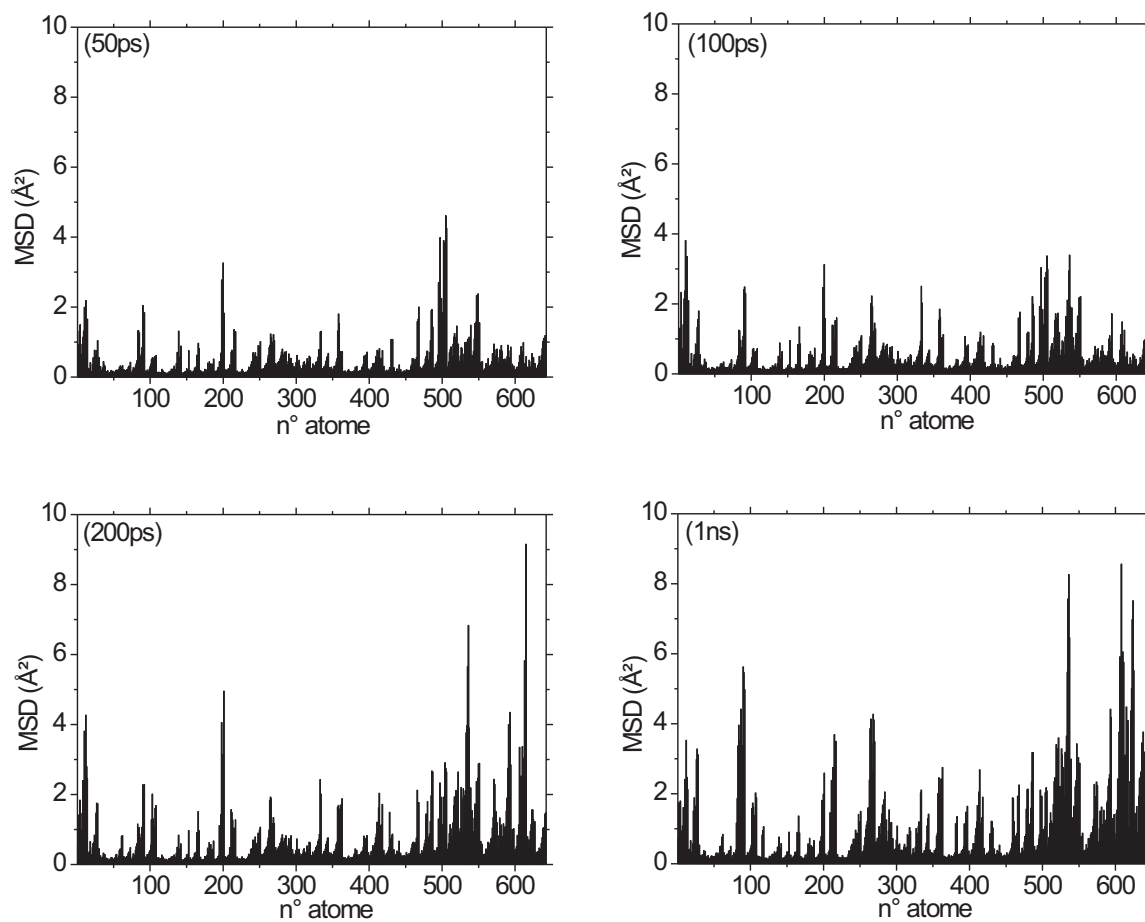


FIG. 3.14 – Variation de la mobilité des atomes avec le temps de simulation pour la crambine P22/L25 hydratée à 300K.

3.2.3 Mobilité des résidus

Dans l'étude de la mobilité des résidus des protéines hydratées, nous avons conservé le même critère de sélection qu'au chapitre 2, partie II, à savoir : sera considéré comme mobile, tout résidu dont son $\frac{MSD}{T}$ est supérieur à $\frac{MSD_{max}}{2T}$. Dans le tableau suivant, nous répertorions les résidus les plus mobiles pour les sept thionines hydratées.

Protéine	Résidus les plus mobiles						
P22/L25	THR1	ILE7	LEU25	TYR29	ALA38	TYR44	
$\frac{MSD}{T}(\text{Å}^2/K)$	0.199	0.253	0.146	0.222	0.147	0.227	
S22/I25	ARG17	THR21	TYR29	ASN46			
$\frac{MSD}{T}(\text{Å}^2/K)$	0.282	0.599	0.356	0.441			
P22/I25	THR1	LEU18	THR21	TYR29	PRO41	ALA45	
$\frac{MSD}{T}(\text{Å}^2/K)$	0.211	0.255	0.226	0.212	0.221	0.132	
S22/L25	ARG17	LEU25	TYR29	TYR44			
$\frac{MSD}{T}(\text{Å}^2/K)$	0.463	0.273	0.504	0.376			
VA3	LYS1	ARG17	LEU18	ARG23	LYS28	LYS33	LYS46
$\frac{MSD}{T}(\text{Å}^2/K)$	0.240	0.229	0.193	0.235	0.287	0.196	0.244
VA2	LYS1	PHE18	ARG23	PRO41	LYS46		
$\frac{MSD}{T}(\text{Å}^2/K)$	0.805	0.308	0.692	0.315	0.347		
VB	ARG23	ARG25	PRO45	LYS46			
$\frac{MSD}{T}(\text{Å}^2/K)$	0.362	0.305	0.359	0.704			

TAB. 3.7 – Résidus considérés comme mobiles pour les thionines hydratées sur une trajectoire de 1ns. Les résidus hydrophobes sont représentés en jaune. La valeur du MSD/T correspond au MSD du résidu à $T=300K$ divisé par $300K$.

La viscotoxine VB montre une flexibilité très localisée. L'analyse de ses résidus mobiles met en évidence la forte contribution du résidu LYS46, avec un $\frac{MSD}{T} \simeq 0.70\text{Å}^2/K$, alors que la mobilité des autres résidus chute à $0.40\text{Å}^2/K$. C'est pourquoi seulement 4 de ses 46 résidus seront considérés comme les plus mobiles pour cette protéine. De la même façon, l'arginine, résidu n°23, de la viscotoxine VA2 a une très forte mobilité à $300K$, (due à l'activation des mouvements de ses chaînes latérales, décrite dans le paragraphe suivant), avec un $\frac{MSD}{T} \simeq 0.70\text{Å}^2/K$. Puis, ses autres résidus ont un $\frac{MSD}{T} \simeq 0.30\text{Å}^2/K$. La viscotoxine VA3 a de nombreux sites mobiles, correspondant essentiellement à des lysines ou arginines. Seul un de ses résidus mobiles, la leucine en 18ième position, est hydrophobe. Inversement, pour la crambine P22/I25, la mobilité de cette protéine se localise tout au long de sa séquence. La crambine P22/L25 est la seule des thionines étudiées qui présentent autant de résidus hydrophobes dans la sélection des résidus mobiles.

En comparant nos résultats aux travaux d'Eyal, développés au chapitre 2, partie II [Eyal *et al.* 03], nous pouvons remarquer que nous obtenons les mêmes conclusions pour les résidus polaires (Arg et Lys), c'est à dire que ce sont des résidus mobiles. Cependant, contrairement à son analyse, les acides aminés aromatiques ont également un $\frac{MSD}{T}$ élevé dans l'étude de tous les résidus des thionines.

D'après les résultats de R. Lavery et S. Sacquin-Mora, les résidus jouant un rôle clé dans la fonction biologique sont rigides [Sacquin-Mora et Lavery 06]. Si on compare les valeurs du

$\frac{MSD}{T}$ du résidu 25 des viscotoxines, résidu impliqué dans la toxicité [Coulon *et al.* 03], nous obtenons les résultats du tableau 3.8.

	<i>Diminution de la toxicité</i> →		
Viscotoxine	VA3	VA2	VB
$MSD (\text{Å}^2)$	12.33	40.55	91.51
	<i>Augmentation de la mobilité</i> →		

TAB. 3.8 – Etude de la mobilité du résidu 25 des viscotoxines hydratées.

Cette propriété se vérifie pour le 25ième résidu des viscotoxines. La mobilité de cet acide aminé augmente lorsque la toxicité des protéines diminue.

Dans le tableau suivant, nous comparons les résidus les plus mobiles pour les protéines déshydratées et dans l'eau.

Protéine		Résidus les plus mobiles							
P22/L25	vide	THR1	ILE7	LEU25	ALA27	ILE34			
	eau	THR1	ILE7	LEU25	TYR29	ALA38	TYR44		
S22/I25	vide	ILE7	VAL8	VAL15	LEU18	THR30	ILE33	ILE35	THR39
	eau	ARG17	THR21	TYR29	ASN46				
P22/I25	vide	ILE25	TYR29	THR30	ILE33	ILE34			
	eau	THR1	LEU18	THR21	TYR29	PRO41	ALA45		
S22/L25	vide	ILE7	VAL15	LEU25	ILE33	PRO41			
	eau	ARG17	LEU25	TYR29	TYR44				
VA3	vide	THR7	ILE12	LEU18	THR19	ALA21	LEU29	ILE35	
	eau	LYS1	ARG17	LEU18	ARG23	LYS28	LYS33	LYS46	
VA2	vide	PHE18	VAL25	LEU29	ILE34	ALA37			
	eau	LYS1	PHE18	ARG23	PRO41	LYS46			
VB	vide	ILE12	THR15	ARG25	LEU29	ILE35	PRO41		
	eau	ARG23	ARG25	PRO45	LYS46				

TAB. 3.9 – Comparaisons des résidus les plus mobiles pour les thionines déshydratées et hydratées sur une trajectoire de 1ns.

D'après le tableau 3.9, nous pouvons remarquer que l'hydratation modifie la distribution de la mobilité le long de la séquence. Très peu de résidus mobiles dans le vide le sont également dans l'eau. Pour la crambine P22/L25 seulement, 3 résidus mobiles se retrouvent dans les deux phases. La contribution des résidus hydrophobes est moins évidente dans l'eau. La forte mobilité des acides aminés hydrophiles peut être expliquée par un couplage avec le solvant. En effet, les mouvements des molécules d'eau entraînent ces résidus hydrophiles via des liaisons hydrogène. Plusieurs types d'acides aminés sont considérés comme les plus mobiles d'une thionine à une autre : il s'agit de l'arginine (acide aminé le plus hydrophile), la tyrosine et la lysine. Ces résidus ont des mouvements de grande amplitude que nous développerons au paragraphe suivant.

Nous avons représenté la mobilité des résidus par des histogrammes, de la même façon que dans l'étude des protéines déshydratées ; sur la Fig. 3.15, nous représentons le nombre de résidus suivant leur valeur du $\frac{MSD}{T}$ (les mêmes courbes sont tracées pour les autres thionines en annexe R). Nous pouvons ainsi constater que la transition dynamique commence à environ 220K (histogramme rose). En effet, le nombre de résidus dont leur valeur du $\frac{MSD}{T}$ est inférieure à $0.10 \text{Å}^2/K$ chute à cette température, ceci implique une augmentation de leur flexibilité et l'activation des mouvements des atomes.

Contrairement aux résultats des protéines déshydratées, ce ne sont plus les mouvements des méthyles (activés dès 100/130K) qui sont responsables de l'augmentation de la flexibilité des protéines dans l'eau. Même s'ils contribuent toujours à cette mobilité, il existe des mouvements responsables de la transition dynamique pour d'autres atomes.

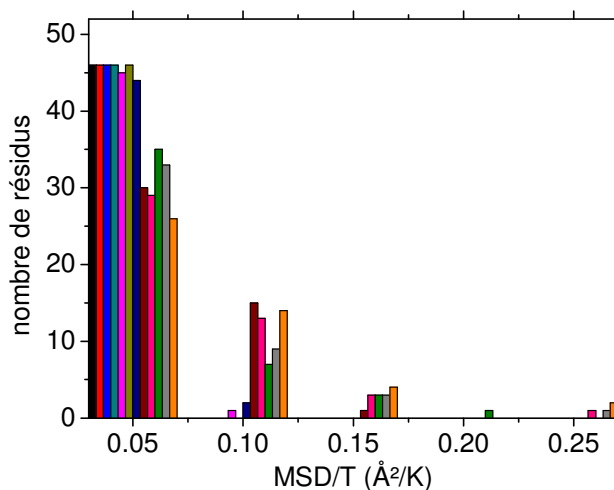


FIG. 3.15 – Variation du nombre de résidus suivant leur valeur du $\frac{MSD}{T}$ pour la crambine P22/L25 hydratée, sur une trajectoire de 1ns. Les histogrammes correspondent aux températures suivantes : noir = 5K, rouge = 30K, bleu = 70K, bleu-vert = 100K, fushia = 130K, kaki = 150K, bleu marine = 170K, marron = 200K, rose = 220K, vert = 250K, gris = 280K, orange = 300K.

Sur la Fig. 3.16, nous avons tracé l'évolution du MSD de chaque acide aminé en fonction de la température pour la crambine P22/L25 dans le vide et hydratée. Cette représentation de la mobilité met clairement en évidence les différences provoquées par l'eau. Alors que dans le vide, la flexibilité de la protéine se traduit essentiellement par l'activation des mouvements des méthyles dès 100/130K, dans l'eau, ce type de mouvement n'est plus le seul à contribuer à la flexibilité de la protéine. L'augmentation de la flexibilité de la protéine hydratée démarre à environ 200K avec les trois types de comportement décrits dans le chapitre 2, partie II.

Certains résidus ont des transitions dynamiques très prononcées (avec par exemple pour la crambine P22/L25 l'acide aminé ASN46 qui a un MSD de 5.05\AA^2 à 250K , 8.55\AA^2 à 280K et 30.84\AA^2 à 300K), d'autres ont une évolution du MSD avec des "sauts" (comme le résidu TYR29 avec un MSD de $47,59\text{\AA}^2$ à 220K , $15,85\text{\AA}^2$ à 250K et $62,24\text{\AA}^2$ à 280K) et enfin d'autres résidus ne subissent aucune élévation de la mobilité (comme le résidu ALA9 qui n'augmente pas plus de 3.50\AA^2 avec la température).

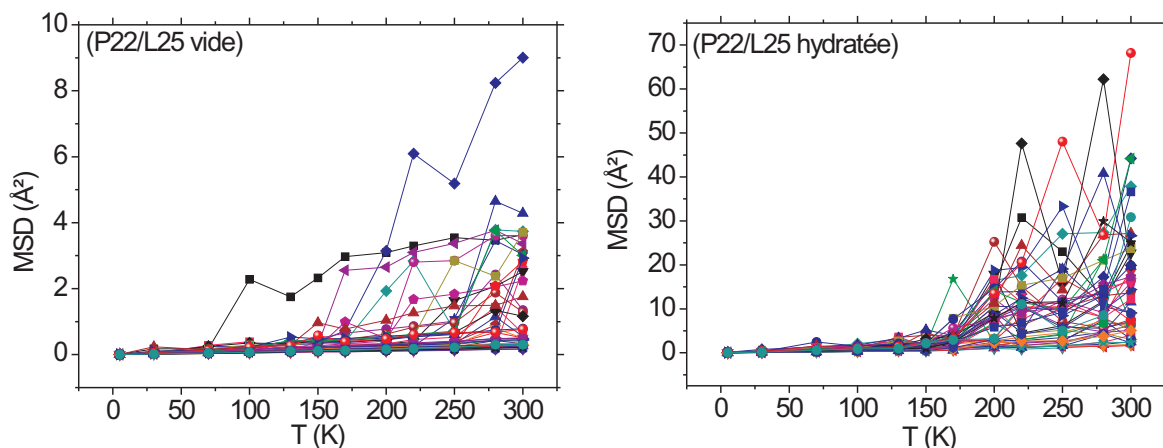


FIG. 3.16 – Variation du MSD de chaque acide aminé avec la température pour la crambine P22/L25 simulée dans le vide à gauche et hydratée à droite, sur une trajectoire de 1ns.

L'évolution du MSD pour chaque acide aminé en fonction de la température est représentée en annexe R pour les autres thionines. Les conclusions sont similaires à celles obtenues pour la crambine P22/L25.

3.2.4 Analyses des types de mouvements

De la même façon que pour les protéines déshydratées, nous avons visualisé les types de mouvements que subissent les protéines à l'aide du logiciel VMD. Alors que dans le vide, les méthyles sont les seuls types d'atomes susceptibles de subir des mouvements de rotations et de translations, dans l'eau, tous les atomes sont susceptibles de bouger.

De manière générale, les acides aminés avec de grandes chaînes latérales subissent de nombreux mouvements. En effet, l'agitation créée par les molécules d'eau entraîne plus facilement des chaînes latérales avec un fort encombrement dans l'espace. La tyrosine, l'arginine ou encore la lysine possèdent des chaînes latérales avec soit des cycles aromatiques, soit un enchaînement de groupements CH-NH-NH_2 qui basculent d'un côté et de l'autre. Ces mouvements de grandes amplitudes contribuent fortement aux fluctuations structurales des protéines dans le régime anharmonique (caractérisé par la constante K_F).

Par exemple, sur la Fig. 3.17, nous pouvons voir la différence de l'encombrement dans l'espace des chaînes latérales entre deux résidus, l'un considéré comme mobile : la tyrosine (pour la crambine S22/I25, TYR29 a un $\frac{\text{MSD}}{T} \simeq 0.189\text{\AA}^2/\text{K}$), l'autre rigide : l'alanine (pour

la crambine S22/I25, ALA9 a un $\frac{MSD}{T} \simeq 0.022 \text{Å}^2/K$). Nous pouvons constater que plus la chaîne latérale sera grande, plus le résidu sera mobile de part les mouvements que subiront les chaînes latérales entraînées par le solvant. Des groupements de NH_3 et CH_3 subissent, en plus des basculements, des rotations autour de l'atome d'azote ou de carbone.

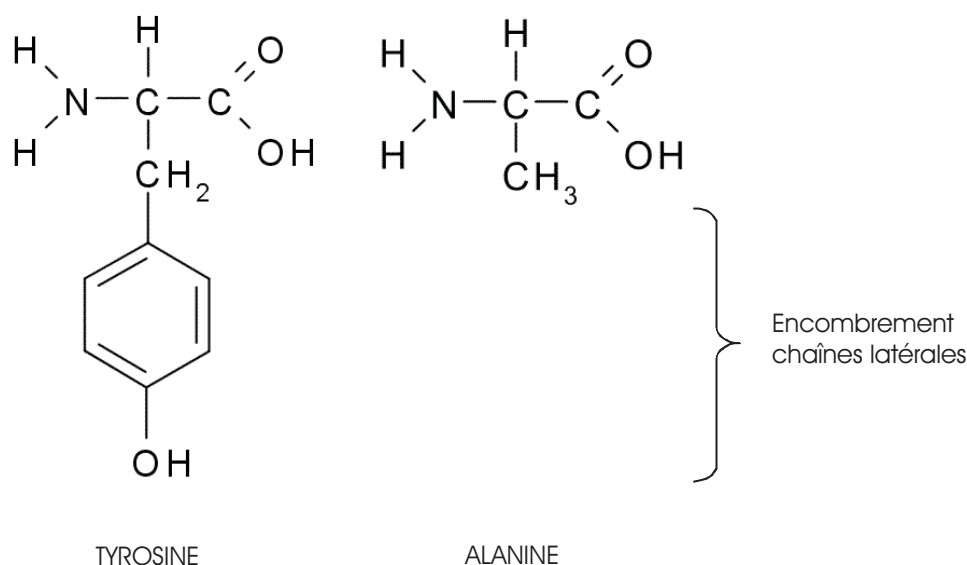


FIG. 3.17 – Comparaisons entre deux acides aminés, la tyrosine à gauche, l'alanine à droite, de l'encombrement spatial des chaînes latérales.

Nous pouvons également remarquer que pour toutes les protéines, même si les résidus mobiles diffèrent d'une protéine à une autre, ils restent localisés dans les mêmes zones de la protéine. En effet, la boucle reliant les deux hélices (résidus 18 à 23) et la zone directement opposée dans l'espace (résidus 33 à 46) contribuent majoritairement à la flexibilité de la protéine.

Les hélices, fixées par des ponts disulfures, sont très rigides, et donc très stables à l'élévation de la température. Les acides aminés situés en surface des thionines sont également plus mobiles que ceux situés au coeur de la protéine, ils sont en contact direct avec l'eau et l'agitation du solvant implique une activation des mouvements des atomes.

3.3 Rôle de l'eau : analyse des paramètres macroscopiques

3.3.1 Rayons de Giration

A l'aide du même programme fortran utilisé pour les protéines dans le vide, nous avons calculé à partir du MSD de chaque atome, l'évolution du rayon de giration des protéines hydratées.

3.3.1.1 Crambines

Dans l'étude des crambines, nous avons comparé nos calculs à des résultats expérimentaux. En effet, M. Teeter a publié sur la Protein Data Bank des structures de crambines obtenues par diffraction X à différentes températures [Teeter *et al.* 93, Teeter *et al.* 01]. A partir des valeurs des B-facteurs, nous avons recalculé le rayon de giration avec l'Eq. 2.24 de la partie I. D'après les cristaux utilisés, ses résultats seront comparables à la crambine P22/L25.

Sur la Fig. 3.18, nous avons donc comparé les évolutions des rayons de giration de la crambine P22/L25 hydratée et dans le vide aux calculs réalisés d'après les résultats expérimentaux de M. Teeter. Les deux premiers points, représentés par des carrés rouges, ont été réalisés sur des cristaux différents. Dans nos calculs, nous remarquons que l'allure de l'évolution du MSD est semblable à la courbe déduite des structures expérimentales, avec la présence d'un pic du rayon de giration à $200K$, température de transition. Ce pic étant également présent dans le calcul de R_{gmean} , nous en concluons que la protéine subit un changement structural à la température de transition dynamique. Cette nette augmentation de la flexibilité à $200K$ est due à l'activation des mouvements de la boucle, en particulier des résidus 19 (Proline) et 20 (Glycine), cf Fig. 3.13.

La coïncidence de ce pic à la température de transition est clairement visible sur la Fig. 3.19, où nous comparons l'évolution du rayon de giration avec le $\frac{MSD}{N}$. Nous remarquons que pour la crambine S22/I25, nous n'observons pas de pic. Dans les deux protéines, nous constatons que le rayon de giration dans le vide est supérieur au rayon de giration dans l'eau à très basses températures ($< 100K$ pour P22/L25 et $< 200K$ pour S22/I25) et inférieur au-delà de la température de transition dynamique.

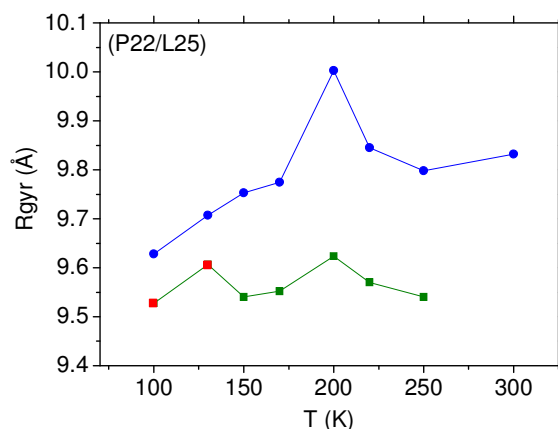


FIG. 3.18 – Variation du rayon de giration de la crambine P22/L25 dans l'eau (ronds bleus), et à partir des structures expérimentales de M. Teeter (carrés verts et rouges). Les résultats expérimentaux représentés par les carrés verts ont été obtenus sur le même échantillon.

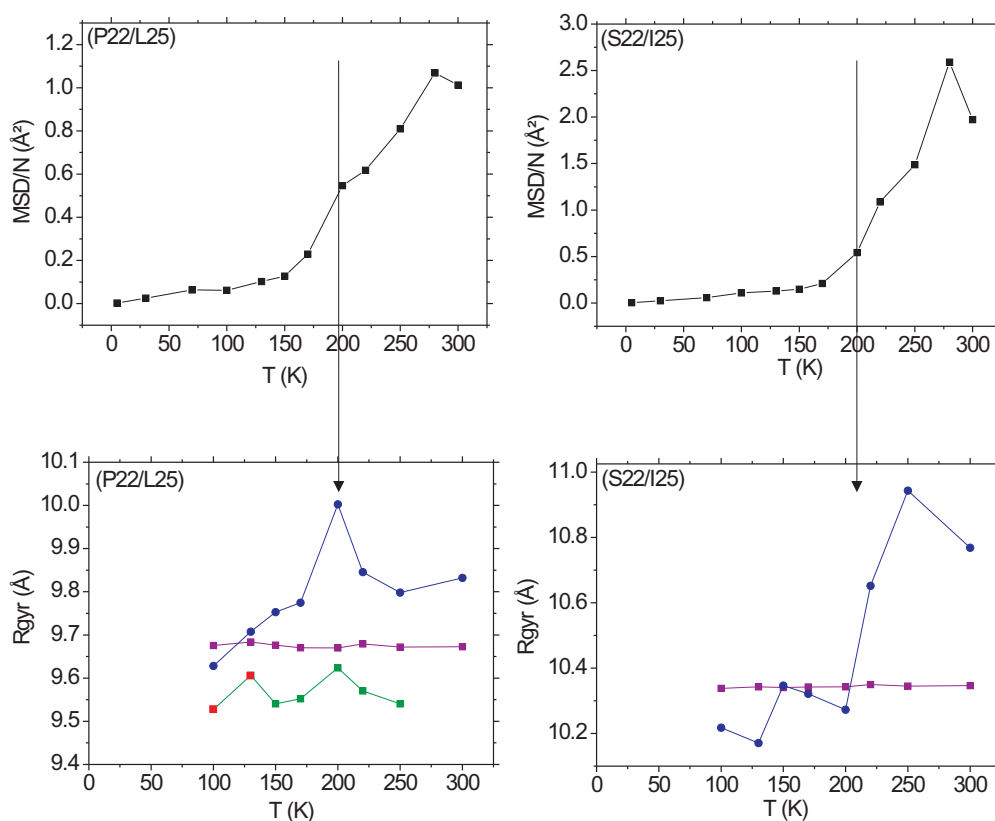


FIG. 3.19 – Variation des rayons de giration pour P22/L25 à gauche, S22/I25 à droite : dans l'eau (ronds bleus), dans le vide (carrés violets) et à partir des structures DRX de M. Teeter (carrés verts et rouges), comparés à la variation de l'écart quadratique moyen (carrés noirs).

Sur la Fig. 3.20, nous avons représenté la variation des rayons de girations pour les deux mélanges de crambines : P22/I25 et S22/L25, dans l'eau et dans le vide. Nous pouvons constater que seule P22/I25 a une évolution différente avec une brusque diminution à 250K et son rayon de giration dans le vide n'est pas supérieur à celui dans l'eau pour $T < 130K$. Cependant, à la température de transition dynamique, les deux crambines ont également une augmentation du rayon de giration. D'autres recherches ont également permis d'observer que le rayon de giration dans le vide est plus petit que dans l'eau [Zanotti *et al.* 99, Steinbach et Brooks 93, Ahlström *et al.* 87].

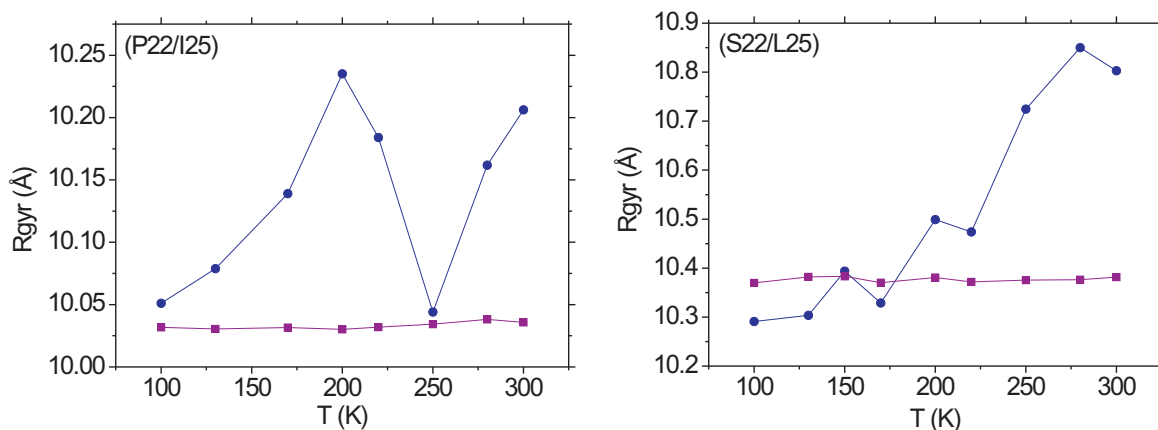


FIG. 3.20 – Variation du rayon de giration de P22/I25 à gauche et S22/L25 à droite dans le vide (carrés violets) et dans l'eau (ronds bleus).

Sur la Fig. 3.21, nous comparons les différentes évolutions des rayons de giration des crambines hydratées aux résultats de M. Teeter.

Nous pouvons tout d'abord remarquer que les deux crambines présentant une proline en 22ième position (courbes rouge et marron), et celles avec une serine en 22ième position (courbes bleue et rose) ont des valeurs de leur rayon de giration semblables. De plus, pour P22/I25 et P22/L25, nous obtenons bien un pic à 200K, comme dans les résultats expérimentaux. Sachant que les résultats DRX ont été obtenus sur des cristaux de P22/I25, nous pouvons constater que la mutation sur le 25ième résidu entre une leucine (P22/L25) et une isoleucine (P22/I25) n'affecte pas la variation de ce rayon de giration.

Les différences engendrées par la mutation sur le résidu 22 entre une serine et une proline provient du fait que ces deux acides aminés diffèrent par leur surface d'hydratation, mais aussi car la proline est composée d'un cycle spécial, dont sa structure se referme sur la chaîne latérale ce qui la rend très peu flexible.

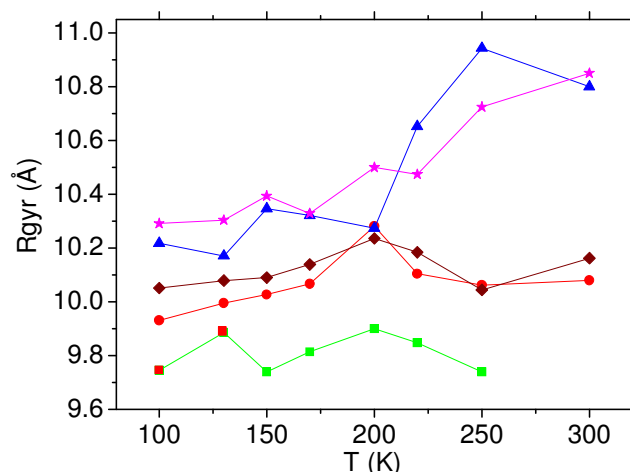


FIG. 3.21 – Comparaison de la variation du rayon de giration à partir des résultats expérimentaux de M. Teeter (carrés verts et rouges) et des protéines hydratées : P22/L25 (ronds rouges), S22/I25 (triangles bleus), P22/I25 (losanges marrons), S22/L25 (étoiles roses).

Dans le calcul suivant, nous avons analysé les variations du rayon de giration pour deux trajectoires de longueur différente : une sur 100ps et l'autre sur 1ns. Sur la Fig. 3.22, nous pouvons remarquer que l'évolution du rayon de giration est semblable entre une trajectoire de 100ps (courbe noire) et celle d'1ns (courbe rouge). Le pic obtenu à 200K existe quelque soit le temps de production.

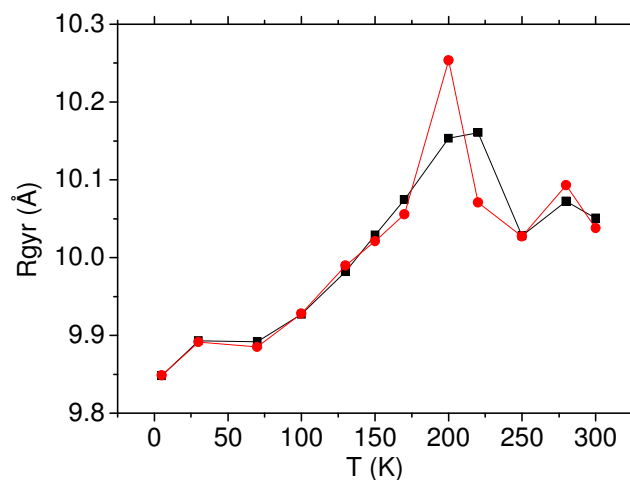


FIG. 3.22 – Comparaison de l'évolution du rayon de giration sur une trajectoire de 100ps (courbe noire) et une trajectoire de 1ns (courbe rouge) pour crambine P22/L25 hydratée.

3.3.1.2 Viscotoxines

Dans le cas des viscotoxines, le rayon de giration montre une brusque élévation à la température de transition dynamique, cf Fig. 3.23. Comme pour les crambines, le rayon de giration du vide (courbe violette) est supérieur à celui de l'eau (courbe bleue) pour $T < 130\text{K}$. Une fois les mouvements des chaînes latérales activés, le rayon de giration de la protéine hydratée devient supérieur. Ceci confirme des travaux réalisés sur la ribonucléase par des simulations de dynamique moléculaire par l'équipe de Curtis en 2004 [Curtis *et al.* 04].

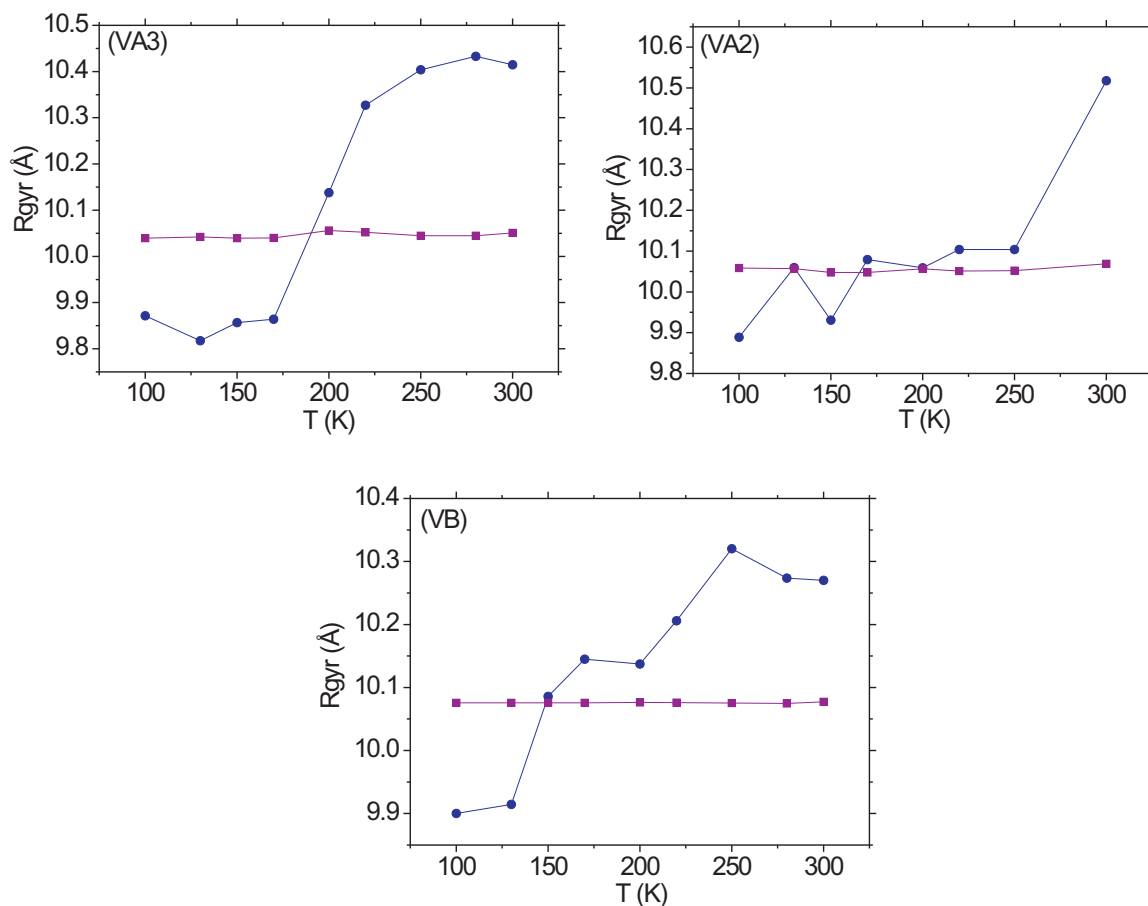


FIG. 3.23 – Variation du rayon de giration des viscotoxines dans l'eau (ronds bleus) et dans le vide (carrés violets).

Sur la Fig. 3.24, nous comparons les rayons de giration des viscotoxines hydratées. Nous pouvons remarquer que l'ordre de grandeur est le même quelque soit la viscotoxine, soit $\sim 10.1\text{\AA}$ à 200K et l'augmentation due à la transition dynamique est surtout marquée pour VA3 (courbe noire).

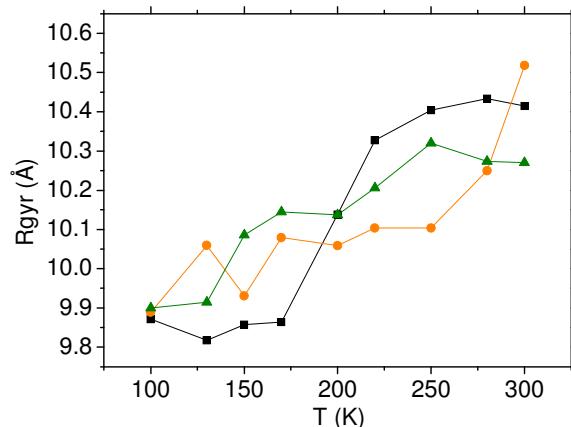


FIG. 3.24 – Variation du rayon de giration des viscotoxines hydratées : VA3 (carrés noirs), VA2 (ronds oranges) et VB (triangles verts).

3.3.2 Variation du volume

Nous avons ensuite estimé le volume de la protéine avec la formule suivante, en considérant qu'elle pouvait être assimilée à une sphère :

$$V_p = \frac{4}{3}\pi R_{gyr}^3, \quad (3.1)$$

où V_p représente le volume de la protéine et R_{gyr} son rayon de giration.

Afin d'évaluer la taille de la protéine par rapport à la taille de la boîte $V_{bo\hat{t}e}$, nous avons calculé le rapport entre le volume de la protéine (cf Eq. 3.1) et le volume de la boîte d'eau (qui varie avec la température dans l'ensemble NPT), calculé sur la moyenne des 5 trajectoires de 1ns.

Sur la Fig. 3.25, nous représentons ce calcul pour les 4 crambines. Nous pouvons constater que l'évolution du rayon de giration et du rapport $\frac{V_p}{V_{bo\hat{t}e}}$ sont les mêmes. Notons que $V_p \simeq 0.10V_{bo\hat{t}e}$, le volume de la protéine est petit par rapport à celui de la boîte d'eau, permettant ainsi une hydratation complète de toute la surface de la protéine.

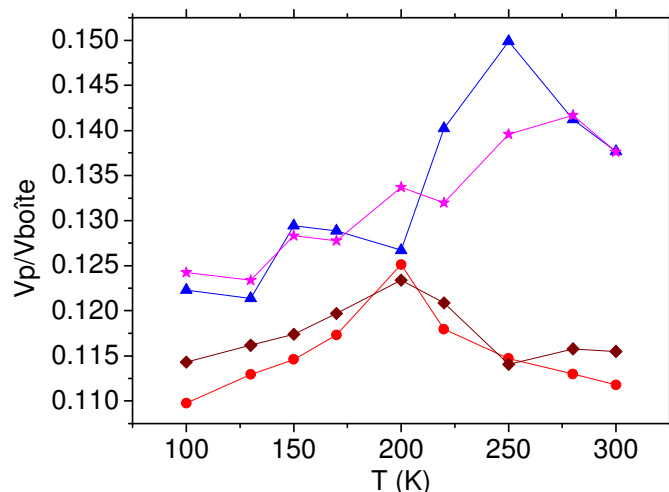


FIG. 3.25 – Variation du rapport entre le volume de la protéine (V_p) et le volume de la boîte ($V_{bo\grave{it}e}$) pour les crambines hydratées : P22/L25 (ronds rouges), S22/I25 (triangles bleus), P22/I25 (losanges marrons), S22/L25 (étoiles roses).

Puis nous avons estimé ce même rapport pour les viscotoxines. Sur la Fig. 3.26, nous pouvons remarquer que la variation de ce rapport est la même que celui du rayon de giration également pour les viscotoxines. L'ordre de grandeur est le même que pour les crambines, soit $V_p \sim 0.10V_{bo\grave{it}e}$.

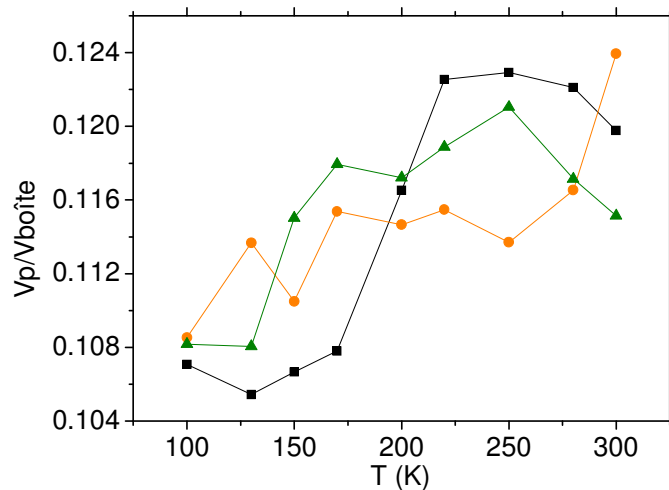


FIG. 3.26 – Variation du rapport entre le volume des viscotoxines hydratées et celui de la boîte d'eau : VA3 (carrés noirs), VA2 (ronds oranges) et VB (triangles verts).

3.3.3 Variation du volume de la boîte de simulation

Nous avons représenté sur la Fig. 3.27 les variations des volumes des boîtes simulées pour les thionines. Nous avons calculé la moyenne du volume sur les 5 trajectoires de 1ns générées. Nous pouvons constater que ce volume diminue jusqu'à 150–200K puis augmente brusquement à la température de transition. La diminution de volume observée en-dessous de 150–200K est en sens opposé à celui du volume de la protéine à ces températures, puisque le volume de la protéine augmente ou reste approximativement constant en-dessous de 150K comme on peut l'observer sur la Fig. 3.24 (le volume V_p étant défini par l'équation 3.1). L'augmentation du rapport $\frac{V_p}{V_{bo\grave{it}e}}$ en-dessous de 150K observé sur les Figs. 3.25 et 3.26 est dû essentiellement à une diminution du volume de la boîte de simulation (Fig. 3.27). Cette contraction de la cellule de la boîte de simulation est donc due à une augmentation de la densité de l'eau simulée à ces températures. Par contre, l'expansion de la boîte qui commence aux températures supérieures à la température de la transition dynamique s'explique par une augmentation du volume de la protéine et ne peut s'expliquer par les propriétés intrinsèques de l'eau de volume.

Pour nous en convaincre, nous avons simulé une boîte d'eau sans protéine à différentes températures et nous avons calculé la densité de l'eau simulée. Les résultats sont représentés sur la Fig. 3.28 et sont comparés aux résultats expérimentaux que nous avons trouvés dans la littérature [Watanabe 91]. Dans l'eau simulée, le volume de la boîte d'eau diminue (la densité augmente) avec la température et atteint un maximum autour de 273K pour rediminuer ensuite. La courbe simulée ne reproduit pas quantitativement les propriétés de l'eau macroscopique (courbe en bleu) mais bien son comportement qualitatif. (Le nombre de molécules qui est simulé reste très petit dans la boîte : 216 molécules d'eau pour un volume de 6700Å³).

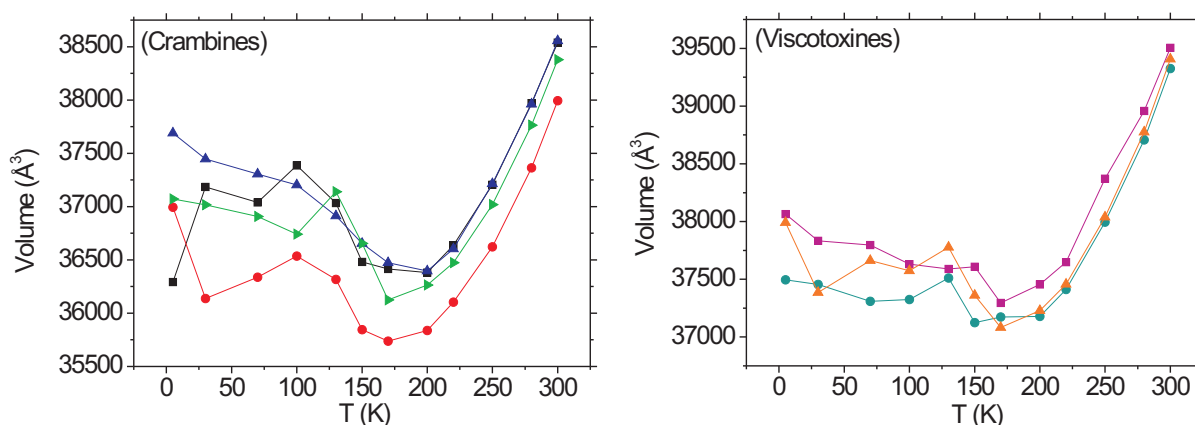


FIG. 3.27 – Variation du volume des boîtes de simulation avec l'augmentation de la température pour les crambines à gauche : P22/L25 (carrés noirs), S22/I25 (ronds rouges), P22/I25 (triangles bleus), S22/L25 (triangles verts) et pour les viscotoxines à droite : VA3 (carrés roses), VA2 (ronds verts) et VB (triangles oranges).

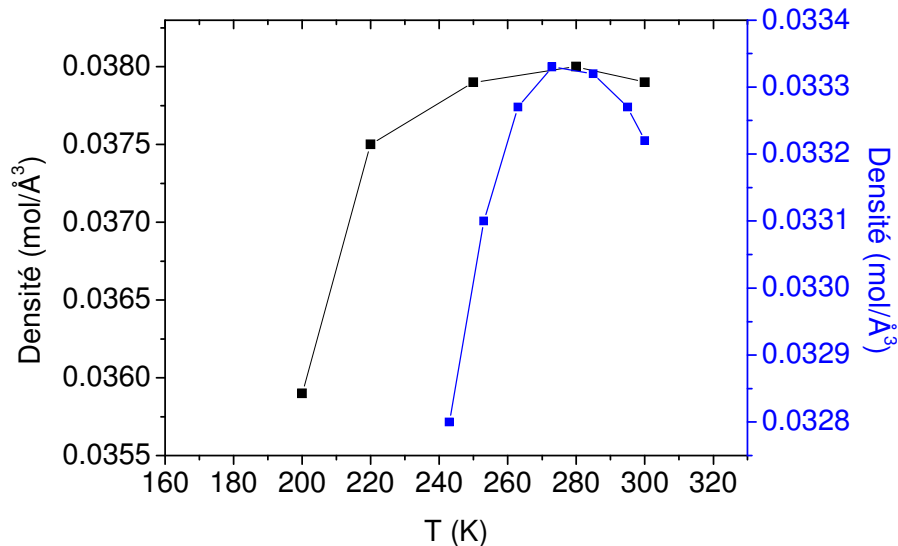


FIG. 3.28 – Variation de la densité de l'eau avec la température, calculée à partir de nos simulations de dynamique moléculaire (courbe noire) et comparée à celle calculée d'après la littérature (courbe bleue) [Watanabe 91].

3.3.4 Etude de la répartition de l'eau autour des protéines

Plusieurs auteurs mentionnent une transition de phase de l'eau à la transition dynamique [Kumar *et al.* 06, Zanotti *et al.* 05, Zanotti *et al.* 07]. Afin de caractériser le rôle de l'eau, nous avons étudié comment les propriétés structurales de l'eau d'hydratation évoluaient avec la température. Nous avons choisi d'étudier la répartition des molécules autour de la protéine de manière statistique (les molécules d'eau étant indiscernables) en utilisant la procédure suivante.

Nous avons divisé l'eau d'hydratation entourant la protéine arbitrairement en 10 couches successives. Comme illustré sur la Fig. 3.29, à partir de la distance maximale, d_{\max} , entre une molécule d'eau et la protéine, la couche n a été définie au temps $t = 0$ (conditions initiales de l'étape de production des résultats) par tous les atomes d'oxygène à une distance comprise entre $(n - 1)\frac{d_{\max}}{10}$ et $n\frac{d_{\max}}{10}$.

Nous avons alors analysé la répartition des molécules d'eau autour des protéines au cours du temps. Pour toutes les molécules d'une couche n définie au temps $t = 0$, nous avons évalué la quantité suivante aux instants t ultérieurs :

$$P_{\text{couche } n}(t, x) = \sum_{at \in \text{couche } n} \frac{1}{t} \sum_{t'=0}^t p(at, x, t'), \quad (3.2)$$

où $x = m\frac{d_{\max}}{10}$ ($m = 0, \dots, 10$) et $p(at, x, t') = 1$ si à t' , $at \in \left] x, x + \frac{d_{\max}}{10} \right]$ (0 sinon), at représente l'atome d'oxygène (de la molécule d'eau).

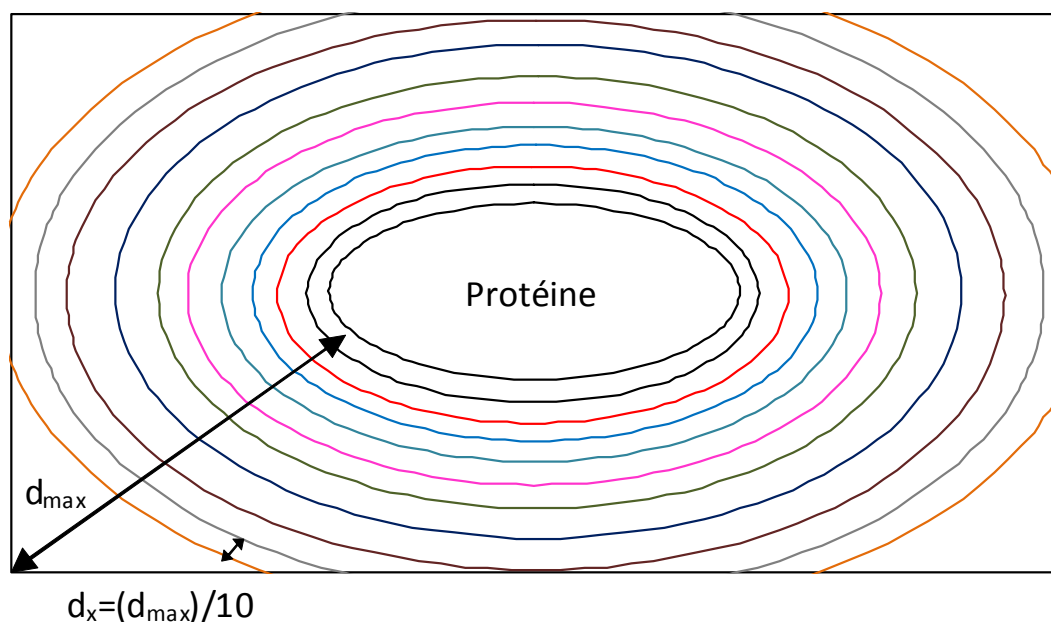


FIG. 3.29 – Schéma des couches d'hydratation, les couleurs des n couches correspondent à celles données dans le tableau 3.11.

Sur la Fig. 3.30, nous avons représenté la variation de cette probabilité pour la crambine P22/L25 à gauche et la viscotoxine VA3 à droite en fonction de la température : 150, 220 et 300K (les mêmes résultats ont été trouvés pour les autres thionines). Nous pouvons constater qu'il existe pour toutes les protéines un "pic" dans la distribution de la probabilité à environ 4Å. La dernière couche n'existe pas à basses températures (celle à environ 17/18Å) ; en effet, le système évoluant dans l'ensemble NPT, à basse température cette couche n'existe pas puisque la boîte est plus petite.

La diminution de la probabilité observée à grande distance est un artéfact dû à la boîte d'eau, lorsque l'on se rapproche des extrémités de la boîte, il y a de moins en moins de molécules d'eau.

couleur	noire	rouge	bleue	bleu-vert	rose	kaki	bleu marine	marron	gris	orange
couche (Å), P22/L25	1.871	3.743	5.614	7.486	9.357	11.229	13.100	14.972	16.843	18.714
couche (Å), VA3	1.757	3.514	5.270	7.027	8.784	10.541	12.297	14.054	15.811	17.568

TAB. 3.10 – Représentation des couches d'hydratation.

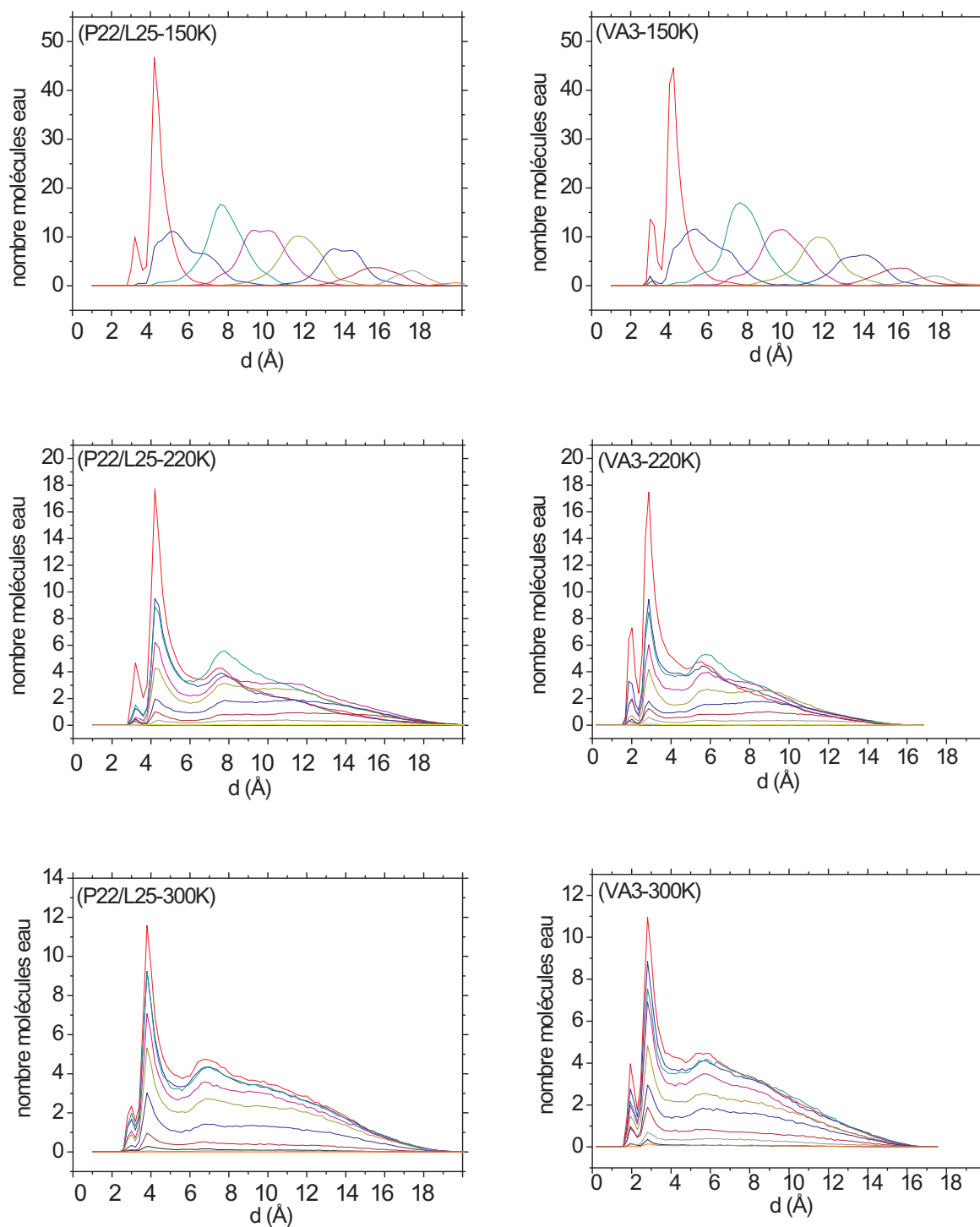


FIG. 3.30 – Etude de la répartition de l'eau avec la distribution de probabilité de présence des molécules d'eau pour P22/L25 à gauche et VA3 à droite. Le calcul étant réalisé dans l'ensemble NPT, les couches ne seront pas les mêmes suivant le système étudié (cf Tab. 3.11).

Sur la Fig. 3.31, nous avons représenté cette même variation mais en sommant toutes les couches afin d'évaluer la distribution globale des molécules d'eau (les courbes des autres thionines sont en annexe S).

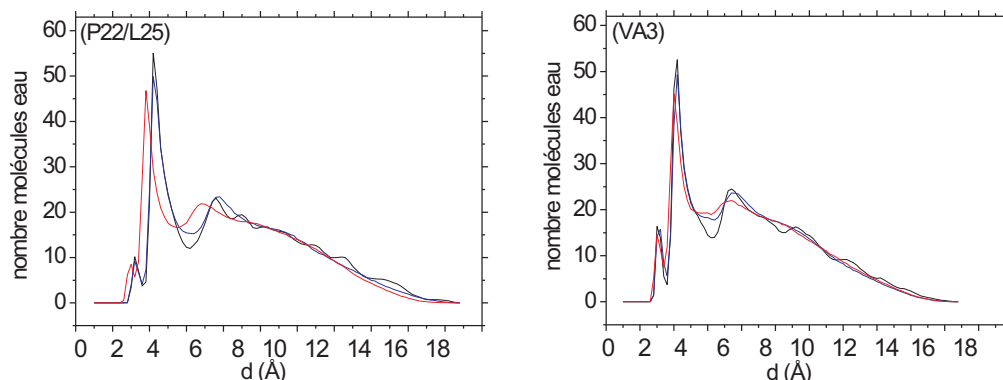


FIG. 3.31 – Somme des distributions de probabilité de présence des molécules d'eau pour P22/L25 à gauche et VA3 à droite, à 150K (courbe noire), 220K (courbe bleue) et 300K (courbe rouge).

Les Figs. 3.30 et 3.31 mettent en évidence l'activation de la diffusion des molécules d'eau à 200K. Les molécules proches de la protéine diffusent ensuite dans tout l'espace (il y aura alors des échanges permanents entre les molécules éloignées et les molécules proches). La structure de l'eau, bien visible à 150K par de nombreux maxima dans la fonction étudiée, disparaît au-delà de 200K. Ceci est dû aux phénomènes d'amorphisation et de diffusion. La température de transition coïncide bien avec la transition de phase de l'eau. Ceci confirme les résultats obtenus par Zanotti en 2007 [Zanotti *et al.* 07].

3.3.5 Variation du taux d'hydratation

Afin d'étudier les contributions de l'hydratation sur la dynamique des protéines, nous avons simulé deux autres boîtes d'eau (cf chapitre 4, partie I) : une première d'une taille supérieure aux calculs précédents ($46.55 \times 40.82 \times 37.10 \text{ \AA}^3$, boîte n°2, $h = 7.5$), et une plus petite représentant les conditions expérimentales en DRX ($40.824 \times 18.498 \times 22.371 \text{ \AA}^3$, boîte n°3, $h = 0.9$) où se trouvent deux protéines par cristal dans une symétrie P21 (nous avons réalisé cette étude sur la crambine P22/L25). Cette symétrie implique une deuxième protéine placée aux coordonnées $(-X, \frac{1}{2} + Y, -Z)$ par rapport à la première protéine située aux coordonnées (X, Y, Z) . Les mêmes protocoles d'hydratation ont été réalisés dans le cas de la boîte n°2. Pour les simulations du cristal, il n'a pas été possible de respecter la symétrie dans l'ensemble NPT avec le champ de force CHARMM. Nous avons donc dû prendre un ensemble NVT. De plus, pour ces simulations, nous avons utilisé la condition initiale n°1.

3.3.5.1 Courbes générales de la mobilité de la crambine P22/L25

Sur la Fig. 3.32, nous avons comparé les écarts quadratiques moyens de la crambine P22/L25 hydratée, dans les calculs précédents (courbe rouge) pour un taux d'hydratation $h = 4$, pour $h = 7.5$ (boîte d'eau supérieure, courbe verte) et pour $h = 0.9$ (équivalente au cristal, courbe bleue).

Nous pouvons constater que le comportement global ne subit que très peu de différences entre les deux premières boîtes. Le taux d'hydratation précédent étant déjà élevé, ajouter des molécules d'eau n'apporte que très peu de différences.

Le pic obtenu à $250K$ pour $h = 7.5$ est probablement dû aux conditions initiales utilisées. A l'opposé, la variation de l'écart quadratique moyen du cristal diffère considérablement des deux autres boîtes. L'ordre de grandeur est diminué de moitié dans la partie anharmonique.

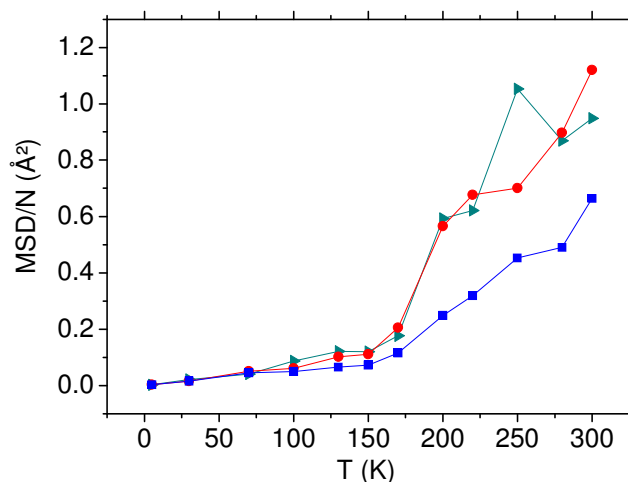


FIG. 3.32 – Comparaison de l'écart quadratique moyen de la crambine P22/L25 hydratée dans la boîte n°1, $h = 4$ (courbe rouge), dans la nouvelle boîte d'eau supérieure, $h = 7.5$ (courbe verte) et dans le cristal, $h = 0.9$ (courbe bleue) sur une trajectoire de 1ns.

A partir de la Fig. 3.33, nous pouvons conclure que les deux protéines présentes dans le cristal ont le même comportement global. Pour la suite de notre étude, nous considérerons alors la première protéine, située aux coordonnées (X, Y, Z) .

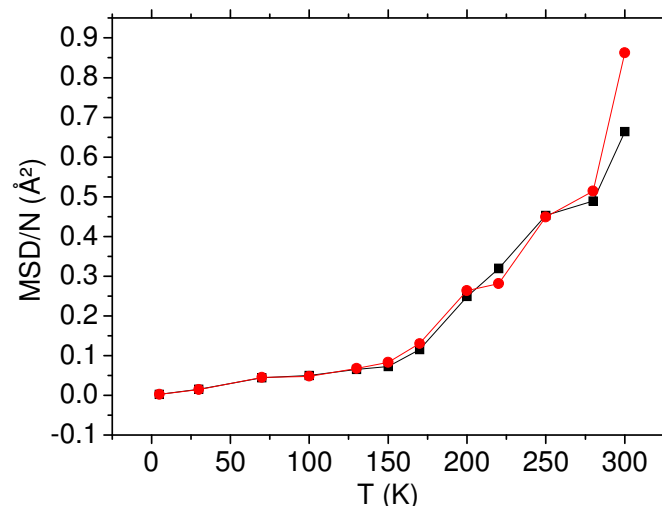


FIG. 3.33 – Comparaison de l'écart quadratique moyen des deux protéines contenues dans le cristal de la boîte n°3, sur une trajectoire de 1ns.

Nous avons calculé les constantes effectives, K_R et K_F , pour la protéine dans les différentes boîtes, avec la méthode des interpolations linéaires décrites dans le chapitre 2, partie II. Les résultats sont donnés dans le tableau suivant :

	$K_R(N/m)$	$K_F(N/m)$	$\Delta S_{FR}(kcal/mol)$	Densité ($mol/\text{Å}^3$)
boîte n°1, $h = 4$ (en solution)	5.7	0.7	6.3	0.032
boîte n°2, $h = 7.5$ (en solution)	4.9	0.7	5.6	0.032
boîte n°3, $h = 0.9$ (cristal)	8.2	1.1	6.1	0.026

TAB. 3.11 – Paramètres effectifs de la crambine P22/L25 hydratée avec trois boîtes de simulation différentes.

Les paramètres effectifs de la crambine P22/L25 dans les deux premières boîtes sont très proches. En effet, les variations de leur écart quadratique moyen sont semblables. L'augmentation d'entropie créée par la transition dynamique reste du même ordre de grandeur quelque soit le taux d'hydratation. La protéine est plus rigide dans le cristal, pour une densité la plus petite.

Sur la Fig. 3.34, nous avons étudié la dépendance en temps de simulation de l'écart quadratique moyen. Comme dans la boîte n°1, la flexibilité globale de la protéine augmente avec le temps de production.

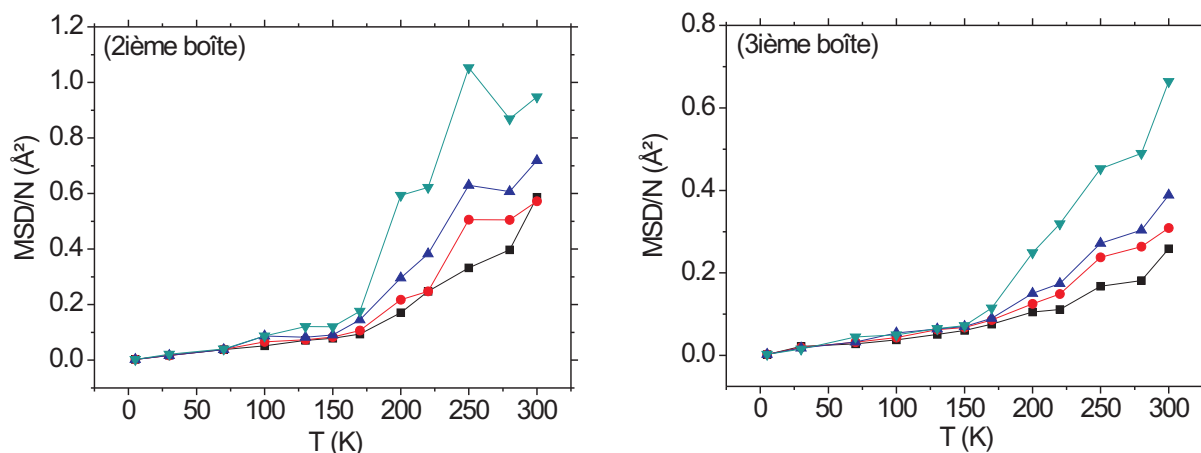


FIG. 3.34 – Variation de l'écart quadratique moyen de la crambine P22/L25 hydratée dans la boîte d'eau de taille supérieure à gauche ($h = 7.5$) et dans le cristal à droite ($h = 0.9$), en fonction du temps de simulation : pour une trajectoire de 50ps (courbe noire), de 100ps (courbe rouge), de 200ps (courbe bleue) et de 1ns (courbe verte).

3.3.5.2 Identifications des atomes mobiles

Sur la Fig. 3.35, nous avons représenté le MSD de chaque atome à 300K de la crambine P22/L25 en solution pour les calculs précédents à gauche ($h = 4$) et pour $h = 7.5$ à droite. L'étude détaillée de cette mobilité nous indique qu'il n'y a que très peu de différences entre ces deux simulations.

Les "pics" de mobilité interviennent pour les mêmes atomes (malgré les différences, les résultats concordent). Par exemple, le CH_3 du résidu LEU25 (atomes n°360, 361, 362 et 363) a un MSD d'environ 3Å^2 dans la boîte n°1 et un MSD de 5Å^2 dans la boîte n°2. La majeure partie des différences interviennent en fin de séquence sur les résidus 44 et 45, où cette partie était très mobile dans la boîte n°1, avec par exemple, le CH du résidu TYR44 (atomes n°607 et 608) qui avait un MSD de 8Å^2 dans l'étude précédente et de 1.5Å^2 dans cette deuxième boîte.

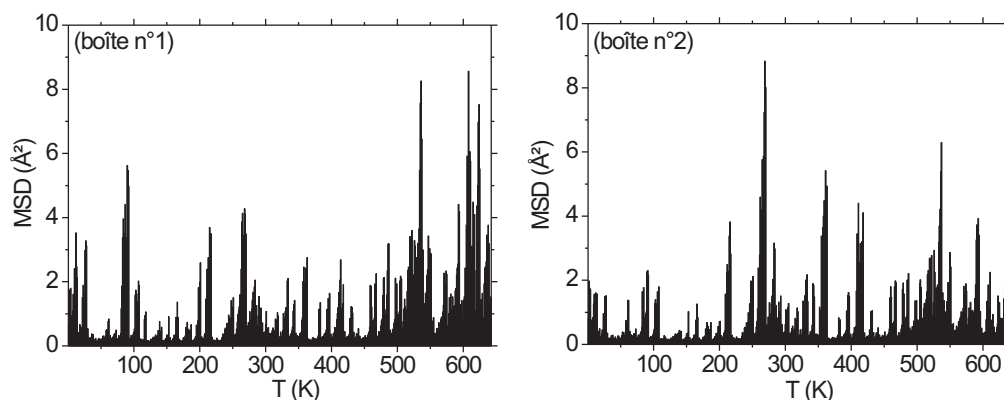


FIG. 3.35 – Mobilité des atomes pour la crambine P22/L25 simulée dans la boîte d'eau n°1 à gauche ($h = 4$) et dans la boîte d'eau n°2 à droite ($h = 7.5$), à 300K et sur une trajectoire de 1ns.

Les différences engendrées par une diminution du taux d'hydratation ($h = 0.9$) sont beaucoup plus importantes. Sur la Fig. 3.36, nous pouvons constater que la distribution de la flexibilité est totalement modifiée. Les atomes mobiles ne sont plus les mêmes.

D'après ce résultat, nous pouvons en conclure que le solvant joue un rôle crucial dans la transition dynamique puisque le taux d'hydratation modifie complètement la mobilité des atomes et de la protéine. Ce taux d'hydratation étant insuffisant pour recouvrir toute la protéine dans l'étude du cristal, les molécules d'eau ne sont pas en contact avec tous les atomes de la protéine et elles ne peuvent donc pas activer tous les mouvements qui existent avec un taux d'hydratation plus grand. La forte mobilité de la fin de séquence pour $h = 4$ a complètement disparu dans le cristal ($h = 0.9$). Par exemple, le CH_3 du résidu ASN46 a un MSD d'environ 7Å^2 dans la boîte n°1 alors qu'il chute à 0.3Å^2 dans la boîte n°3.

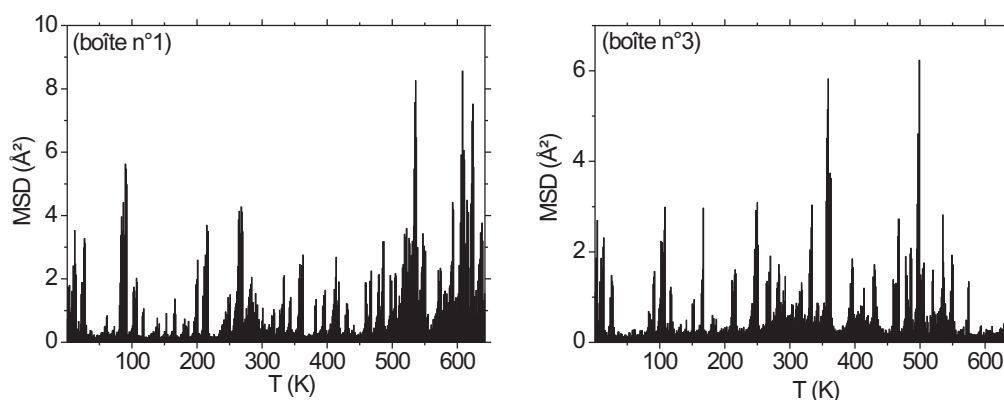


FIG. 3.36 – Mobilité des atomes pour la crambine P22/L25 simulée dans la boîte d'eau n°1 à gauche ($h = 4$) et dans la boîte d'eau n°3 à droite ($h = 0.9$), à 300K et sur une trajectoire de 1ns.

3.3.5.3 Résidus mobiles

Dans le tableau suivant, nous répertorions les résidus les plus mobiles (le critère de sélection reste le même : le $\frac{MSD}{T}$ doit être supérieur au $\frac{MSD_{max}}{2T}$).

Boîte 1, $h = 4$	THR1	ILE7	(LEU18)	LEU25	TYR29	ALA38	TYR44	(ASN46)
$\frac{MSD}{T} (\text{Å}^2/K)$	0.199	0.253	(0.126)	0.146	0.222	0.147	0.227	(0.103)
Boîte 2, $h = 7.5$	(THR1)	ILE7	LEU18	LEU25	TYR29	ALA38	(TYR44)	ASN46
$\frac{MSD}{T} (\text{Å}^2/K)$	(0.109)	0.161	0.210	0.161	0.244	0.203	(0.144)	0.326
Boîte 3, $h = 0.9$	ARG17	LEU25	ILE35	THR39	ASN46			
$\frac{MSD}{T} (\text{Å}^2/K)$	0.077	0.129	0.106	0.091	0.181			

TAB. 3.12 – Comparaison des résidus les plus mobiles de la crambine P22/L25 simulée avec trois boîtes d'hydratation de dimensions différentes. La valeur du MSD/T correspond au MSD du résidu à $T=300K$ divisé par $300K$. Les résidus entre parenthèses sont légèrement en-dessous du critère de sélection.

Nous pouvons constater qu'augmenter le volume de la boîte d'eau ne modifie pas la distribution de la mobilité de la crambine P22/L25. Les différences mineures observées dans le tableau 3.12 entre les résultats pour un taux d'hydratation élevé ne sont pas significatifs.

Dans le cas du cristal, les résidus mobiles diffèrent puisque nous avons pu voir précédemment que toute la distribution de la flexibilité diffère plus avec un taux d'hydratation plus faible.

3.3.5.4 Variation du volume des boîtes simulées

Sur la Fig. 3.37, nous avons comparé les évolutions des volumes des deux premières boîtes avec l'augmentation de la température. Nous pouvons voir que ces deux boîtes ont la même évolution quelque soit leur taille. (Ce calcul n'a pas été réalisé dans le cristal puisque les simulations ont été faites dans l'ensemble NVT).

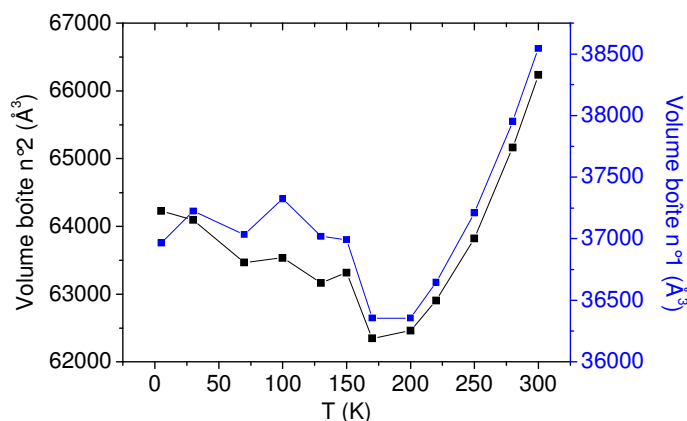


FIG. 3.37 – Comparaison des variations des volumes des boîtes d'eau avec l'augmentation de la température pour la première boîte, $h = 4$ (courbe bleue) et la seconde, $h = 7.5$ (courbe noire).

3.3.5.5 Répartition de l'eau

Nous retrouvons exactement la même évolution de la distribution de probabilité de présence dans les boîtes d'eau, comme le montre la figure 3.38. Le même "pic" à environ 3\AA existe.

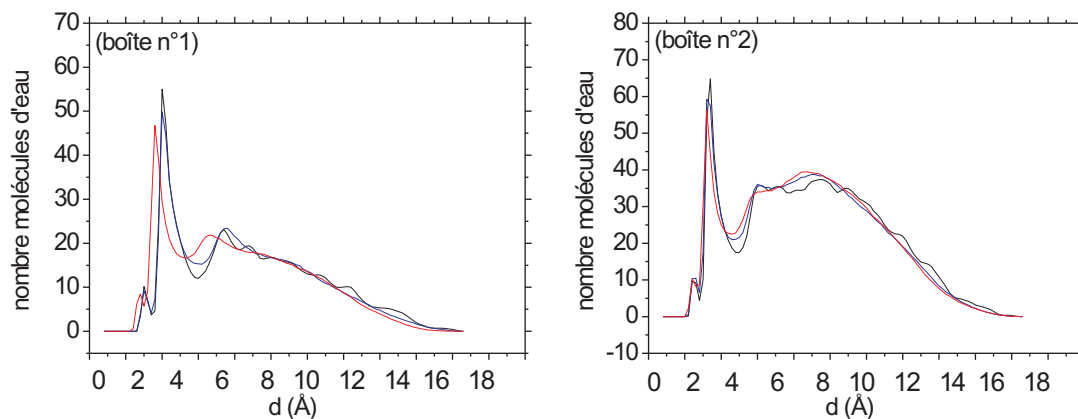


FIG. 3.38 – Somme des distributions de probabilité de présence des molécules d'eau pour P22/L25 dans la boîte n°1 ($h = 4$) à gauche et dans la boîte n°2 ($h = 7.5$) à droite, à 150K (courbe noire), 220K (courbe bleue) et 300K (courbe rouge).

De cette étude, nous pouvons conclure que le taux d'hydratation joue un rôle important dans la dynamique de la protéine. A partir d'un taux "critique" où toute la protéine se retrouve complètement hydratée, ce taux n'apporte plus de modifications si on l'augmente. En-dessous de ce taux critique, si la protéine présente des parties de sa séquence mal hydratées, la distribution de sa flexibilité sera complètement différente puisque le solvant ne pourra plus activer les mouvements des atomes.

3.3.6 Discussion

Les propriétés de l'eau sont modifiées à la surface des protéines. Cette eau est appelée eau interfaciale par opposition à l'eau de volume. La structure et la dynamique de l'eau interfaciale sont perturbées par les interactions hydrophiles et hydrophobes avec les systèmes biologiques [Bellissent-Funel 04]. Nous avons pu voir également que le degré d'hydratation joue un rôle important dans la dynamique des systèmes.

Nos résultats mettent en évidence la contraction de l'eau d'hydratation conforme à l'eau de volume. Le volume de la maille augmente en même temps que le rayon de giration des protéines à la transition. L'eau d'hydratation subit une expansion à des températures où l'eau de volume subirait une contraction (cf Figs. 3.24, 3.25, 3.26 et 3.27). La distribution de l'eau change également à la transition (cf Figs. 3.30 et 3.31), l'eau commence à diffuser dans tout le volume et seules les molécules proches de la protéine ont une "structure" qui varie peu avec la température. Ces molécules sont continuellement échangées avec celles du volume.

Dans le vide, les protéines sont plus compactes, d'où un rayon de giration plus petit, ceci confirme les résultats de Ahlström [Ahlström *et al.* 87]. L'agitation des molécules d'eau entraîne plus facilement les chaînes latérales. Celles-ci vont se déplacer vers l'extérieur de la protéine et devenir de plus en plus mobiles avec le solvant [Zanotti *et al.* 99].

Le taux d'hydratation joue un rôle primordial. Nos résultats sur l'étude du cristal, où $h = 0.9$, montrent une flexibilité des protéines totalement différente. Il existe donc un taux critique, à partir duquel la protéine est totalement hydratée, nécessaire pour activer tous les mouvements de la protéine.

3.4 Origine de l'effet des mutations

3.4.1 Modification du réseau de liaisons hydrogène

En respectant les mêmes critères que pour les protéines déshydratées (ie, il existe une liaison hydrogène si $d(O-H) < 2.4\text{\AA}$), chapitre 2, partie II, nous avons étudié les réseaux des liaisons hydrogène. Dans le tableau 3.13 sont répertoriées les liaisons hydrogène à 300K des crambines et dans le tableau 3.14 celles des viscotoxines. Comme précédemment, les chiffres correspondent aux numéros des acides aminés.

Les liaisons hydrogène qui existaient déjà dans le cas des protéines dans le vide sont représentées en jaune. Nous pouvons tout d'abord constater que l'hydratation a considérablement modifié le réseau de liaisons hydrogène. Dans l'étude des crambines, P22/L25 est celle qui présente le plus de modifications : alors que le réseau était constitué de 16 liaisons hydrogène dans le vide, nous en dénombrons 21 dans l'eau, et seulement 6 sont communes aux deux phases. Pour les trois autres crambines, le nombre de liaisons hydrogène est proche entre le vide et l'eau mais seules deux liaisons pour S22/L25 et S22/I25 et une seule pour P22/I25 se retrouvent dans les deux phases.

Dans le cas des viscotoxines, le réseau de VA3 passe de 17 liaisons dans le vide à 24 dans l'eau, 8 se retrouvant dans les deux phases. Pour VA2, l'hydratation provoque une diminution du réseau, avec 21 liaisons dans le vide et 16 dans l'eau. De plus, 5 liaisons existent dans l'eau et le vide. Enfin, pour VB, le nombre varie peu, avec 17 liaisons dans le vide et 15 dans l'eau, mais seules deux liaisons se retrouvent dans les deux phases.

De manière générale, l'hydratation a modifié tout le réseau de liaisons hydrogène. Les liaisons conservées dans les deux phases sont souvent les liaisons hydrogène de la première hélice alpha (résidus 7 à 18) : en effet, ces liaisons font partie des structures secondaires et sont nécessaires à la création des hélices. Les modifications ne concernent pas des résidus mobiles. Nous pouvons nous attendre à ce qu'un résidu change de configuration dans l'espace avec les mouvements qu'il subit et ainsi "casser" des liaisons hydrogène qui existaient dans le vide, mais ce n'est pas le cas pour les thionines hydratées.

P22/L25						
1-35	3 – 33	3 – 33	6-6	6-10	7 – 11	8 – 12
9 – 13	10-14	11-15	12-16	13 – 17	14-18	17-20
22-26	23-27	24-28	25-29	26-30	37-21	27-32
S22/I25						
1-35	3-33	3-33	6-10	7-11	8-12	9-13
10-14	14 – 18	26-30	27-31	27 – 32		
P22/I25						
1-35	3-33	3-33	6-10	7-11	8-12	9-13
10-14	11-15	12-16	13-17	23-27	24-28	26-30
27-31	27 – 32	37-37				
S22/L25						
1-35	3-33	3-33	6 – 10	7-11	8-12	9-13
10-14	13 – 17	14-18	23-25	26-30	27-31	27-32

TAB. 3.13 – Liaisons hydrogène des crambines hydratées à 300K. Les liaisons hydrogène qui existaient dans le vide sont représentées en jaune.

VA3							
1-35	3 – 33	3 – 33	4-46	6-6	6-10	7 – 11	8-12
9-13	10-14	11-15	12-16	13 – 17	14 – 18	15-19	16-20
16 – 21	22-26	23-27	24-28	25-29	26 – 30	27-31	44 – 41
VA2							
1-35	3 – 33	3-33	6 – 10	8 – 12	9 – 13	10-14	11 – 15
12-16	13-17	14 – 18	21-21	25-29	26-30	27-31	27-32
VB							
1-36	3-35	3-33	9-13	10-14	12 – 16	13-17	20-20
22 – 26	23-27	24-28	25-29	26-30	27-31	27-32	

TAB. 3.14 – Liaisons hydrogène des viscotoxines hydratées à 300K. Les liaisons hydrogène qui existaient dans le vide sont représentées en jaune.

3.5 Fluctuations structurales et paysage d'énergie libre en solution

3.5.1 Définition du paysage d'énergie libre

L'énergie libre dépend de la température, de l'énergie moyenne (dominée par l'énergie potentielle) et de l'entropie (dominée par les fluctuations). Dans l'ensemble canonique, l'énergie libre F est égale à :

$$F = U - TS, \quad (3.3)$$

où T est la température, S l'entropie et U l'énergie interne. Une grande variation de F avec la température est en général due à une variation importante d'entropie.

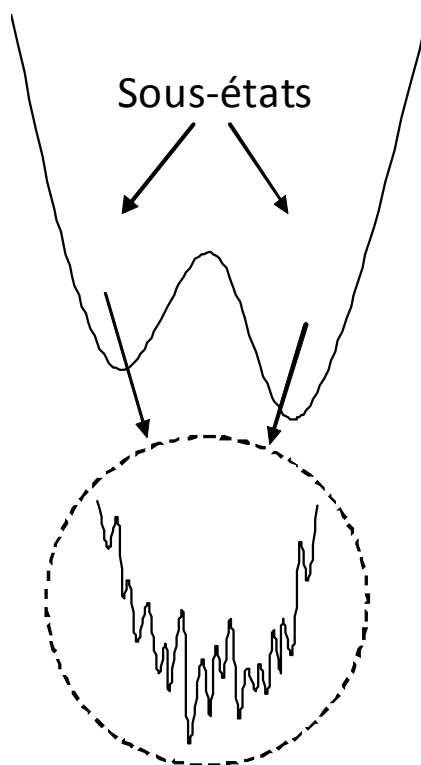
L'énergie libre d'une protéine est fonction d'un grand nombre de variables et possède un nombre important de minima locaux. Une protéine en solution dans son état natif, c'est-à-dire repliée, existe dans un ensemble de conformations stables qui sont proches en énergie. Chaque conformation occupe un minimum de la surface d'énergie libre de la protéine. L'énergie libre ressemble à un paysage de vallées séparées par des montagnes et cols qui est appelée littéralement « paysage d'énergie libre ». La notion de paysage d'énergie libre permet d'interpréter de nombreux phénomènes tels que le repliement des protéines, la fonction biologique des protéines, les mécanismes allostériques et la transition dynamique comme nous le démontrerons plus loin.

Le paysage d'énergie libre est construit dans $3N$ dimensions, où N est le nombre d'atomes de la molécule et ne peut donc pas être facilement visualisé. Une représentation schématique à une dimension est généralement utilisée pour illustrer la structure caractéristique et hiérarchique du paysage d'énergie libre des protéines.

La Figure 3.39 illustre la hiérarchie des états conformationnels. Dans l'exemple de la Fig. 3.39, une protéine repliée existe dans un puits d'énergie libre (état natif). Dans ce puits, la protéine peut adopter deux conformations différentes séparées par une barrière d'énergie importante (états conformationnels). Dans chacun de ces deux puits, la protéine existe dans de nombreuses conformations différentes (sous-états conformationnels). Le nombre de niveaux dans la hiérarchie peut varier suivant les protéines.

Le schéma de la Fig. 3.39 est en réalité un profil d'énergie libre le long d'une certaine coordonnée effective. Dans l'interprétation du repliement protéique, le paysage d'énergie libre est aussi représenté à trois dimensions par un « entonnoir ». Ces représentations simplifiées seront rendues concrètes dans nos simulations présentées plus bas.

Des expériences, comme la diffraction X, les neutrons ou la spectroscopie de Mössbauer mais aussi des simulations de dynamique moléculaire démontrèrent l'existence de ces sous-états conformationnels [Frauenfelder *et al.* 79, Frauenfelder *et al.* 91]. Dans la myoglobine, deux régions distinctes existent : la première est représentée par un noyau compact et la seconde est une région semi-liquide située en surface qui permet le passage des ligands.


 FIG. 3.39 – Représentation schématique du paysage d'énergie libre [Ishikawa *et al.* 08].

La hiérarchie des états conformationnels d'une protéine se reflète dans sa dynamique, soit vibrationnelle (mouvements autour d'un minima), soit diffusive (transitions entre minima). Les échelles de temps des mouvements dépendent à la fois des barrières d'énergie et de la masse des fragments en mouvements. Le tableau suivant en donne un aperçu.

Echelle de temps	Amplitude	Mouvements
courte femtoseconde, picoseconde $10^{-15} - 10^{-12}_s$	$0.001 - 0.1\text{Å}$	Liaisons, Angles, Angles Dihédraux
picoseconde, nanoseconde $10^{-12} - 10^{-9}_s$	$0.01 - 10\text{Å}$	Chaînes latérales, Mouvements collectifs
longue nanoseconde, microseconde $10^{-9} - 10^{-6}_s$	$1 - 100\text{Å}$	Repliement de petites protéines Transition des hélices alpha ("Helix-coil transition")
longue microseconde, seconde $10^{-6} - 10^{-1}_s$	$10 - 100\text{Å}$	Repliement protéines

TAB. 3.15 – Echelle de temps des mouvements de la protéine.

3.5.2 Paysage d'énergie libre et dynamique du backbone

3.5.2.1 Coordonnées internes γ

Dans le chapitre 2 de la partie II, nous avons montré que la transition dynamique des viscotoxines et des crambines pouvait être reproduite par le modèle simplifié de Doster. Ce modèle représente le paysage d'énergie libre de la protéine par un double puits le long d'une coordonnée spatiale effective (cf Fig. 2.1 du chapitre 2). En-dessous de la température de transition dynamique, la protéine est bloquée dans un puits et donc dans une conformation particulière. Au-dessus de cette température, la structure se réorganise et la protéine change de conformation en franchissant la barrière de potentiel entre deux puits.

Le modèle de Doster a le mérite de capturer la physique du phénomène mais il ne permet pas d'identifier quelle est l'origine des mouvements et barrières impliquées dans la transition. Dans ce qui suit, nous allons démontrer l'existence de plusieurs minima locaux et identifier les barrières d'énergie du paysage d'énergie libre du backbone des crambines et des viscotoxines.

Nous avons démontré que la transition dynamique des atomes du backbone est une propriété de la transition dynamique des protéines en solution : elle n'existe pas pour des protéines dans le vide. Les positions des atomes du backbone permettent de définir les éléments de structure secondaires (hélices α , feuillets β , ...) et la structure tertiaire de la protéine. Les mouvements du backbone sont assez lents et jouent un rôle dans le repliement de la protéine mais aussi dans la fonction biologique.

La conformation du backbone est définie dans notre travail par les angles diédraux γ_n définis à partir de 4 atomes C^α successifs ($n - 1, n, n + 1, n + 2$) le long de la structure primaire (cf partie I, chapitre 2). Le second atome $C^\alpha(n)$ est associé au résidu n de la séquence. Les angles diédraux γ sont des variables à gros grains utilisés pour décrire les fluctuations de peptides [Nishikawa *et al.* 74] et font parties d'un modèle à gros grains du repliement des protéines [Oldziej *et al.* 04]. Ces variables sont utiles pour décrire les relations compliquées entre la dynamique et la fonction biologique d'une protéine. Par exemple, les fluctuations des angles γ sont corrélées aux changements conformationnels des enzymes [Henzler-Wildman *et al.* 07]. L'avantage de l'utilisation des coordonnées internes γ au lieu des coordonnées cartésiennes des atomes de C^α est que leurs valeurs sont indépendantes des mouvements globaux de translation et de rotation de la protéine. Il n'est donc pas nécessaire d'appliquer la procédure de construction variationnelle d'une trajectoire de référence décrite dans la partie I.

Il faut noter que les fluctuations de la position de l'atome $C^\alpha(n)$ sont fortement corrélées aux fluctuations de l'angle diédral γ_n correspondant. En effet, sur la Fig. 3.40, nous avons représenté la variation de l'écart quadratique moyen des carbones alpha (courbe noire) et des angles γ (courbe bleue) pour une trajectoire de 1ns à 300K pour la crambine P22/L25 hydratée. Nous constatons que les fluctuations des $C^\alpha(n)$ et γ_n sont fortement corrélées. Le backbone de la protéine possède deux zones mobiles : la boucle reliant les deux hélices et les résidus des feuillets beta. Les mêmes conclusions peuvent être tirées des simulations des autres thionines.

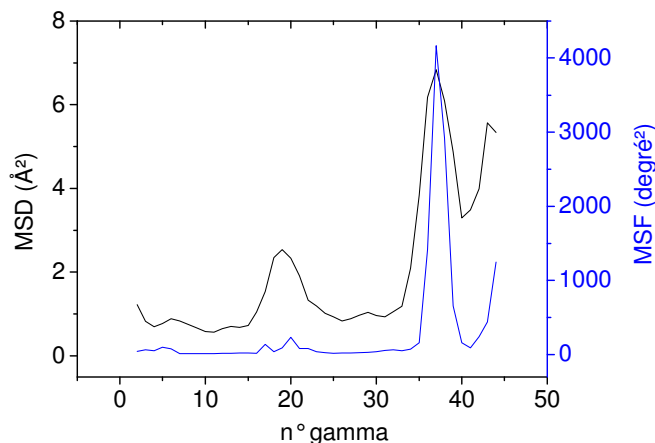


FIG. 3.40 – Variation du MSD des gammas (courbe noire) comparée à celle des fluctuations des gammas (courbe bleue) pour la crambine P22/L25 hydratée à 300K, sur une trajectoire de 1ns.

3.5.2.2 Profil d'énergie libre le long de la séquence primaire

À partir des trajectoires de dynamique moléculaire, nous avons calculé le profil d'énergie libre ou Potential of Mean Force (PMF) le long de chaque coordonnée γ_n en utilisant la procédure décrite dans la partie I, chapitre 2.

3.5.2.3 Mouvement diffusif des angles γ

Chaque angle $\gamma_n(t)$ extrait de chaque trajectoire de dynamique moléculaire reflète le taux et l'amplitude des fluctuations de la protéine dans l'environnement du résidu n .

L'exploration des valeurs possibles de $\gamma_n(t)$, i.e. l'exploration du profil d'énergie libre ou PMF, est similaire à celle d'un mouvement aléatoire, diffusif, sur un cercle. Il est caractérisé par le déplacement quadratique moyen au temps t de l'angle diédral $\gamma_{\{n\}}$ associé au résidu :

$$\Delta_{\{\gamma_{\{n\}}\}}(t) = \left\langle [\gamma_{\{n\}}(t) - \gamma_{\{n\}}(0)]^2 \right\rangle, \quad (3.4)$$

où la moyenne $\langle \rangle$ est calculée sur toutes les conditions initiales $\gamma_{\{n\}}(0)$ de la trajectoire de la même manière que dans l'équation 2.26 de la partie I, chapitre 2. La variation de l'angle diédral au cours du temps peut être décrite comme résultant du mouvement d'une particule fictive sur un cercle. $\Delta_{\{\gamma_{\{n\}}\}}(t)$ mesure la diffusion de cette particule au cours du temps. L'analyse de la loi $\Delta_{\{\gamma_{\{n\}}\}}(t)$ nous renseigne sur le mécanisme de diffusion de la particule, i.e. sur la manière dont la protéine explore ses conformations au cours du temps.

Pour le mouvement de diffusion d'une particule libre sur un cercle, le déplacement $\Delta_{\{\gamma_{\{n\}}\}}(t)$ devrait obéir à la loi d'Einstein [Perrin 28] :

$$\Delta_{\{\gamma_{\{n\}}\}}(t) = 2D_{\{n\}}t, \quad (3.5)$$

où $D_{\{n\}}$ est un coefficient de diffusion rotationnelle pour l'angle n . Comme nous le démontrons ci-dessous, l'équation d'évolution de $\Delta_{\{\gamma_{\{n\}}\}}(t)$ est en réalité :

$$\Delta_{\{\gamma_{\{n\}}\}}(t) = 2D_{\{n\}}(\beta)t^{(2\beta)}, \quad (3.6)$$

où $\beta < 0.5$ est un exposant qui varie le long de la séquence primaire. Une série temporelle qui obéit à l'équation 3.6 est dite self-affine avec une dimension fractionnaire (dite de Hurst) β [Mandelbrot et Ness 68].

Nous avons calculé l'exposant dans l'Eq. 3.6 en appliquant l'algorithme DFA (Detrended Fluctuation Analysis) des séries temporelles $\gamma_n(t)$ (cf partie I, chapitre 2) [Peng *et al.* 94]. Plus précisément, nous avons calculé l'exposant de l'intégrale des séries temporelles, soit la quantité $\alpha \approx \beta + 1$ [Hu *et al.* 01]. Pour un mouvement Brownien d'une particule libre, $\beta = 0.5$ et $\alpha = 1.5$.

L'algorithme DFA permet d'analyser les propriétés statistiques d'une série temporelle non-stationnaire, c'est-à-dire dont la valeur moyenne varie au cours du temps. C'est toujours le cas des trajectoires de dynamique moléculaire de protéines car d'une part les temps de simulations sont souvent trop courts pour atteindre un équilibre stationnaire et d'autre part car la nature anharmonique de la dynamique d'une protéine implique un tel comportement.

3.5.3 Dynamique du backbone et PMF à 300K

3.5.3.1 Crambine P22/L25

Les résultats suivants ont été réalisés sur des trajectoires de $5ns$ à $300K$, pour la crambine P22/L25. Sur la Fig. 3.41 nous représentons les trois types de comportements dynamiques que nous avons observé, et leur PMF (Fig. 3.42) : le gamma 11 a un exposant de 0.80, son PMF est harmonique, le gamma 17 a un exposant de 1.07, son PMF est anharmonique et enfin, le gamma 37 a un exposant de 1.22 avec un PMF qui est en double-puits. Pour plus de clartés dans la présentation des résultats, les PMF sont représentés en kT (Dans cet exemple à $300K$, $1kT = 0.6kcal/mol$).

Quelque soit la protéine, nous observons trois types de PMF et de dynamique :

- Si $\alpha < 1$, le PMF est harmonique avec un simple puits,
- Si $\alpha \cong 1$, le PMF est représenté par un seul minimum mais anharmonique,
- Si $\alpha > 1.1$, le PMF présente une forte anharmonicité, avec de multiples minima.

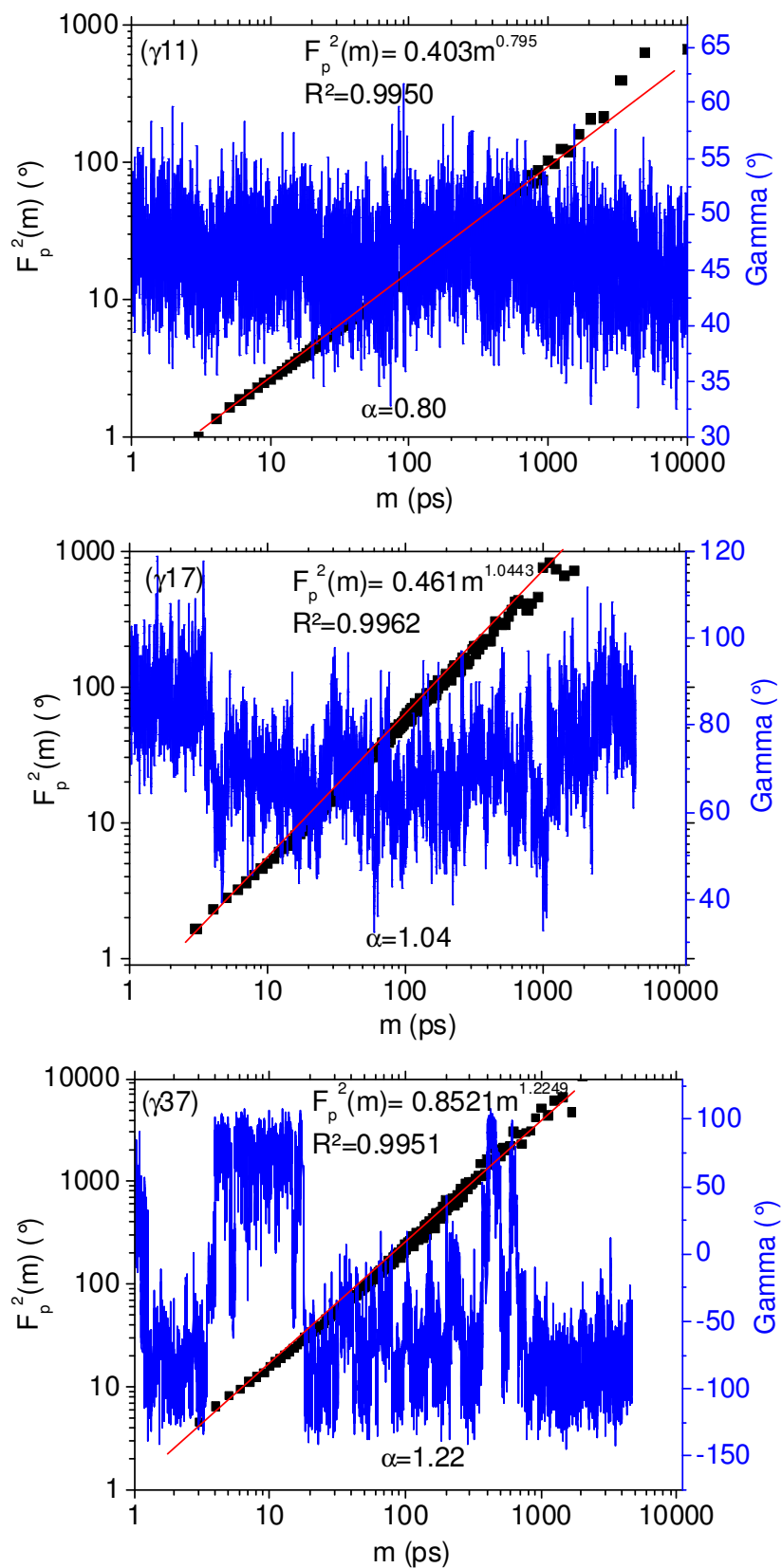


FIG. 3.41 – Lois en puissance de γ_{11} , γ_{17} et γ_{37} de la crambine P22/L25 hydratée à 300K.

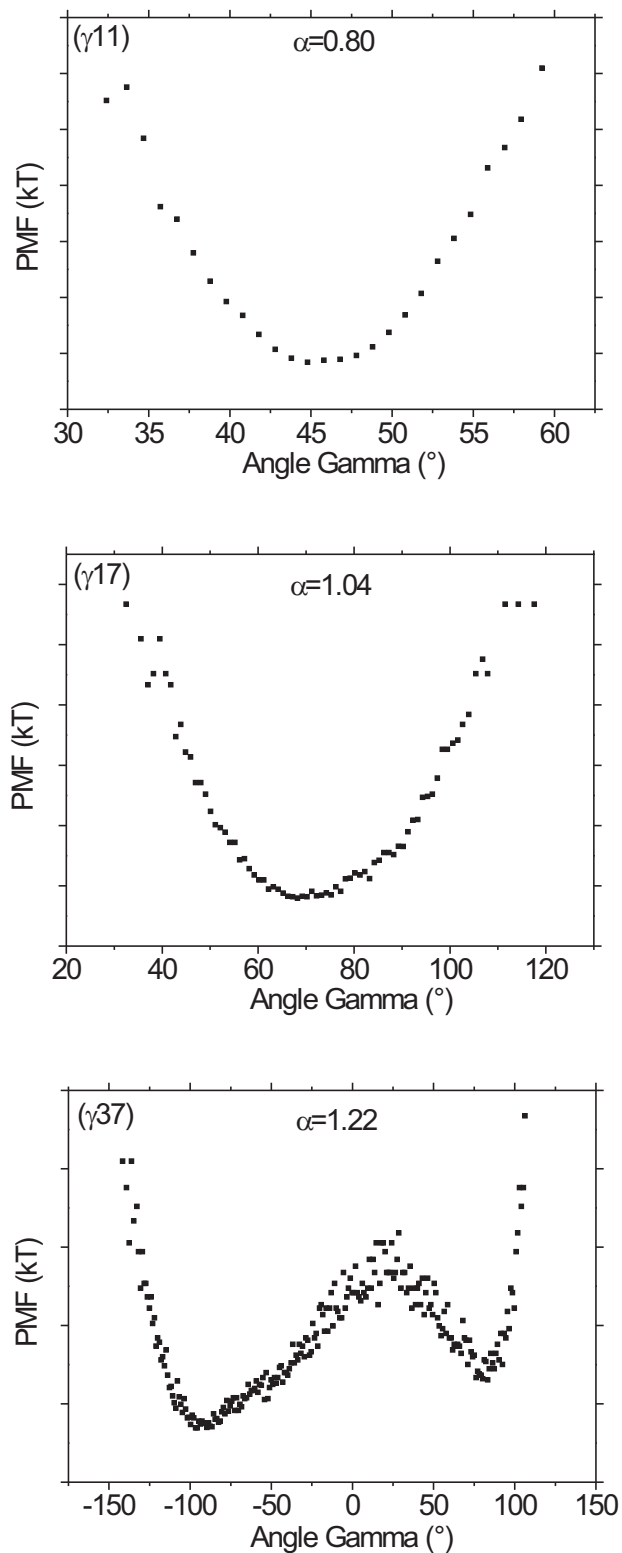


FIG. 3.42 – Représentation des 3 types de PMF possibles avec 3 angles gamma de la crambine P22/L25 hydratée à 300K sur une trajectoire de 5ns.

La Fig. 3.41 illustre les lois en puissance pour ces 3 cas différents : deux pour $\alpha < 1.1$: γ_{11} et γ_{17} et une pour $\alpha > 1.1$: γ_{37} . De cette figure, nous pouvons conclure qu'un exposant $\alpha > 1.1$, γ_{37} , correspond à une trajectoire présentant des "sauts". Aucun "saut" n'est visible pour γ_{11} ($\alpha = 0.8$), alors que pour γ_{17} ($\alpha = 1$), des "sauts" avec de courtes périodes d'"attente" sont observés. Ces observations sont en accord avec la signification de l'exposant calculé par la méthode DFA, il mesure la "rugosité" de la trajectoire. Les mêmes conclusions peuvent être tirées des trajectoires simulées pour les autres thionines.

Sur la Fig. 3.43, nous présentons les PMF et exposants de tous les gammas de la crambine P22/L25 hydratée à 300K, sur une trajectoire de 5ns. Nous pouvons constater que l'évolution des PMF le long de la séquence primaire suit la structure secondaire des thionines. En effet, les PMF des hélices alpha (représentées en jaune) ont des valeurs d'angles très proches (environ 50°). Les doubles puits interviennent pour des exposants élevés et dans les zones les plus mobiles de la protéine : dans la boucle (résidus 18 à 23) et de manière générale en fin de séquence (résidus 36 à 44).

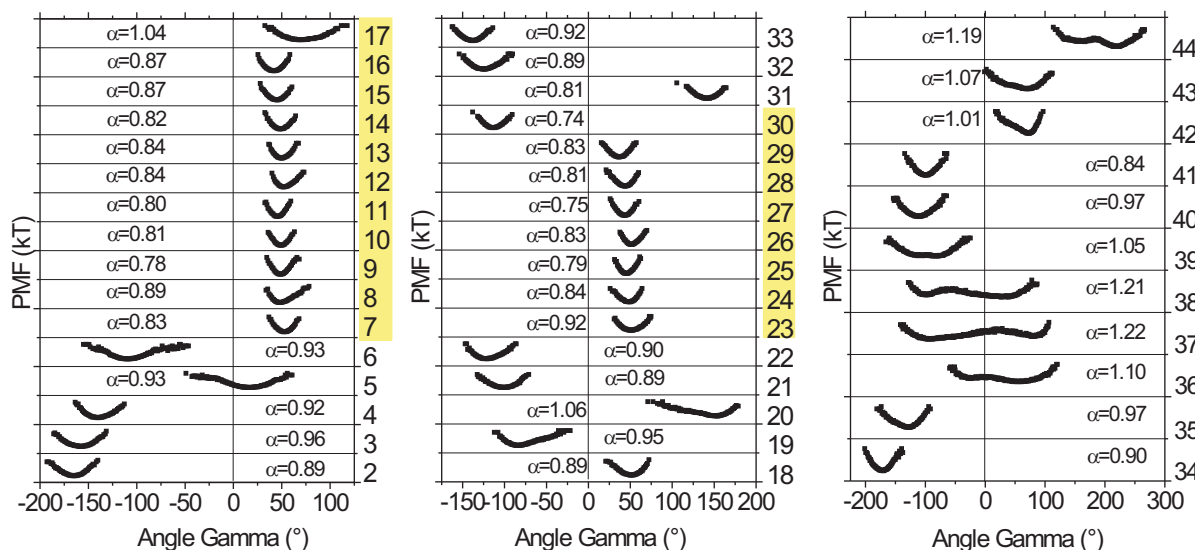


FIG. 3.43 – Variation du PMF de la crambine P22/L25 hydratée sur une trajectoire de 5ns à 300K. L'exposant α est donné pour chaque gamma, les hélices alpha sont localisées en jaune.

Note : Un angle gamma est aussi équivalent à un angle gamma $+2\pi$, mais pour plus de visibilité sur les graphes du PMF, nous n'avons pas ajouté les potentiels translatsés de 2π .

La Fig. 3.44 représente la séquence primaire des exposants α de la crambine P22/L25 hydratée à 300K. Nous différencions deux domaines :

- $\alpha > 1.1$ pour les résidus dont le PMF possède de multiples minima, et
- $\alpha < 1.1$ où le PMF est un simple puits, harmonique ou non.

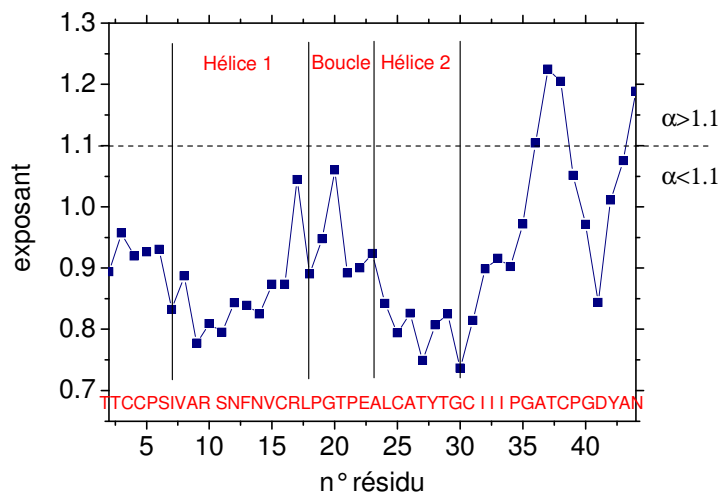


FIG. 3.44 – Séquence primaire des exposants de la crambine P22/L25 hydratée à 300K, sur une trajectoire de 5ns. Les labels rouges représentent les acides aminés de la séquence.

3.5.3.2 Effets des mutations sur le PMF

Sur la Fig. 3.45, nous présentons la séquence primaire des exposants des crambines hydratées à 300K sur une trajectoire de 1ns. Nous avons localisé les éléments des structures secondaires et la limite $\alpha > 1.1$ où les PMF deviennent multi-puits. Nous pouvons constater qu'une seule mutation dans la séquence primaire modifie l'allure de la séquence des exposants.

De manière générale, la boucle et la fin de séquence sont des zones où $\alpha > 1.1$. Nous retrouvons les parties mobiles décrites dans l'étude de l'écart quadratique moyen des protéines hydratées (cf chapitre 3, partie II).

Les différences entre P22/L25 et P22/I25 sont nombreuses. Premièrement, 5 résidus possèdent des exposants supérieurs à 1 dans l'hélice 2 de P22/I25 et pas dans P22/L25 (tous les exposants sont inférieurs à 0.9). Deuxièmement, les exposants sont plus élevés dans la boucle de P22/I25 (les exposants α de γ_{19} , γ_{21} , γ_{22} et γ_{23} sont proches de 1.1) que dans la boucle de P22/L25 (seul γ_{20} a un exposant proche de 1.1). En étudiant les PMF sur la Fig. 3.46, nous constatons que dans le cas de P22/I25, la boucle ne montre pas de multi-puits. Même si ses exposants sont supérieurs à P22/L25, il n'y a pas de multi-puits dans cette partie de la protéine.

On peut conclure qu'une seule mutation sur le résidu 25, entre deux résidus très similaires, la leucine et l'isoleucine, modifie la séquence complète des exposants, pas seulement au niveau de la mutation mais tout au long de la protéine.

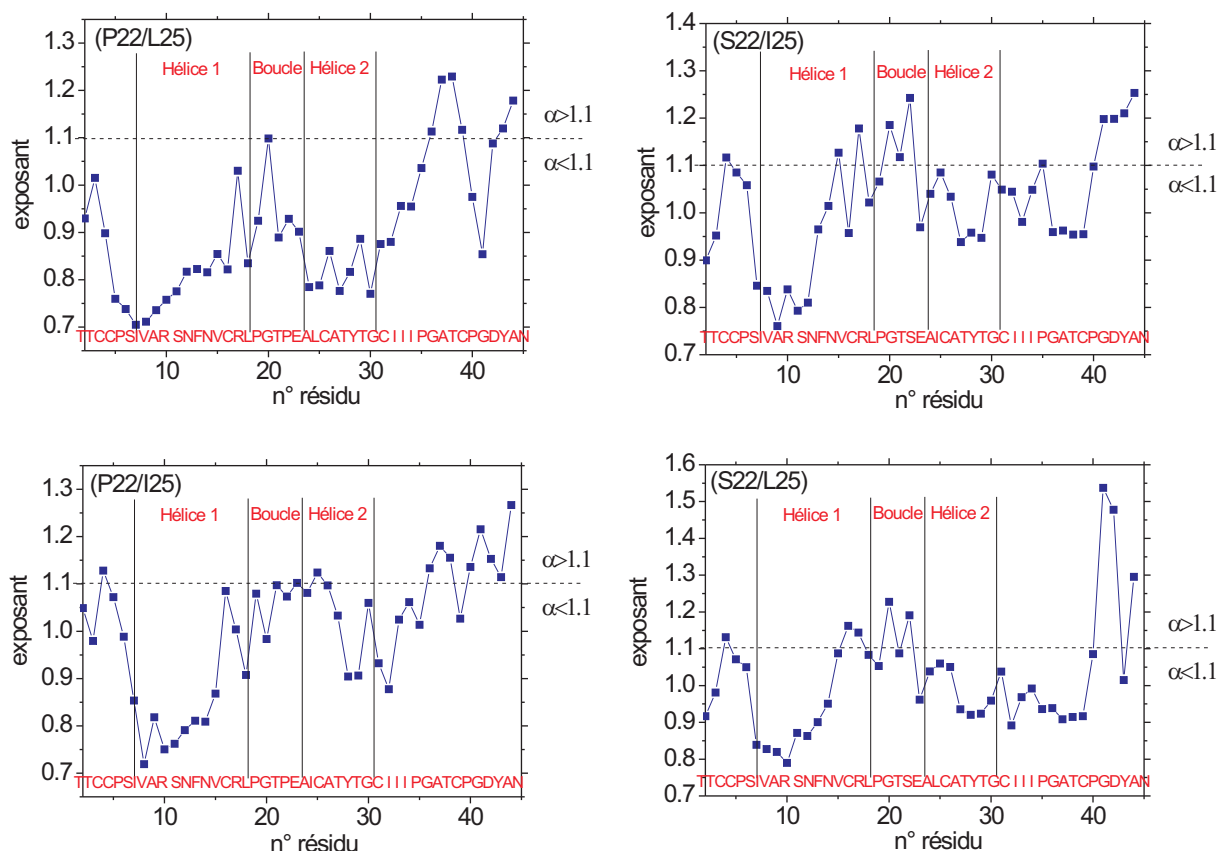


FIG. 3.45 – Séquences primaires des exposants des crambines hydratées à 300K sur une trajectoire de 1ns. Les labels rouges représentent les acides aminés de la séquence, la ligne noire en traits interrompus indique la limite pour les exposants $\alpha > 1$ et la structure secondaire est localisée.

Pour les protéines S22/L25 et S22/I25, les exposants $\alpha > 1.1$ sont globalement localisés dans les mêmes régions (g4, fin de l'hélice 1, boucle et la région C terminale). Il existe cependant des différences ; la valeur des exposants dans la région C-terminale est assez différente, par exemple, pour $\gamma 41$ et $\gamma 42$, $\alpha \simeq 1.5/1.6$ pour S22/L25 et seulement 1.2 pour S22/I25. Mais cela ne change pas les PMF dans cette région puisque les deux protéines ont des multi-puits, cf Fig. 3.47. Les mêmes résultats ont été trouvés sur les 5 conditions initiales.

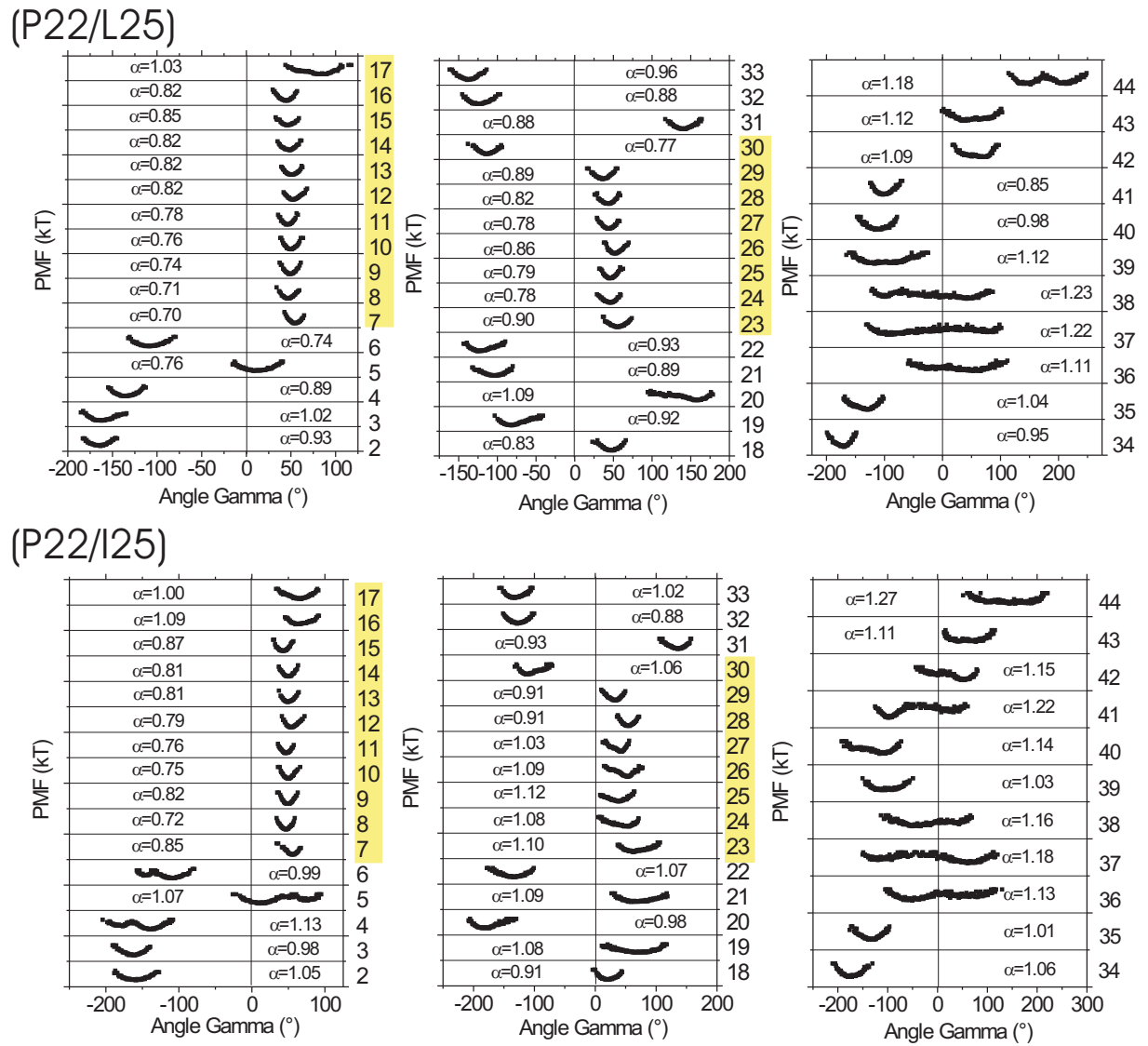


FIG. 3.46 – Comparaisons des PMF de la crambine P22/L25 en haut et de la crambine P22/I25 en bas, hydratées sur une trajectoire de 1ns à 300K. L'exposant α est donné pour chaque gamma, les hélices alpha sont localisées en jaune.

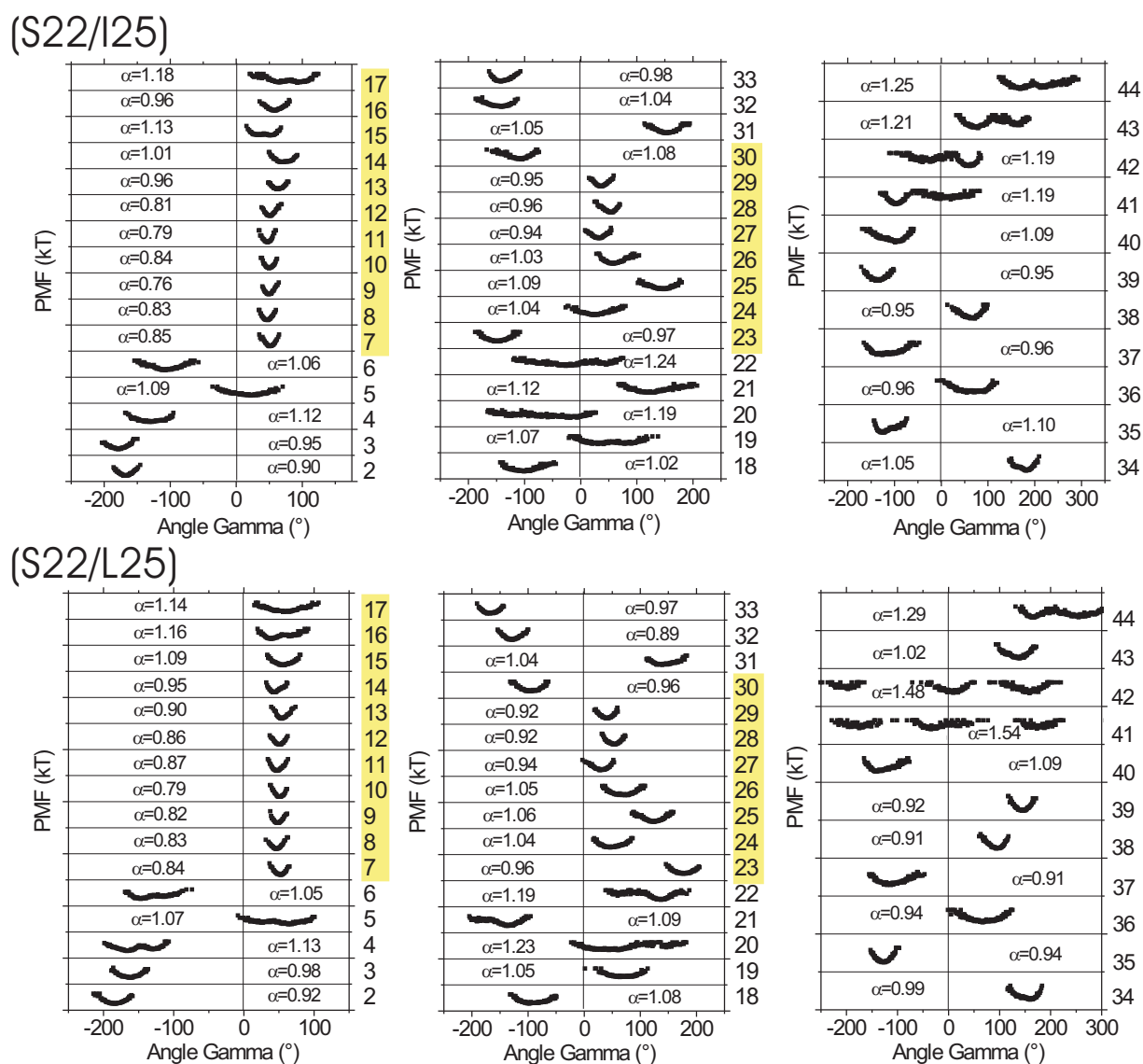


FIG. 3.47 – Comparaisons des PMF de la crambine S22/I25 en haut et de la crambine S22/L25 en bas, hydratées sur une trajectoire de 1ns à 300K. L'exposant α est donné pour chaque gamma, les hélices alpha sont localisées en jaune.

Le remplacement d'une proline par une serine en position 22 a un effet structural important, visible dans le calcul des angles diédraux gamma sur la Fig. 3.47, sur l'hélice 2. En effet, le résidu 22 a un double puits et les résidus dans l'hélice 2 ont des minima qui ne sont plus alignés sur la position d'environ 50° (correspondant à une hélice α droite idéale). La proline 22 stabilise l'hélice α 2, la serine a tendance à déstabiliser cette structure. Il faut rappeler que l'hélice 2 dans les crambines est liée par des ponts disulfures (liaisons entre les cystéines 16 et 26) à l'hélice 1, la dénaturation de la structure tertiaire n'est pas possible.

3.5.4 Transition dynamique du backbone

3.5.4.1 Exposants et anharmonicité

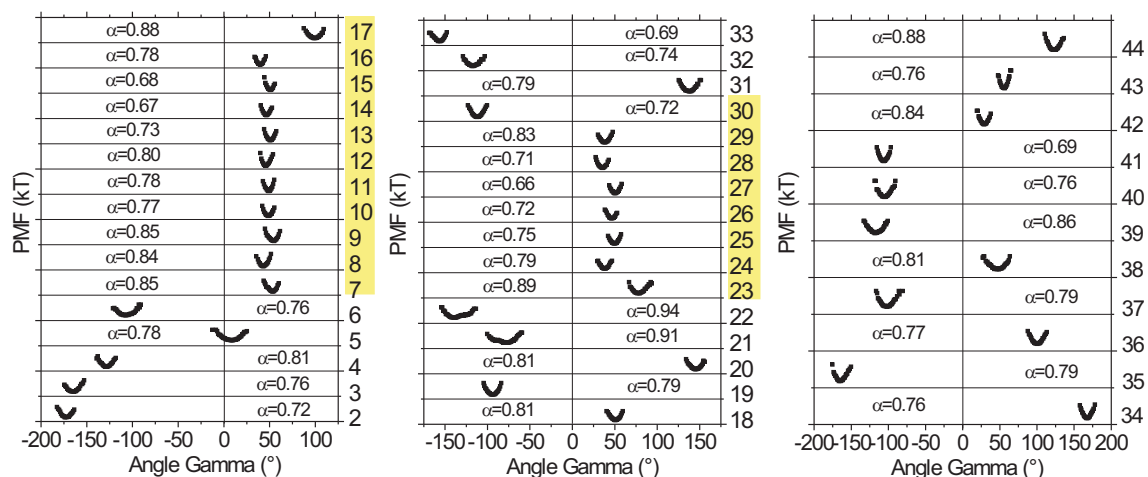
Afin d'étudier les modifications des profils d'énergie libre avec l'élévation de la température, nous représentons sur la Fig. 3.48 les PMF de la crambine P22/L25 à 150K, 220K et 300K, sur 1ns et ceux de la viscotoxine B sur la Fig. 3.49. Ces analyses sont caractéristiques des résultats calculés pour les autres thionines étudiées.

Nous observons clairement l'apparition de potentiels avec plusieurs minima au-delà de la température de transition dynamique (220K et 300K). En-dessous de cette température, tous les puits de potentiels sont harmoniques. Alors que ces puits apparaissent généralement dans les boucles, il faut noter pour VB l'apparition de plusieurs minima dans l'hélice 2 à 220K. Ces minima disparaissent à 300K. De nouveau, nous avons observé que la dynamique suit une loi en puissance (Equation 2.33 de la partie I) ; les exposants les plus élevés (environ 1.1 et supérieurs) correspondent à une dynamique où les puits de potentiel ont plusieurs minima.

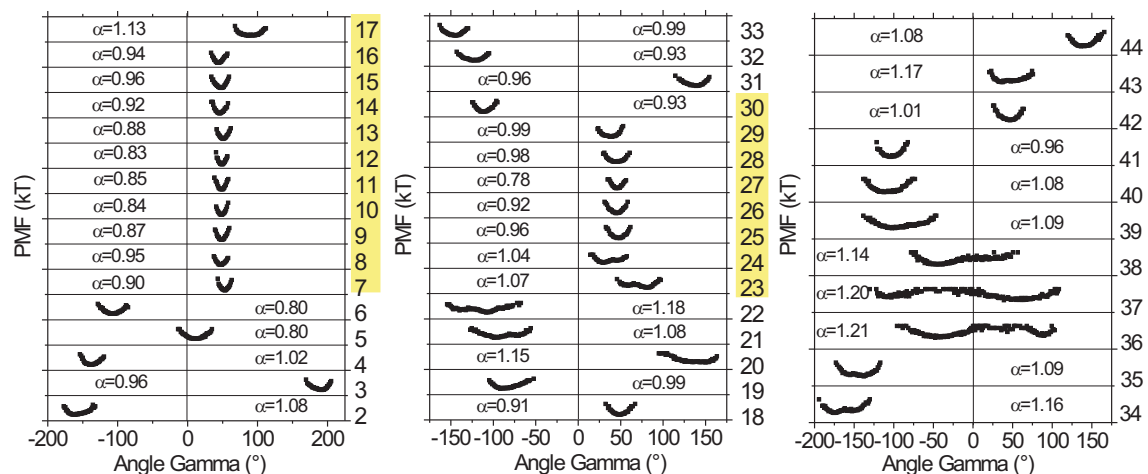
L'étude des PMF des angles diédraux γ du backbone en fonction de la température nous permet de constater la formation de potentiels anharmoniques avec un seul minimum (soit les gammas 10 et 11 pour crambine P22/L25 et les gammas 28 et 29 pour VB) ou plusieurs minima au-delà d'une température de 150K (soit les gammas 37 et 38 pour crambine P22/L25 et les gammas 20 et 21 pour VB). En-dessous de 150K, tous les profils d'énergie libre de la viscotoxine B et crambine P22/L25 sont harmoniques. Le nombre d'angles diédraux γ ayant un profil d'énergie libre anharmonique augmente avec la température.

Non seulement le profil d'énergie libre mais aussi la manière dont la protéine explore ses profils d'énergie libre se modifient avec la température. Nous pouvons constater que les exposants α , qui caractérisent la diffusion sur ces PMF (cf Eq. 3.6, p.165), sont tous inférieurs à 1.0 à T plus petit ou égaux à 150K, c'est-à-dire dans le régime harmonique. Lorsque les PMF ont plusieurs minima, au-delà de 150K, les exposants α deviennent supérieurs à 1.1.

(150K)



(220K)



(300K)

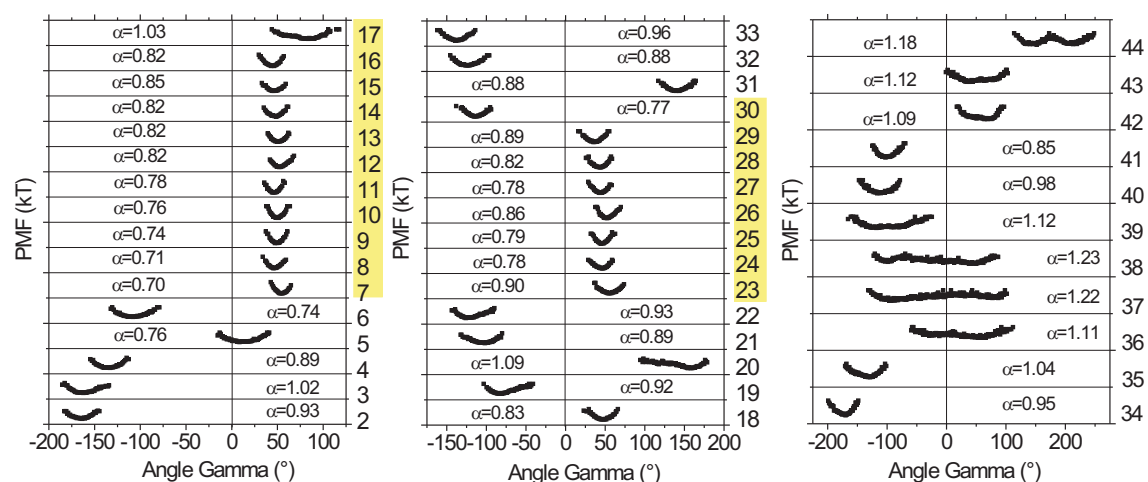
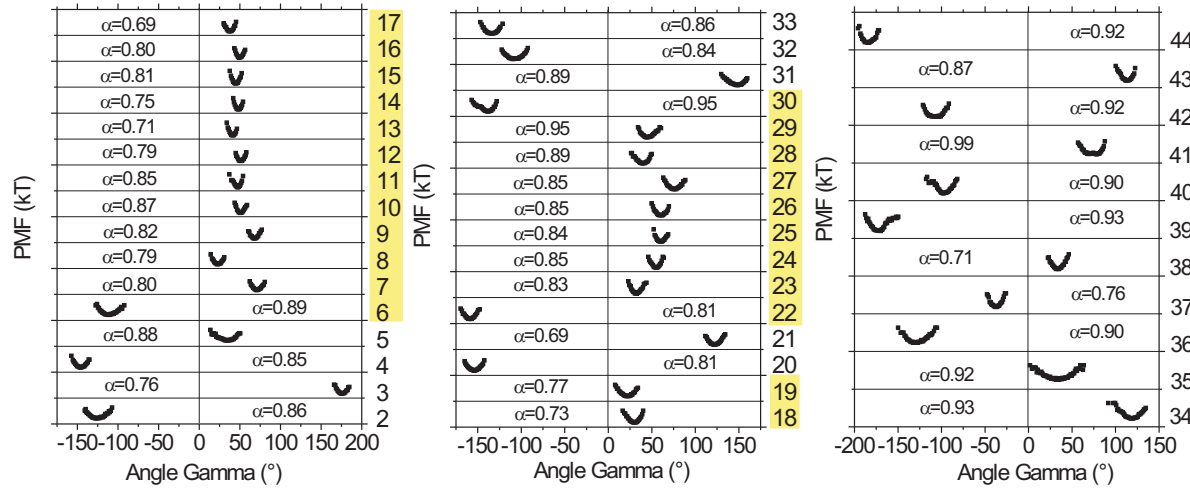
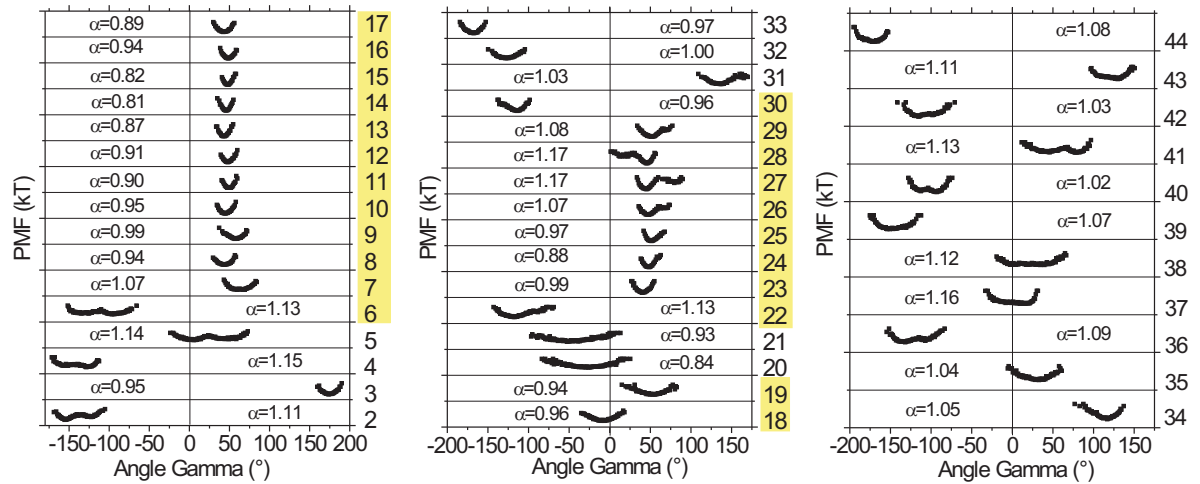


FIG. 3.48 – Variation du PMF de la crambine P22/L25 hydratée sur une trajectoire de 1ns en fonction de la température : la première cartographie correspond à 150K, la seconde à 220K et la dernière à 300K. L'exposant α est donné pour chaque gamma, les hélices alpha sont localisées en jaune.

(150K)



(220K)



(300K)

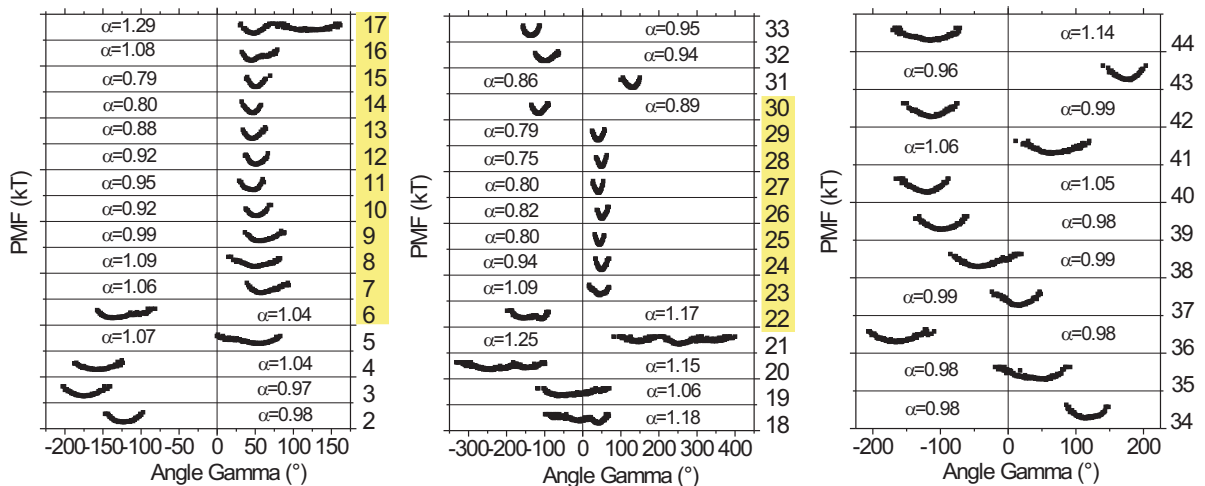


FIG. 3.49 – Variation du PMF de la viscotoxine B hydratée sur une trajectoire de 1ns en fonction de la température : la première cartographie correspond à 150K, la seconde à 220K et la dernière à 300K. L'exposant α est donné pour chaque gamma, les hélices alpha sont localisées en jaune.

Une exploration Brownienne des PMF correspondrait à un exposant $\alpha = 1.5$ (cf paragraphe 3.5.2.3, [Mandelbrot et Ness 68]). Nous pouvons donc conclure de notre étude que tous les mouvements du backbone associés aux angles diédraux γ sont sous-diffusifs quelque soit la température étudiée (en-dessous de $300K$). Des mouvements sous-diffusifs (« diffusion anormale », [Bouchaud et Georges 90]) pourraient être une caractéristique générale de la dynamique des protéines. Le phénomène de diffusion anormale a été observé pour les fluctuations de la distance entre deux résidus dans un complexe de protéine entre $1ms$ et $20s$ [Min *et al.* 05], pour la diffusion de l'eau à l'interface entre une protéine et l'eau d'hydratation sur une échelle de temps inférieure à $20ps$ [Settles et Doster 96], pour la dynamique de filaments d'actine sur une échelle de temps de la seconde [LeGoff *et al.* 02] et théoriquement dans différentes simulations de dynamique moléculaire sur des échelles de temps de la nanoseconde [Luo *et al.* 06, Kneller 04, Neusius *et al.* 08].

Cependant, le mécanisme conduisant à une diffusion anormale peut avoir de multiples origines. Dans notre cas, il est vraisemblable qu'il soit lié au phénomène de « piégeage » de la structure dans un minimum local (sous-états « cachés ») [Luo *et al.* 06] qui n'est pas visible sur le profil d'énergie libre à 1-D représenté sur la Fig. 3.39.

Lorsque l'angle « diffuse » le long du profil d'énergie, ou PMF à 1-D, la structure diffuse en réalité sur une surface complexe multidimensionnelle. Il existe de nombreux chemins différents dans cet espace multidimensionnel qui permettent de passer d'une valeur de l'angle diédral à une autre le long du PMF de l'angle. Le long de certains de ces chemins, l'angle γ se trouvera bloqué à une valeur donnée, correspondant à un minimum local du paysage multidimensionnel d'énergie libre, retardant sa « diffusion » le long du profil et conduisant ainsi à une diffusion plus lente que Brownienne ou « anormale ». Dans notre cas, on observe aussi directement un piégeage de la structure dans les « vrais » minima du PMF, les positions les plus probables de la structure par définition.

3.5.4.2 Origine des mouvements sous-diffusifs

Dans ce paragraphe, nous allons étudier plus en détails l'évolution de la dynamique et du PMF de 4 angles diédraux caractéristiques de la dynamique de l'ensemble du backbone ; soit les angles γ_{37} (PMF multi-minima), γ_{38} (PMF multi-minima), γ_{10} (PMF harmonique), γ_{11} (PMF harmonique) pour la crambine P22/L25 et les angles γ_{20} (PMF multi-minima), γ_{21} (PMF multi-minima), γ_{28} (PMF harmonique), γ_{29} (PMF harmonique) pour la viscotoxine B.

Nous avons représenté l'évolution du PMF avec la température de ces angles pour la crambine P22/L25 et pour la viscotoxine B sur les Fig. 3.50 et 3.51, respectivement. Les PMF à $150K$ sont représentés en bleu, à $220K$ en rouge et en noir à $300K$.

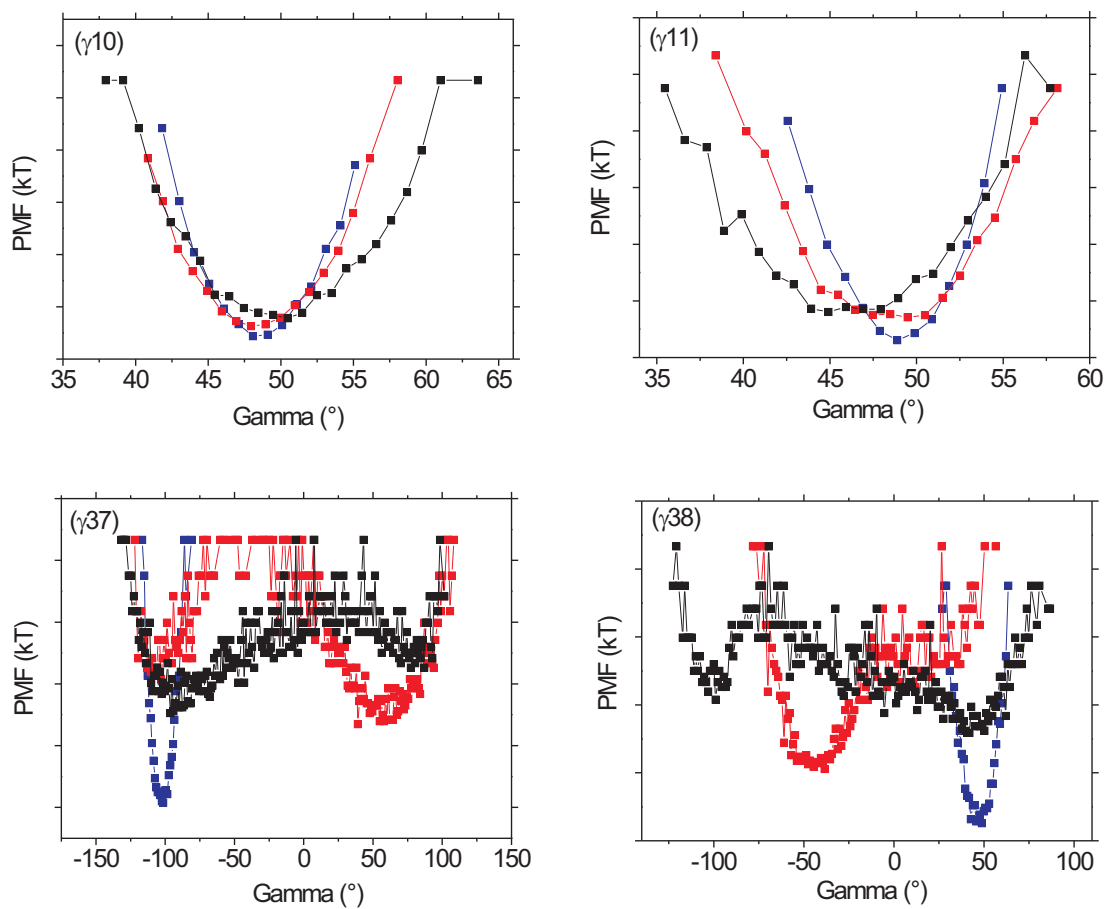


FIG. 3.50 – Représentation des PMF de la crambine P22/L25 à 150K (en bleu), 220K (en rouge) et 300K (en noir) pour deux gammas harmoniques (en haut, n°10 et 11) et deux gammas anharmoniques (en bas, n°37 et 38).

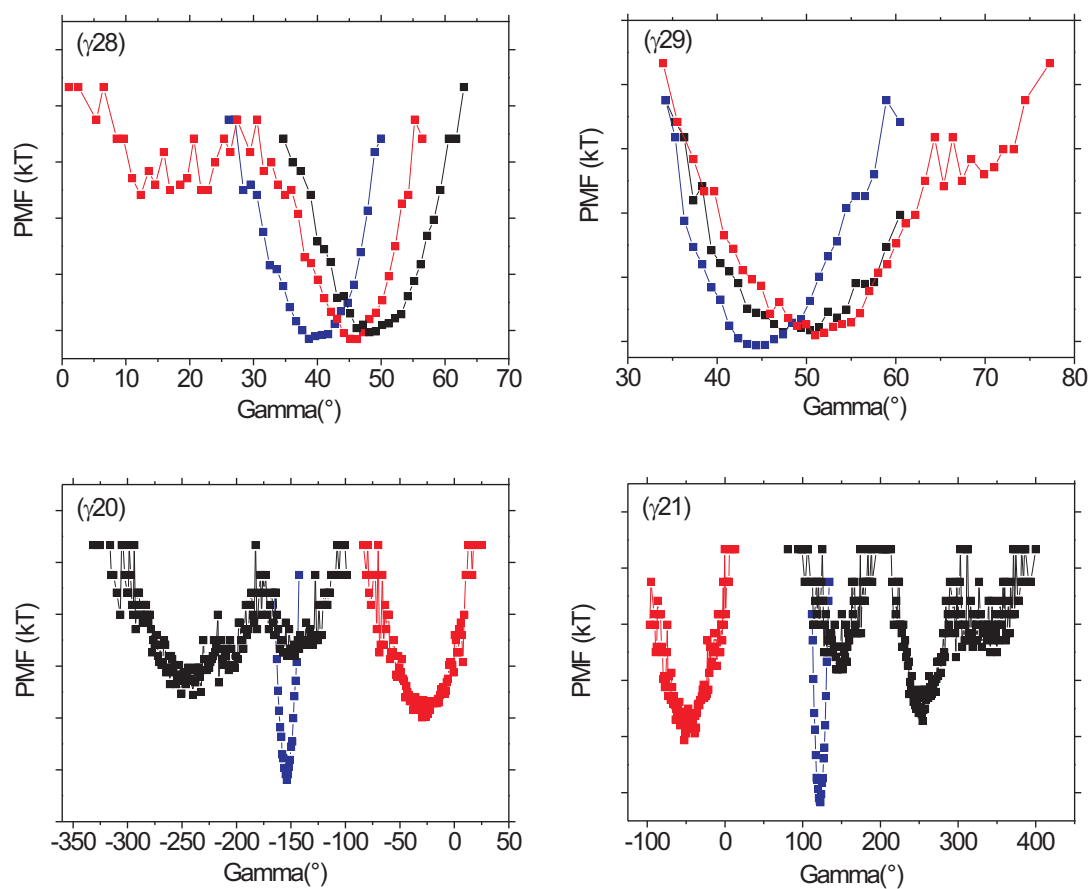


FIG. 3.51 – Représentation des PMF de la viscotoxine B à 150K (en bleu), 220K (en rouge) et 300K (en noir) pour deux gammas harmoniques (en haut, n°28 et 29) et deux gammas anharmoniques (en bas, n°20 et 21).

Nous discutons d'abord des résultats de la crambine. Nous constatons la présence d'un puits harmonique à 150K pour tous les angles, puis avec l'élévation de la température, les PMF s'élargissent et on observe la formation d'un double puits à 300K pour γ_{37} et γ_{38} .

Sur la Fig. 3.52, nous avons également représenté les variations de l'écart quadratique moyen (MSD) des gammas 10, 11, 37 et 38 correspondants à ces angles diédraux (le MSD est la moyenne du MSD des 4 atomes de carbone alpha intervenants dans le gamma).

Nous constatons que les gammas 37 et 38 subissent une transition dynamique prononcée. Pour les gammas 10 et 11, leur MSD ne dépasse pas 1.0\AA , ils n'ont donc pas de transition dynamique même si on observe une élévation de la mobilité. Il faut noter la grande différence d'amplitude des fluctuations (environ un facteur 9) entre les gammas 10, 11 (harmoniques) et 37, 38 (anharmoniques). La forte rupture de pente des gammas 37 et 38, et donc leur transition dynamique, est due aux mouvements diffusifs entre les deux minima de leurs PMF.

Par contre, les fluctuations des gammas 10 et 11 restent confinées dans un seul puits, qui s'élargit avec la température (le changement de constante de force visualisé par la régression linéaire de la courbe en traits rouges interrompus montre le changement de pente).

On observe bien ici un changement de constante de force « effective » pour les 4 gammas pour deux raisons différentes :

- 1) Le changement de profil du potentiel et des constantes de force, mais nous sommes toujours en présence d'un régime harmonique et une amplitude « faible » pour les gammas 10 et 11,
- 2) Le changement très important du MSD (transition dynamique) due à la création d'un chemin entre deux sous-états conformationnels donne des mouvements diffusifs pour les gammas 37 et 38. Nos résultats rendent concret le modèle de Doster (cf chapitre 3, partie II). Mais la réalité est plus complexe que ce modèle effectif car la barrière de Doster moyenne toutes les barrières entre les minima locaux de tous les degrés de liberté de la protéine.

Nous discutons maintenant des résultats de la viscotoxine B (cf Fig. 3.53).

Nous observons une grande similitude de comportements ; les gammas avec un potentiel en multi-puits (γ_{20} et γ_{21}) ont un MSD qui subit une transition dynamique marquée contrairement aux gammas qui possèdent un seul minimum et qui sont fortement harmoniques (γ_{28} et γ_{29}). Notons que le changement de pente des gammas γ_{28} et γ_{29} peut également s'interpréter à partir du changement de courbure de leur PMF et de l'élargissement du puits avec la température.

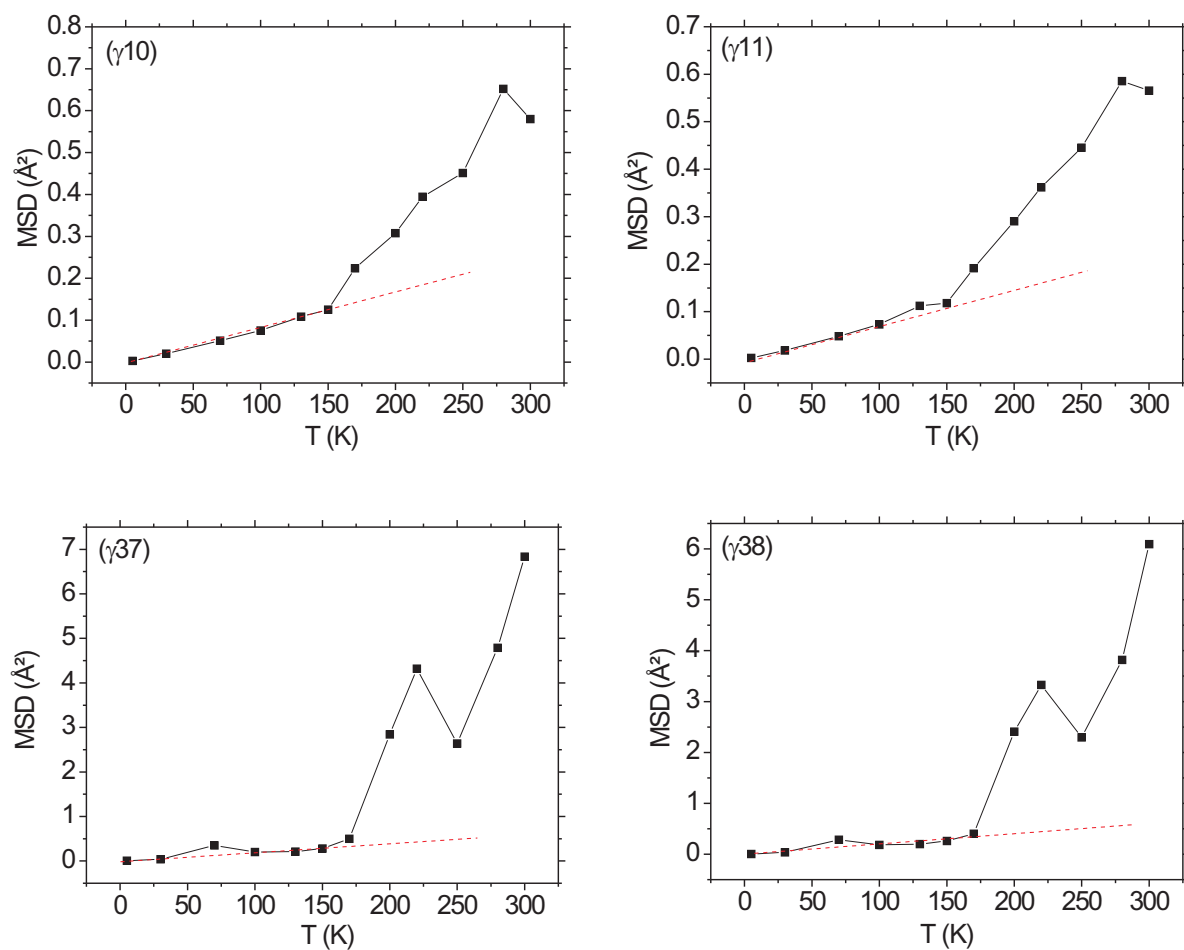


FIG. 3.52 – Variation de l'écart quadratique moyen des gammas 10 et 11, équivalent aux gammas harmoniques et des gammas 37 et 38, pour les gammas anharmoniques, de la crambine P22/L25 hydratée sur une trajectoire de 1ns. Les lignes rouges en traits interrompus correspondent aux régressions linéaires de la courbe dans l'état R.

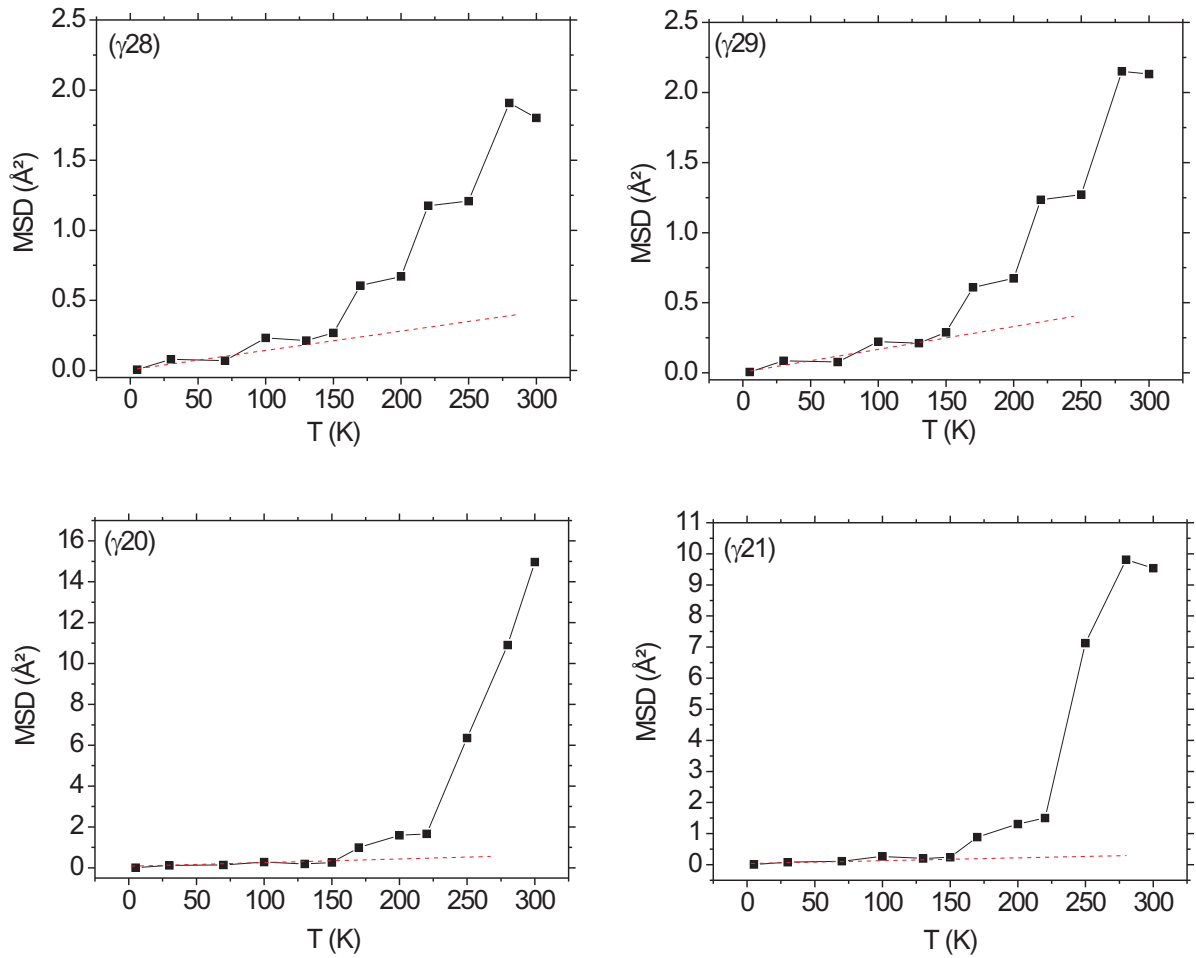


FIG. 3.53 – Variation de l'écart quadratique moyen des gammas 28 et 29, équivalent aux gammas harmoniques et des gammas 20 et 21, pour les gammas anharmoniques, de la viscotoxine B hydratée sur une trajectoire de 1ns. Les lignes rouges en traits interrompus correspondent aux régressions linéaires de la courbe dans l'état R.

3.5.5 Effet de l'hydratation sur la structure

Les Fig. 3.54, 3.55 et 3.56 comparent les PMF de la crambine P22/L25 dans le vide et en solution ($h = 4$). A 150K, 220K et 300K, nous constatons que la succession des PMF le long de la séquence primaire de la crambine simulée dans le vide ne permet pas de mettre clairement en évidence les éléments de structures secondaires. En effet, les minima des PMF des résidus formant les hélices se superposent pour la protéine en solution à toutes les températures mais pas pour la protéine dans le vide. Par exemple, à 300K, les résidus des hélices de la protéine dans le vide ont des angles diédraux γ qui varient entre 40° et 100° pour γ_{23} à γ_{30} alors qu'ils sont tous situés à environ 50° dans l'eau, mettant ainsi en évidence la localisation de l'hélice.

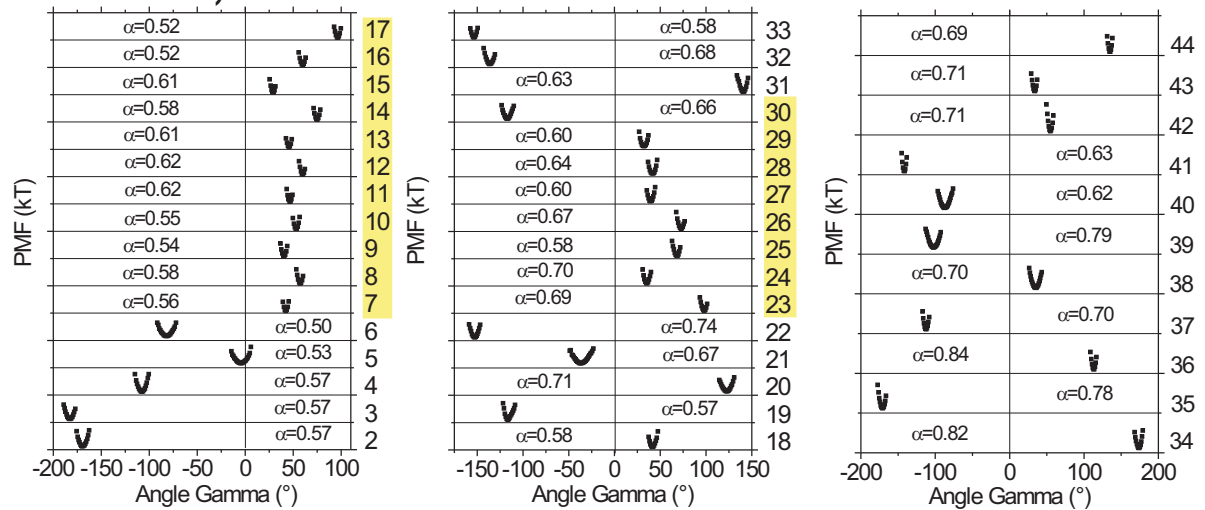
Il est intéressant de constater que le rôle organisateur de l'eau (renforcement de la structure des hélices α 1 et 2) a les mêmes conséquences que l'insertion d'une proline à la place d'une serine en position 22 (cf Figs. 3.54, 3.55, 3.56, 3.46 et 3.47).

A 150K, les exposants pour la protéine dans le vide et dans l'eau ont le même ordre de grandeur (l'exposant maximum dans le vide est égal à 0.84 et dans l'eau est de 0.94) puisqu'à cette température, les puits de potentiel sont tous harmoniques.

A 220K, nous n'observons pas de PMF avec plusieurs minima dans le vide. L'élévation de la température a tendance à augmenter tous les exposants. Ils sont tous inférieurs à 0.9, sauf pour le γ_{18} qui vaut 1.15 mais avec un simple puits harmonique (c'est la seule exception). Pour la protéine en solution, la situation est différente ; plusieurs résidus ont un PMF anharmonique (par exemple le γ_{17} avec $\alpha = 1.13$) et 8 de ces résidus ont un PMF avec plusieurs minima ($\gamma_{17}, \gamma_{20}, \gamma_{22}, \gamma_{34}, \gamma_{36}, \gamma_{37}, \gamma_{38}$ et γ_{43}). Par exemple, le γ_{36} a un PMF avec de multiples minima. L'exposant α est 1.21 dans l'eau mais 0.69 dans le vide avec un PMF harmonique.

A 300K, nous n'observons toujours pas de PMF avec plusieurs minima dans le vide et tous les potentiels restent harmoniques sauf le γ_{18} ($\alpha = 1.08$) qui est anharmonique avec un seul minimum. Tous les exposants dans le vide sont inférieurs à 0.9 (sauf pour le γ_{18}). Par exemple, le PMF du résidu 38 a un double puits dans l'eau ($\alpha = 1.23$) et est harmonique dans le vide ($\alpha = 0.55$).

(VIDE, 150K)



(HYDRATEE, 150K)

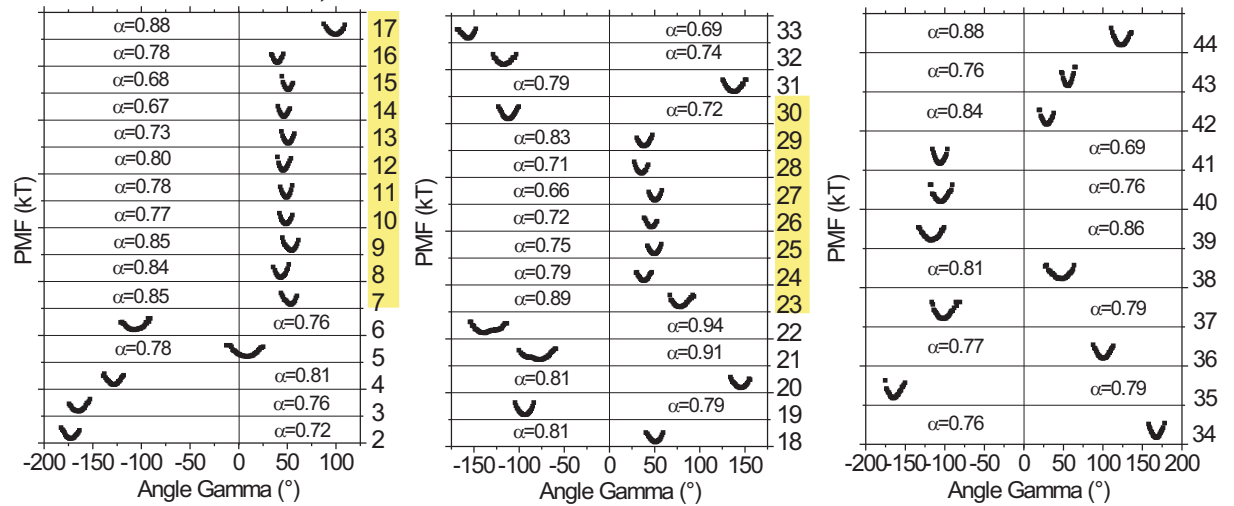
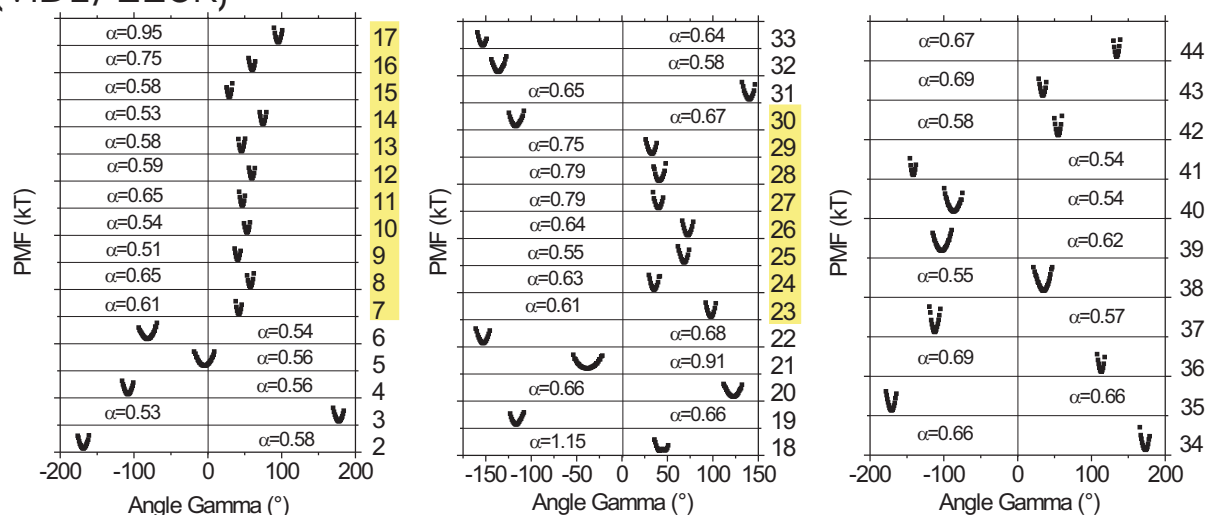


FIG. 3.54 – Variation du PMF de la crambine P22/L25 à 150K sur une trajectoire de 1ns : la première cartographie correspond à la protéine dans le vide et la seconde à la protéine hydratée. L'exposant α est donné pour chaque gamma, les hélices alpha sont localisées en jaune.

(VIDE, 220K)



(HYDRATEE, 220K)

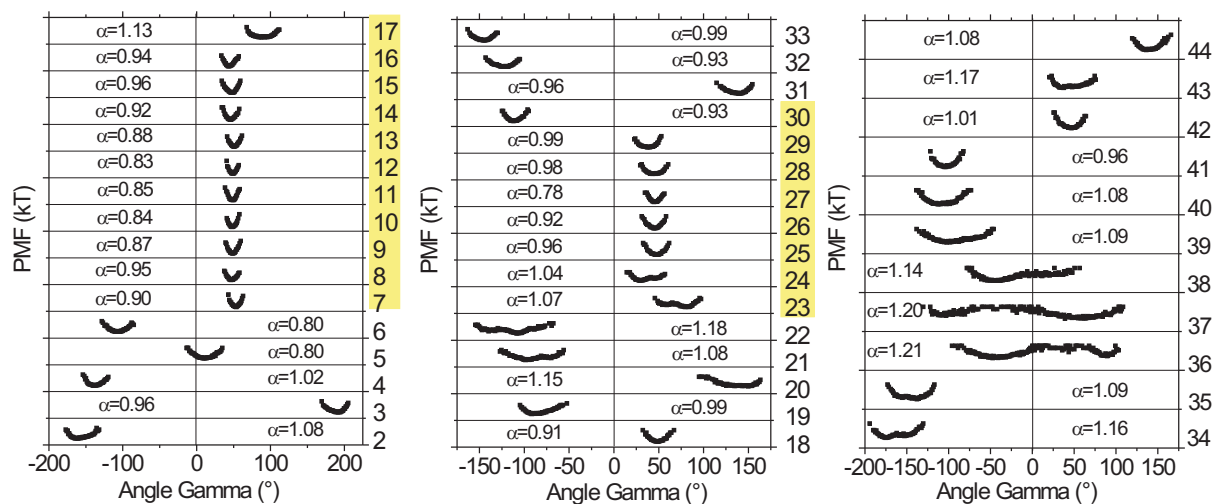
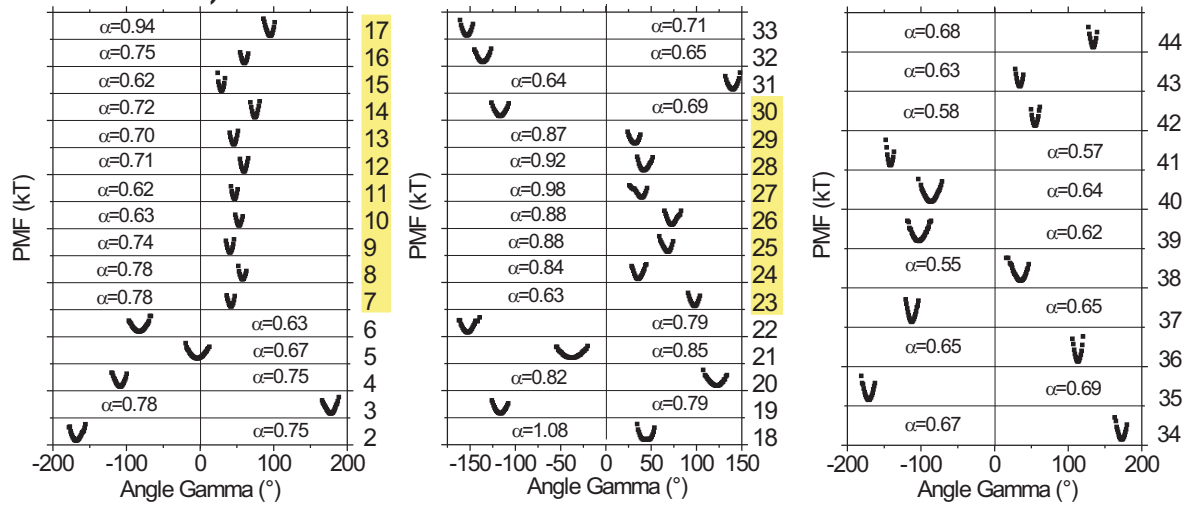


FIG. 3.55 – Variation du PMF de la crambine P22/L25 à 220K sur une trajectoire de 1ns : la première cartographie correspond à la protéine dans le vide et la seconde à la protéine hydratée. L'exposant α est donné pour chaque gamma, les hélices alpha sont localisées en jaune.

(VIDE, 300K)



(HYDRATEE, 300K)

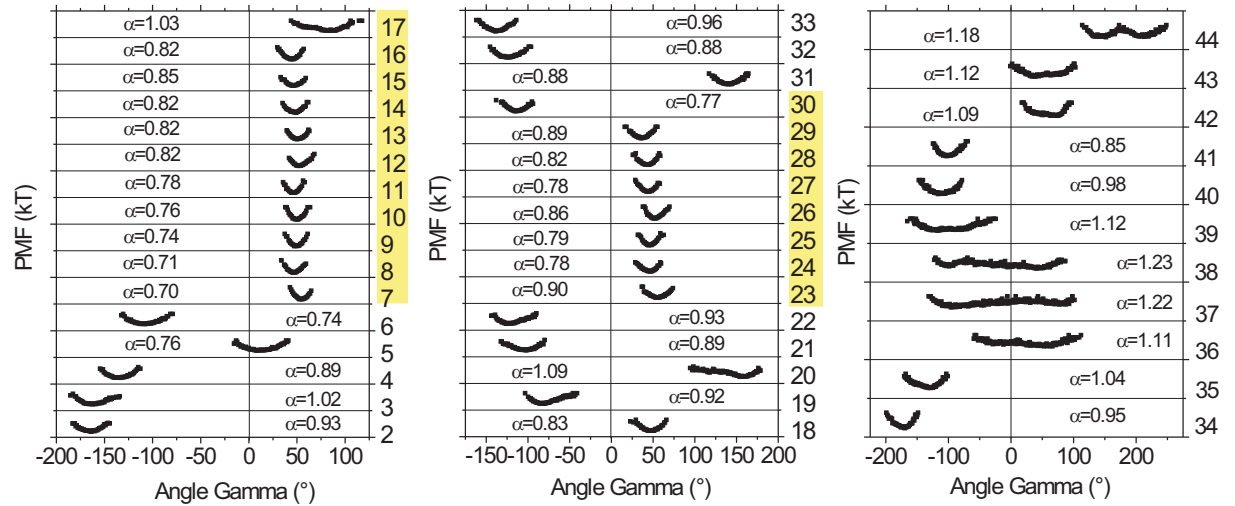


FIG. 3.56 – Variation du PMF de la crambine P22/L25 à 300K sur une trajectoire de 1ns : la première cartographie correspond à la protéine dans le vide et la seconde à la protéine hydratée. L'exposant α est donné pour chaque gamma, les hélices alpha sont localisées en jaune.

3.6 Modèle de la transition dynamique

A partir de nos résultats, nous pouvons ébaucher un modèle de la transition dynamique.

La dynamique des protéines peut être décomposée en deux parties : la première correspond à un terme vibrationnel, qui augmente linéairement avec la température et la seconde est une constante diffusive, avec une évolution non-linéaire.

Nos résultats peuvent se résumer schématiquement par sur figure 3.57. Les protéines présentent deux types de résidus. Le premier a un comportement harmonique, comme le gamma 4 de la crambine P22/L25 hydratée à gauche, son puits s'élargit avec l'augmentation de la température et donc avec la transition dynamique mais il possède un seul minimum. Le second a un comportement anharmonique, comme le gamma 44 à droite. La transition dynamique provoque la formation d'un double-puits qui représente des sous-états conformationnels. Il existe des chemins ouverts vers d'autres états. Ses chemins sont reliés à l'eau, puisque nous avons pu voir que ces multi-puits n'existent pas dans le cas des protéines déshydratées. De plus, il faut noter que la transition dynamique se produit à la température de diffusion de l'eau. Nous avons vu également que l'eau interfaciale a un comportement différent de l'eau de volume et dont les propriétés dynamiques jouent très certainement un rôle dans celles des protéines hydratées. L'eau crée donc des chemins entre les nouveaux sous-états conformationnels. Avec un logiciel de visualisation, nous avons pu constater que les minima des PMF correspondent à des basculements de chaînes dans leur totalité.

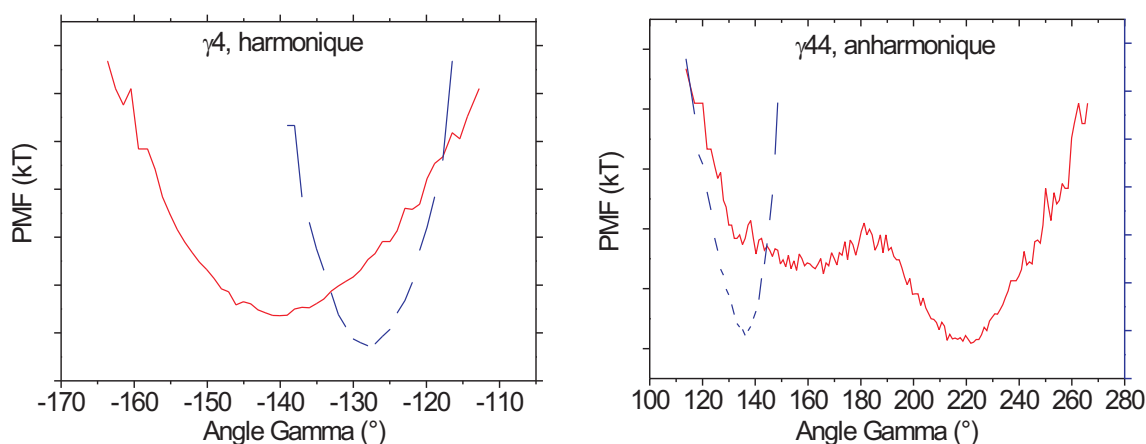


FIG. 3.57 – Evolution du PMF entre 150K (courbe bleue en traits interrompus) et 300K (courbe rouge), pour deux gammas de la crambine P22/L25 hydratée sur une trajectoire de 5ns : g4, gamma harmonique, à gauche et g44, gamma anharmonique, à droite.

Les résidus susceptibles de bouger sont situés préférentiellement à la surface de la protéine. L'agitation des molécules d'eau entraîne plus facilement des résidus proches, par rapport à ceux situés au coeur des thionines mais aussi ceux présentant des chaînes latérales avec un fort encombrement dans l'espace, comme nous l'avons vu dans la section 3.2.4 de la partie II. Plus les mouvements provoqués par le solvant seront grands, plus les sauts entre les différents sous-états conformationnels seront possibles. Il est difficile de prédire quel type d'acide aminé

sera mobile, mais dans le cas des viscotoxines hydratées, nous remarquons que la plupart de leurs résidus considérés comme mobiles sont chargés (6 sur les 7 pour VA3, 3 sur les 5 pour VA2 et 3 sur les 4 pour VB).

Dans le vide, les résidus hydrophobes sont le plus souvent les plus mobiles. Tous les résidus ont un puits harmonique pour le backbone mais les chaînes latérales existent dans différents rotamers. Par exemple, les méthyles ont 3 minima locaux (cf chapitre 2, partie II).

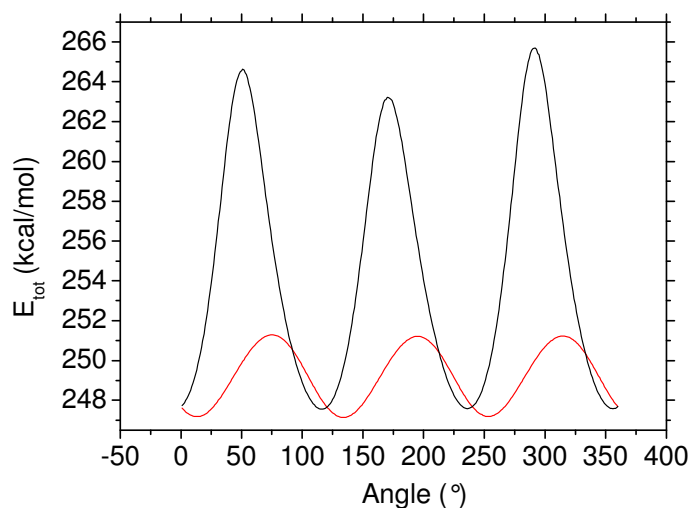


FIG. 3.58 – Barrières d'énergie pour deux méthyles de la structure optimisée de la crambine P22/L25 dans le vide : méthyle du résidu THR2 en noir et méthyle du résidu VAL8 en rouge.

La Fig. 3.58 montre un méthyle avec une haute barrière (courbe noire, résidu THR2) et un méthyle avec une petite barrière (courbe rouge, résidu VAL8). A la transition dynamique, les barrières ne changent pas de hauteur mais la probabilité de sauter cette barrière augmente d'après la loi d'Arrhenius. Ceci provoque l'activation des mouvements des méthyles (cf chapitre 2). Dans l'eau, ce mouvement de rotation existe encore mais sa contribution est masquée par l'augmentation de la mobilité du backbone et des chaînes latérales due à l'hydratation.

Une seule mutation peut changer le comportement dynamique des protéines. Nous avons pu voir que les exposants sont modifiés sur toute la séquence, pas seulement au niveau de la mutation et l'allure des PMF peut également changer. Cependant, les parties mobiles restent les mêmes d'une protéine à une autre.

Les ions comme sondes de la transition dynamique

Lors des simulations des viscotoxines hydratées, nous avons ajouté des ions (Cl^-) afin de neutraliser le système. Nous avons alors tracé la moyenne de l'évolution de l'écart quadratique moyen obtenue sur les 5 trajectoires de 1ns générées.

Sur la Fig. 4.1, nous constatons que les ions ont une transition dynamique très prononcée. Cette transition démarre un peu plus tard que celle de la protéine (soit $\sim 200\text{K}$, au lieu de 170K pour les viscotoxines). En analysant leurs mouvements, nous avons pu remarquer que ces ions sont proches des sites mobiles. Par exemple, la mobilité des ions de VA2 est principalement due à un ion placé près du résidu ARG23, l'ion Cl^- subit alors les mêmes mouvements de basculement que ce résidu.

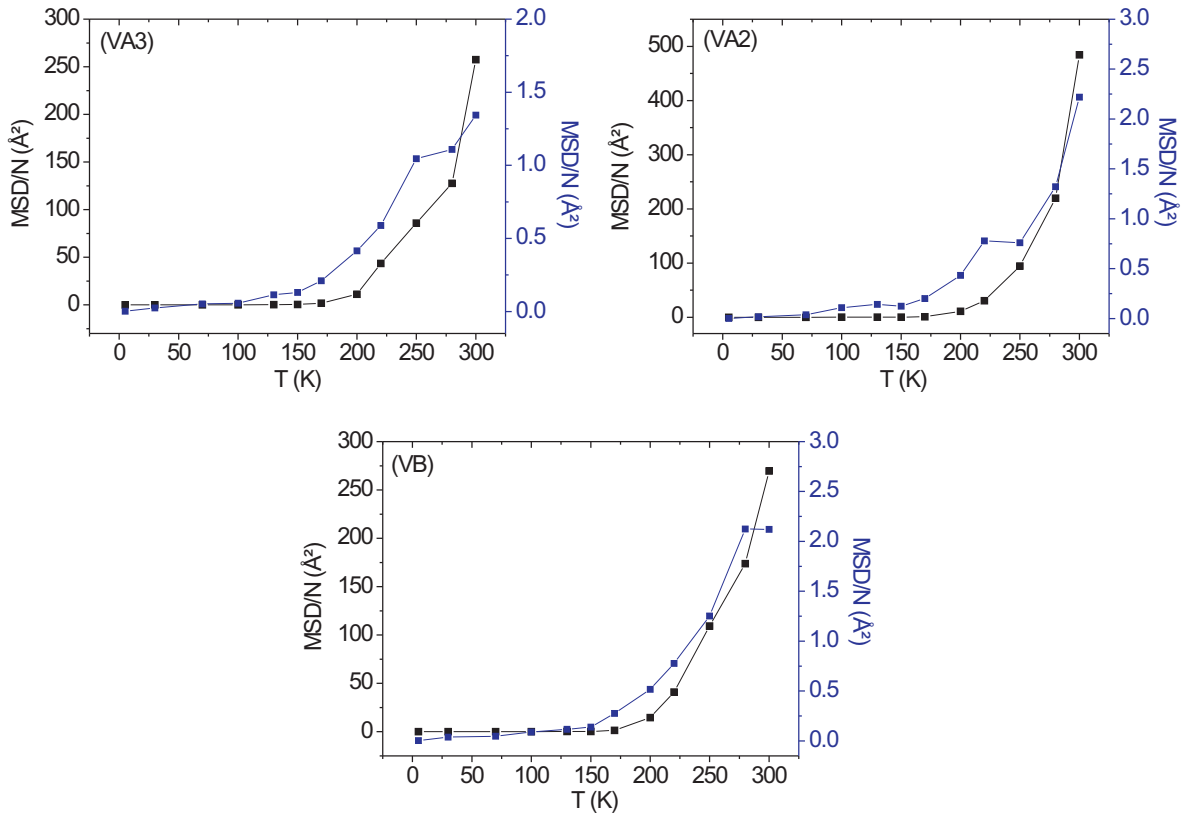


FIG. 4.1 – Evolution de l'écart quadratique moyen des ions (courbe noire) comparée à celui des viscotoxines hydratées (courbe bleue), sur la moyenne des 5 trajectoires de 1ns. N correspond au nombre d'ions.

Sur la Fig. 4.2, nous comparons les comportements moyens des ions. Nous pouvons constater que les évolutions de l'écart quadratique moyen de ces ions sont très semblables d'une viscotoxine à une autre. L'augmentation du $\frac{MSD}{N}$ des ions de VA2 est essentiellement due aux mouvements décrits ci-dessus. Les mouvements des ions étant couplés à ceux des sites très mobiles des protéines (hydrophiles), ceux-ci engendrent une forte mobilité de ces ions.

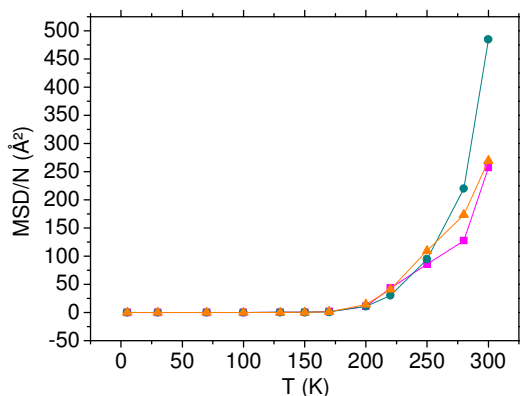


FIG. 4.2 – Evolution de la moyenne des écarts quadratiques moyens des ions mis en solution avec VA3 (carrés roses), VA2 (ronds bleu-verts) et VB (triangles oranges).

Sur la Fig. 4.3, nous présentons l'évolution de l'écart quadratique moyen des ions, simulés au contact de VA3, en fonction du temps de production. Nous remarquons que pour les trajectoires de 50ps (courbe noire), 100ps (courbe rouge) et 200ps (courbe bleue), les valeurs du $\frac{MSD}{N}$ sont beaucoup plus faibles que sur une trajectoire de 1ns (courbe bleu-vert). Ceci implique que l'activation des mouvements des ions démarre sur des temps de simulation plus longs que 200ps. En effet, à 300K, le $\frac{MSD}{N}$ vaut environ 30Å^2 sur 200ps, alors qu'il est de 300Å^2 sur 1ns.

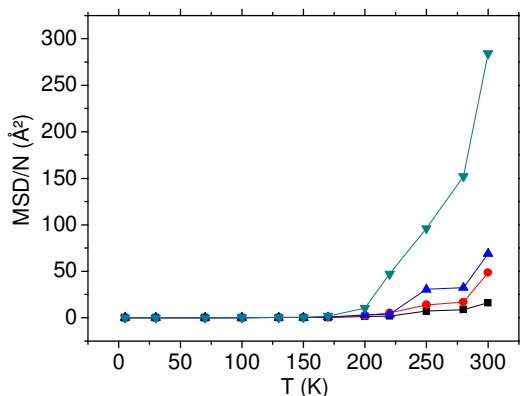


FIG. 4.3 – Dépendance en temps de simulations de l'écart quadratique moyen des ions simulés avec la viscotoxine A3, la courbe noire représente une trajectoire de 50ps, la courbe rouge de 100ps, la courbe bleue de 200ps et la courbe verte de 1ns.

Une mesure expérimentale des ions peut être envisagée pour déterminer la dynamique des protéines puisque nous avons pu voir que leurs mouvements sont couplés aux mouvements des sites hydrophiles. Des techniques de RMN avec l'ajout d'ions PO_4^{3-} ou de marquage par l'iode pourraient être une possibilité pour la détection des mouvements de ces ions.

TROISIÈME PARTIE

Dynamique et cytotoxicité

Données sur les interactions membrane/protéine

La toxicité des thionines est caractérisée par leur capacité à interagir avec une membrane. L'orientation des structures secondaires par rapport aux membranes est déterminante dans le mécanisme d'insertion de ces protéines dans les membranes [Coulon *et al.* 03]. Dans le cas des viscotoxines, certaines propriétés intrinsèques modifient leur capacité à interagir avec la membrane. Par exemple, la surface hydrophobe de VA3 est supérieure à celle de VA2, ceci peut expliquer pourquoi il faudra une concentration plus grande de VA2 pour aboutir à une rupture des membranes.

Il existe deux régions distinctes dans les thionines qui jouent un rôle fondamental : la région hydrophobe dans les hélices et la région hydrophile entre les hélices et les feuillets qui leur procurent un caractère amphiphatique [Giudici *et al.* 03, Caaveiro *et al.* 98]. Les degrés de toxicité diffèrent d'une viscotoxine à une autre : VA3 est la plus toxique, puis VA2 et enfin VB la moins. Le comportement de VB est totalement différent de VA3, ceci peut s'expliquer par les mutations, qui interviennent majoritairement sur des résidus chargés, comme nous l'avons vu au chapitre 4, partie I.

La Fig. 1.1 représente la localisation de ces mutations entre VA3 et VB [Giudici *et al.* 03]. Nous pouvons constater qu'elles se situent majoritairement sur les deux hélices et la boucle les reliant, modifiant ainsi le caractère amphiphatique de VB. De plus, le résidu 25 semble jouer un rôle primordial dans l'insertion des protéines dans la membrane. Suivant l'orientation de ce résidu, l'insertion sera modifiée. En comparant les séquences avec le tableau 4.3 de la partie I, nous constatons que les viscotoxines diffèrent par une mutation sur ce résidu : la viscotoxine A3 a une thréonine en 25^{ième} position, VA2 une valine et VB une arginine. Ces différences modifient l'orientation probable des protéines [Coulon *et al.* 03]. Les différences de charge entre VA3 (+6) et VA2 et VB (+5) sont faibles et n'apportent probablement pas de différence majeure dans l'interaction avec la membrane.

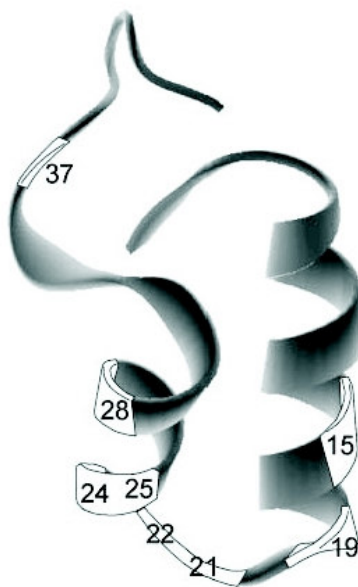


FIG. 1.1 – Localisation des mutations entre VA3 et VB [Giudici *et al.* 03].

1.1 Point de vue expérimental

Giudici étudia en 2003 l'insertion des viscotoxines A3 et B dans des membranes de phospholipides chargées négativement par des techniques infrarouges et de fluorescence. La combinaison des interactions électrostatiques des groupements polaires du phospholipide avec les interactions hydrophobes de la viscotoxine permet une attache et la perturbation de la membrane [Giudici *et al.* 03].

Plusieurs étapes sont nécessaires à l'interaction : la première consiste donc à cette interaction électrostatique avec les charges négatives du phospholipide, puis la stabilisation de la structure à la surface de la membrane se fait par l'orientation des hélices alpha parallèlement au plan de la membrane (opposant ainsi la région hydrophobe des hélices à la membrane) et enfin l'insertion de la protéine dans la membrane, partiellement ou entièrement. Le positionnement de la protéine dans une conformation spécifique dans la membrane cause une perturbation de la structure de la membrane, suivie d'un changement de perméabilité et d'une rupture de la membrane. Suivant l'orientation de la protéine, l'insertion et donc la rupture de la membrane peuvent être modifiées. C'est pourquoi des mutations impliquant un changement d'encombrement dans l'espace peuvent moduler la toxicité.

L'équipe de Coulon a réalisé un deuxième modèle de membrane avec phosphatidylcholine et phosphatidylserine et étudia l'interaction de VA3 avec ce modèle par fluorescence [Coulon *et al.* 01]. La Fig. 1.2 représente un modèle structural (plausible) pour l'orientation des viscotoxines lors de l'insertion. Cette représentation met en évidence le rôle du résidu 25 en contact avec la membrane. Le mode d'action est décrit de la manière suivante : l'insertion de VA3 est induite par des défauts dans la membrane. Quand la surface est saturée en VA3, il y a une déstabilisation complète de la perméabilité menant à la rupture de la membrane. Deux possibilités sont évoquées pour la rupture de la membrane : soit les thionines créent des canaux d'ions avec la membrane, soit les protéines induisent une déstabilisation par une interaction spécifique.

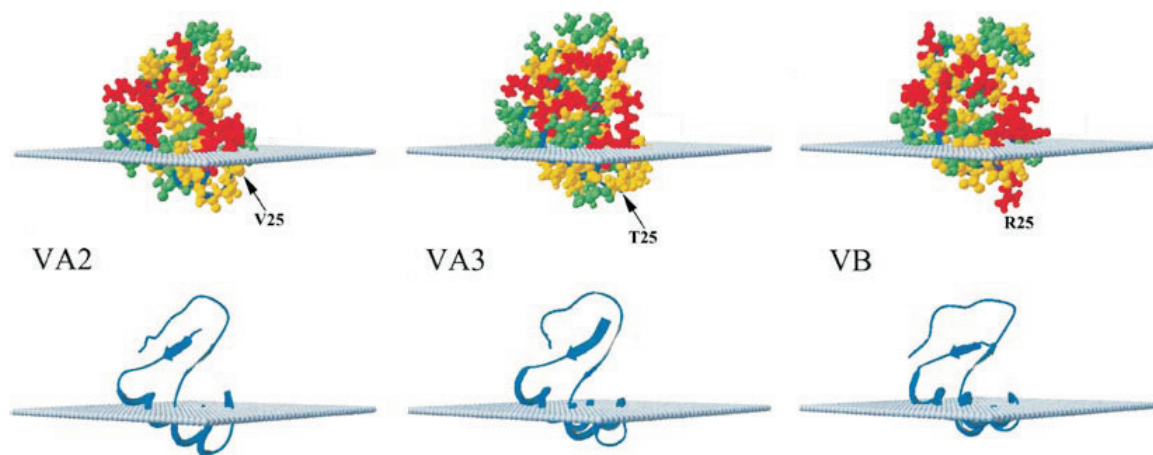


FIG. 1.2 – Modèle d'insertion probable des viscotoxines dans les membranes et localisation du résidu 25 lors de l'interaction [Coulon *et al.* 03].

Les canaux d'ions ont été étudiés par Hughes en 2000 [Hughes *et al.* 00]. En utilisant des mesures électro-physiologiques, il a établi que la toxicité des thionines est due à leur

capacité à générer des canaux d'ions dans les membranes ayant pour résultat la dissipation des gradients de concentration en ions, essentiels pour l'entretien des cellules. Ces canaux d'ions provoquent des formations de pores, conduisant à la rupture de la membrane.

Le résidu Tyr13 joue un rôle primordial dans la toxicité puisque son iodination mène à une perte de toxicité. Le passage des ions dans les canaux est peut-être relié à ce résidu (situé au centre d'un domaine hydrophile). Il faut noter que ce résidu Tyr13 est absent dans la crambine, ceci pourrait expliquer sa perte de toxicité. De plus, la neutralité de la crambine doit apporter d'importantes modifications dans cette interaction.

1.2 Point de vue théorique

Shai fit un résumé des modes d'action possibles pour la perméation de la membrane en 1999 [Shai 99]. Les thionines, solubles dans l'eau, atteignent la membrane et changent de conformations ce qui leurs permet de s'associer à la membrane. La perméation de la membrane par des peptides amphiphatiques avec des hélices alpha peut se produire par l'intermédiaire de l'un de ces 2 mécanismes : soit un mécanisme '*barrel-stave*' (c'est à dire par la formation de pores), soit par un mécanisme '*carpet*' (la destruction de la membrane se fait par un 'tapis' de thionines). Dans le mécanisme *barrel-stave*, il y aura insertion de peptides alors que dans le mécanisme *carpet* la perméation est due à une accumulation de protéines à sa surface, mais sans insertion.

Le mécanisme *barrel-stave* consiste donc en la formation de canaux ou pores à travers la membrane avec les hélices alpha amphiphatiques. Les surfaces hydrophobes interagissent avec le cœur lipide de la membrane et les surfaces hydrophiles produisent les pores. Ce modèle est représenté à droite sur la Fig. 1.3. Quatre critères doivent être respectés :

- 1) les thionines se lient à la membrane grâce à leur structure d'hélices alpha, (ceci constitue une étape cruciale avec des interactions principalement hydrophobes),
- 2) les thionines s'identifient et se positionnent autour de la surface,
- 3) les hélices s'insèrent dans le cœur de la membrane hydrophobe,
- 4) il y a accumulation de thionines pour augmenter la taille des pores.

Le mécanisme *carpet* repose sur des peptides lytiques en contact des groupes phospholipides tout au long du processus de perméation de la membrane. Les peptides se lient à la surface de la membrane cible et la recouvre comme un tapis. Contrairement au '*barrel-stave*', il n'y a pas d'insertion. Ce modèle est représenté à gauche sur la Fig. 1.3. 5 étapes sont nécessaires :

- 1) interactions entre peptides et phospholipides,
- 2) alignement des peptides à la surface de la membrane avec la surface hydrophile face aux phospholipides,
- 3) réorientation des résidus hydrophobes vers le cœur hydrophobe de la membrane,
- 4) dissolution (en augmentant la solubilité) de petits agrégats de lipides recouverts de peptides,
- 5) désintégration de la membrane en perturbant sa courbure.

La perméation de la membrane requiert un changement dans sa courbure. Il existe alors 2 classes de peptides suivant cette courbure :

- ceux qui induisent une courbure négative (classe des L-peptides) : les faces polaires des peptides sont associées aux groupes polaires des phospholipides,
- ceux qui induisent une courbure positive (classe des peptides A) : la courbure est positive dans une dimension perpendiculaire au plan de la membrane.

Il est possible de prédire quel type de mécanisme sera utilisé. Pour le *carpet*, la perméation de la membrane n'est possible que s'il y a une grande concentration de peptides liés à la membrane. Si le peptide induit la perméation de la membrane avant qu'elle ne soit recouverte de peptides, ce n'est pas le mécanisme *carpet* qui est utilisé mais le mécanisme *barrel-stave*.

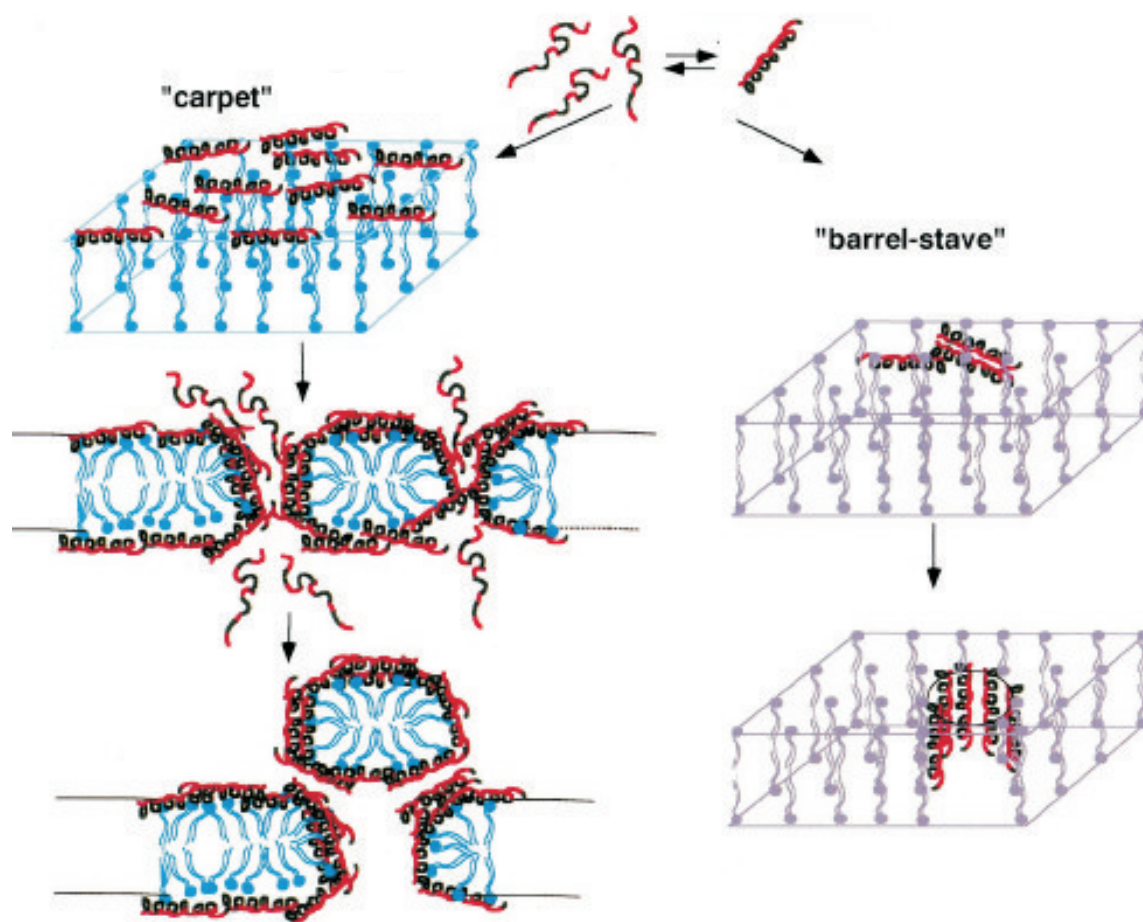


FIG. 1.3 – Illustration des mécanismes d'insertion des thionines dans une membrane, le modèle *carpet* à gauche et le modèle *barrel-stave* à droite [Shai 99].

Dynamique de protéines avec tyrosines iodinisées

Nous avons vu au chapitre 4, partie I, que l'iodination des tyrosines 13 des viscotoxines conduit à une perte de toxicité. Dans ce chapitre, nous présentons l'étude de la transition dynamique de ces viscotoxines modifiées chimiquement afin de constater des différences de dynamique engendrées par cette iodination.

2.1 Courbes moyennes générales

Nous avons tout d'abord analysé la dynamique globale des protéines. Sur la Fig. 2.1, nous comparons les évolutions de l'écart quadratique moyen entre les viscotoxines iodinisées (courbes noires) et les viscotoxines étudiées au chapitre 3, partie II (courbes rouges).

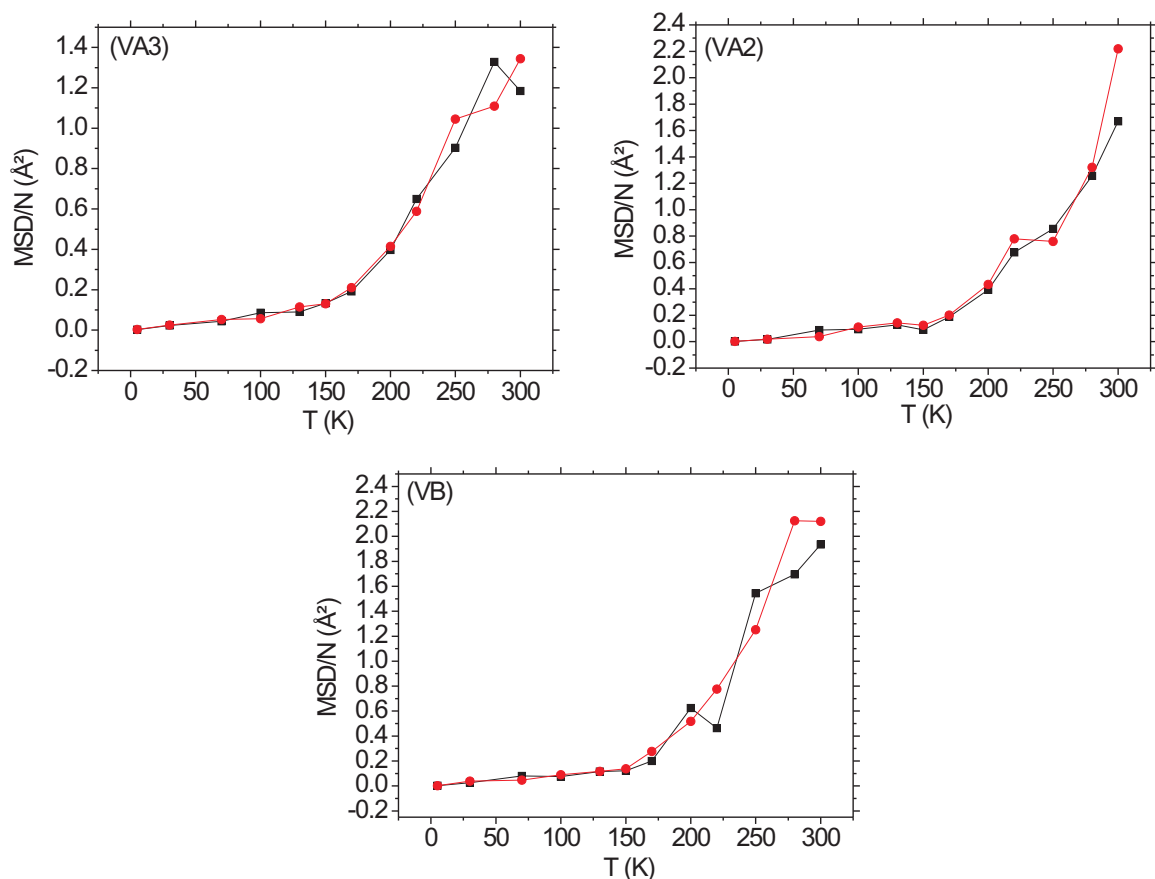


FIG. 2.1 – Comparaisons des écarts quadratiques moyens des viscotoxines (courbes rouges) et des viscotoxines iodinisées (courbes noires), simulées dans une boîte d'eau sur des trajectoires de 1 ns.

Nous remarquons que l'iodination de ces viscotoxines n'apporte pas de différence dans la dynamique globale de ces protéines. La flexibilité globale semble être la même, mais comment se répartit cette mobilité le long de la séquence ? Quelles sont les conséquences sur la mobilité

des atomes dues à la mutation d'un atome d'hydrogène par un atome d'iode sur la tyrosine 13 ?

2.2 Mobilité des atomes

D'après les courbes de la Fig. 2.2, nous constatons que la distribution de la mobilité le long de la séquence est modifiée avec la tyrosine iodinisée. Cependant, en étudiant en détails la mobilité des atomes, nous remarquons que les sites mobiles restent les mêmes dans le cas de VA3. Leur MSD varie mais les atomes les plus mobiles sont les mêmes. Par exemple, le CH_3 du résidu LEU18 (atomes n°268, 269, 270 et 271) a un MSD dans VA3 d'environ 8Å^2 , et sa valeur chute à 5Å^2 pour la VA3 iodinisée, mais ce groupement méthyle fait partie des atomes les plus mobiles des deux protéines. L'étude détaillée de VA2 montre une flexibilité totalement différente. L'iodination modifie toute la distribution de la mobilité. Les atomes des résidus LYS1 et ARG23 (principalement responsables de la flexibilité de VA2 dans la partie II) n'ont plus un MSD élevé. Par exemple, le groupe NH_2 du résidu LYS1 (atomes n°19, 20, 21 et 22) a un MSD de 22Å^2 pour VA2 et seulement 6Å^2 lors de l'iodination. Dans le cas de VB, nous pouvons constater que la viscotoxine iodinisée a beaucoup moins de pics de mobilité, surtout au milieu de la séquence. Cependant, les atomes les plus mobiles de cette viscotoxine iodinisée le sont également pour VB.

Même si la dynamique globale montrait peu de changements, l'iodination a provoqué plus de modifications dans la mobilité des atomes pour VA2, moins pour VB et enfin très peu pour VA3.

2.3 Etude détaillée de la tyrosine 13

Sur la Fig. 2.3, nous avons représenté la mobilité des atomes de la tyrosine 13 des viscotoxines à gauche, comparée à la mobilité de la tyrosine iodinisée. Pour la viscotoxine A3, les mêmes atomes sont mobiles : il s'agit des hydrogènes du cycle, en particulier celui rattaché à l'oxygène. Avoir remplacé un hydrogène par un atome d'iode (atome n°19) n'apporte que très peu de différences. Pour la viscotoxine A2, les différences engendrées par l'iodination sont beaucoup plus importantes. Tout d'abord, l'ordre de grandeur du $\frac{MSD_{\max}}{N}$ passe de 1.2Å^2 pour VA2 à seulement 0.4Å^2 pour la viscotoxine iodinisée. Malgré cette différence dans l'ordre de grandeur, les atomes considérés comme mobiles restent les mêmes et il s'agit également des hydrogènes du cycle. Pour la viscotoxine B, l'atome d'iode a une très grande mobilité (environ 0.8Å^2) mais la distribution de la mobilité reste la même.

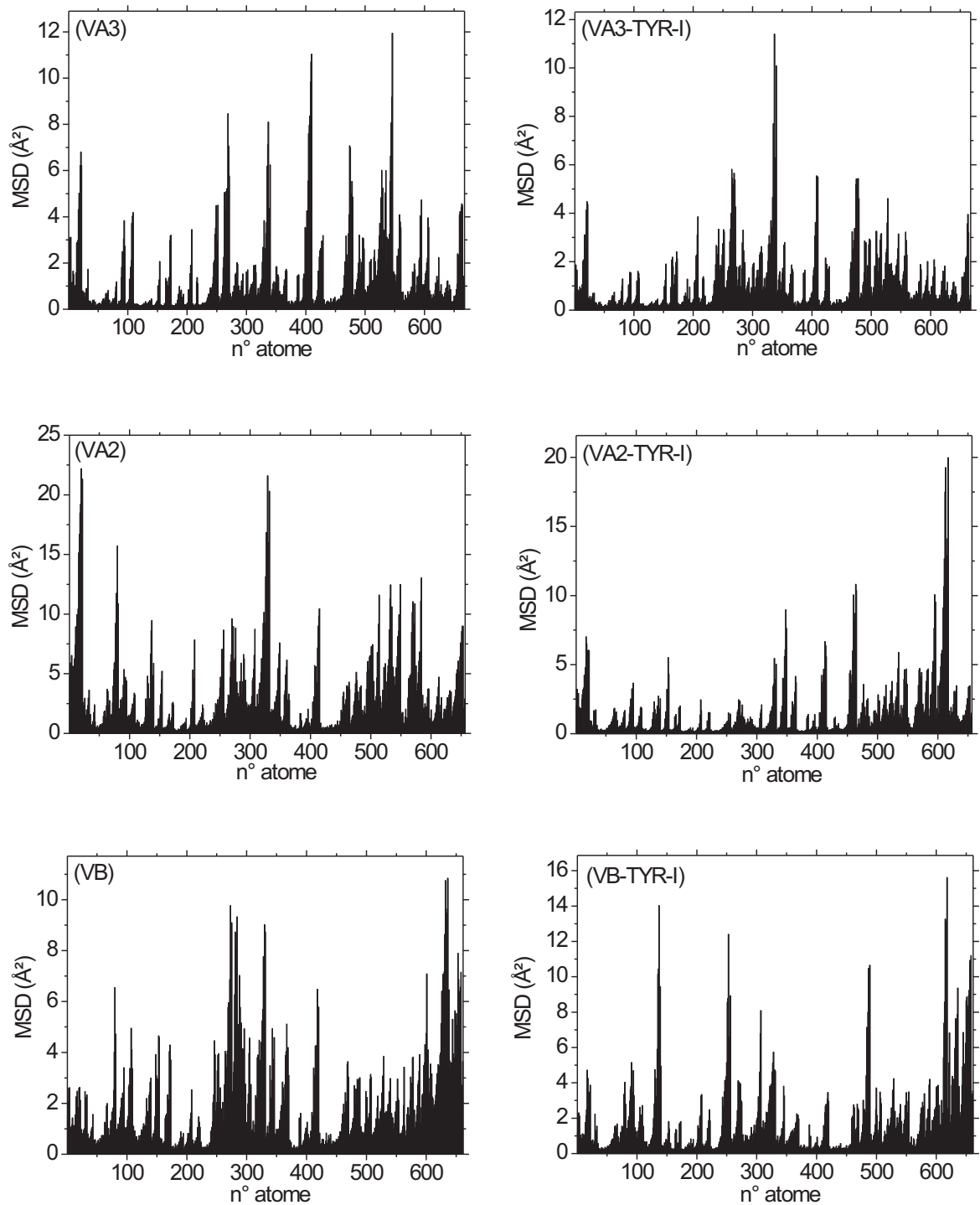


FIG. 2.2 – Comparaisons de la mobilité des atomes entre les viscotoxines à gauche, et les viscotoxines iodinisées à droite à 300K, sur une trajectoire de 1ns.

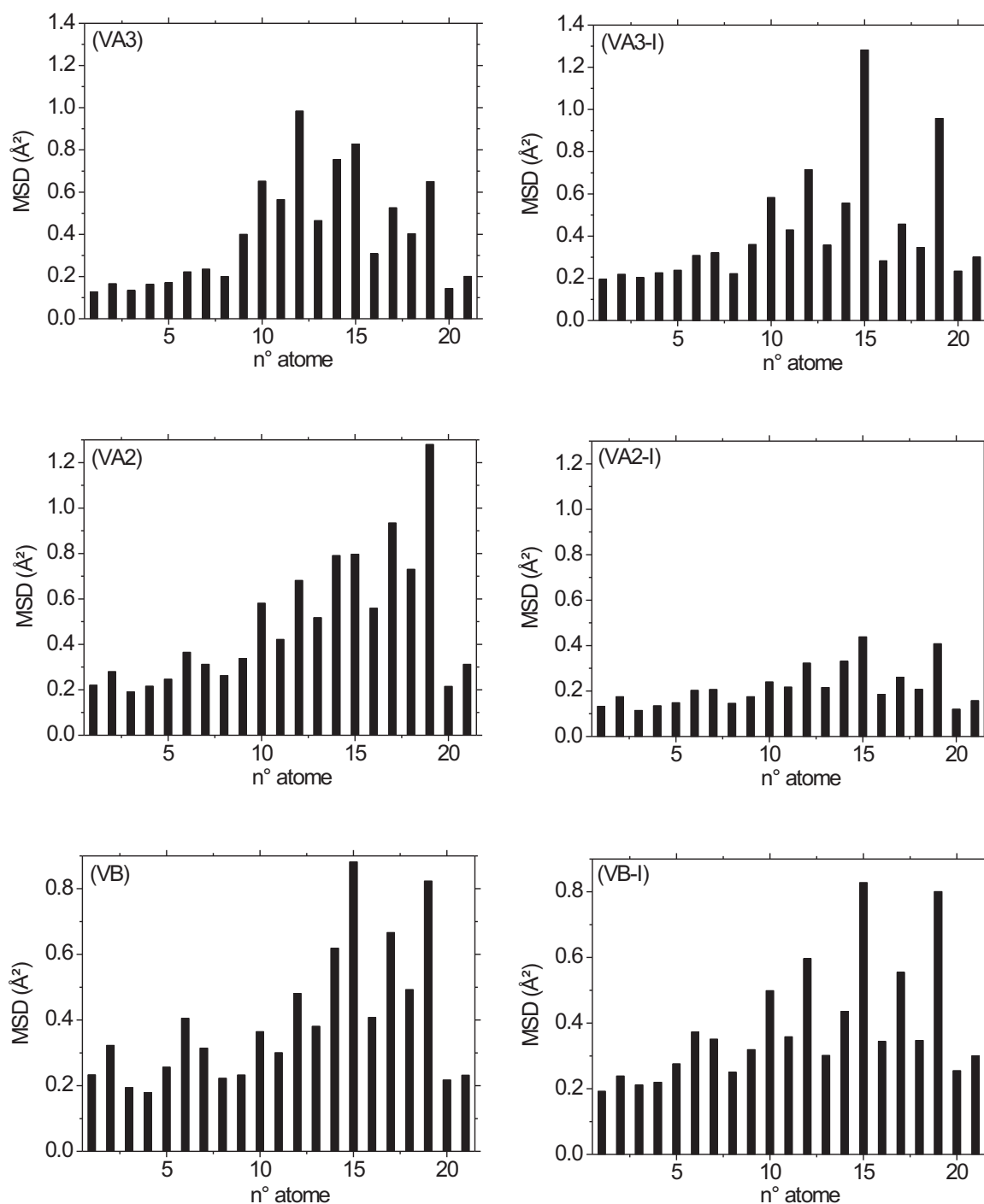


FIG. 2.3 – Mobilité des atomes de la tyrosine 13 des viscotoxines à gauche et des viscotoxines iodinisées à droite, simulées dans une boîte d'eau à 300K et sur une trajectoire de 1ns.

2.4 Mobilité des résidus

Dans le tableau suivant, nous répertorions les résidus considérés comme mobiles des viscotoxines et des viscotoxines iodinisées. Le même critère de sélection a été utilisé (résidu considéré comme mobile si $\frac{MSD}{T} > \frac{MSD_{max}}{2T}$). Nous constatons que dans l'étude de la mobilité des résidus, VA3 a seulement une différence pour LYS46, mais dans le cas de la viscotoxine iodinisée, ce résidu a une valeur très proche du critère de sélection. **L'iodination ne provoque donc aucune différence pour VA3.** Pour VA2, seuls deux résidus se retrouvent dans les deux cas : LYS1 et LYS46, l'iodination provoque beaucoup plus de modifications dans la distribution de la mobilité. Pour VB, seul le résidu LYS46 est considéré comme mobile dans les deux.

En étudiant la localisation dans l'espace des acides aminés avec la figure 4.3, nous pouvons remarquer que tous les résidus mobiles de VB-I sont proches dans l'espace de la tyrosine iodinisée. La même remarque est faite pour VA2-I. Nous avons vu précédemment que la mobilité de la tyrosine n° 13 était surtout modifiée pour VA2 et VB, c'est pourquoi la mobilité des résidus proches dans l'espace est également modifiée. En effet, la mobilité des résidus dépend de la mobilité de son environnement et donc des résidus voisins.

	Résidus considérés comme mobiles						
VA3	LYS1	ARG17	LEU18	ARG23	LYS28	LYS33	LYS46
$\frac{MSD}{T} (\text{Å}^2/K)$	0.240	0.229	0.193	0.235	0.287	0.196	0.244
VA3-I	LYS1	ARG17	LEU18	ARG23	LYS28	LYS33	
$\frac{MSD}{T} (\text{Å}^2/K)$	0.274	0.198	0.525	0.323	0.194	0.217	
VA2	LYS1	PHE18	ARG23	PRO41	LYS46		
$\frac{MSD}{T} (\text{Å}^2/K)$	0.805	0.308	0.692	0.315	0.347		
VA2-I	LYS1	ARG10	ARG17	TYR44	LYS46		
$\frac{MSD}{T} (\text{Å}^2/K)$	0.274	0.089	0.290	0.557	0.163		
VB	ARG23	ARG25	PRO45	LYS46			
$\frac{MSD}{T} (\text{Å}^2/K)$	0.362	0.305	0.359	0.704			
VB-I	LYS1	ARG17	PHE18	LYS33	TYR44		
$\frac{MSD}{T} (\text{Å}^2/K)$	0.321	0.293	0.144	0.189	0.357		

TAB. 2.1 – Comparaison des résidus considérés comme mobiles pour les viscotoxines iodinisées. La valeur du MSD/T correspond au MSD du résidu à $T=300K$ divisé par $300K$.

2.5 Conclusion

Grâce à cette étude des viscotoxines iodinisées, nous pouvons conclure que le remplacement d'un atome d'hydrogène par un atome d'iode n'affecte pas la dynamique de la protéine, du moins dans les échelles de temps que nous avons étudiées et pour des protéines isolées en solution. Certaines différences apparaissent dans la distribution de la mobilité le long de la séquence majoritairement pour VA2, moins notables pour VB et très peu pour VA3. Tous

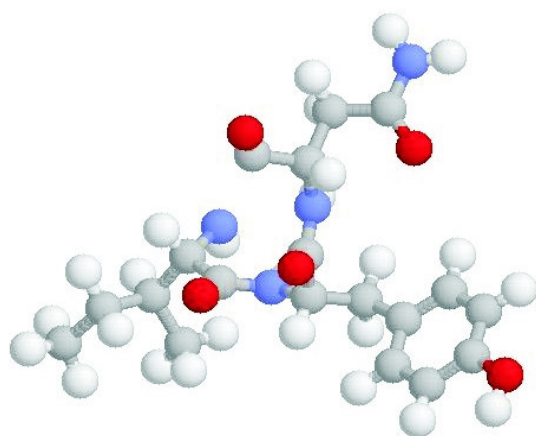
les autres calculs sur le rayon de giration, la dispersion des fluctuations, des liaisons hydrogène ou encore de la mobilité des contre-ions ont également été réalisés pour ces protéines modifiées mais aucune différence n'apparaît avec l'iodination, c'est pourquoi ces résultats, n'apportant aucun autre élément, ne sont pas présentés ici. La perte de toxicité ne semble pas due à un changement de dynamique ; VA3 est la plus toxique et c'est la protéine dont les mouvements internes sont les moins affectés par l'iodination dans nos simulations.

Une autre hypothèse pouvant expliquer la perte de toxicité liée à l'iodination de la TYR13 est que l'encombrement stérique associé à l'iode pourrait bloquer un canal artificiel créé par l'insertion des viscotoxines dans les membranes. Cette hypothèse reste à étudier. D'autres hypothèses sur les conséquences sur la toxicité apportées par l'iodination sont possibles. Une dernière hypothèse repose sur l'interaction possible entre les ions et la tyrosine. En effet, dans le cristal de VA3 (code PDB : 1OKH), un ion PO_4^{3-} est proche de cette tyrosine. Modifier chimiquement la tyrosine 13 change probablement l'interaction avec ces ions.

Des calculs des propriétés électroniques et des potentiels électrostatiques des protéines iodinisées et natives pourraient nous donner des informations sur ces interactions.

Nous avons entrepris des travaux très préliminaires lors d'un récent séjour à l'Université d'Anvers (groupe du Prof. Van Alsenoy) afin d'évaluer les modifications des propriétés électroniques induites par l'iodination. Nous présentons ici nos premiers résultats obtenus pour la modification de la densité électronique induite par l'iodination dans un modèle d'agrégat. Nous avons calculé les charges atomiques de deux modèles, cf Fig. 2.4. Le premier modèle, à gauche, est un tripeptide, équivalent à l'enchaînement des résidus ILE12, TYR13 et ASN14 de la viscotoxine A3. Leur position dans l'espace par rapport à celle qu'ils occupent dans la protéine entière a été reproduite. Le second modèle est un enchaînement de 7 acides aminés. Nous avons ajouté aux trois précédents résidus 4 autres acides aminés qui sont très proches de la TYR13 dans l'espace, à savoir les résidus ARG23, PRO24, CYS26 et ALA27. De la même façon, les positions des résidus les uns par rapport aux autres sont les mêmes que dans la protéine entière. De plus, pour chacun de ces modèles, les calculs ont été également réalisés avec la tyrosine iodinisée.

Modèle n°1



Modèle n°2

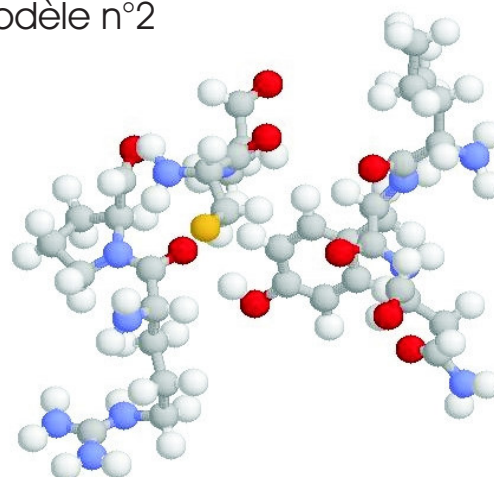


FIG. 2.4 – Modèles utilisés lors des calculs de la densité électronique.

Nos calculs ont été réalisés en utilisant la théorie de la fonctionnelle de la densité (DFT, Density Functional Theory). L'objectif principal de la théorie DFT est d'évaluer l'énergie fondamentale des molécules et solides en utilisant comme fonction à optimiser, non pas la fonction d'onde électronique mais la densité des électrons. Alors que la fonction d'onde dépend de $3N$ variables spatiales (où N est le nombre total de particules du système), la densité est seulement fonction de trois variables x, y, z . La DFT permet de simuler les propriétés des molécules et solides en incluant les effets de la corrélation électronique et des interactions d'échange avec un temps de calcul raisonnable [Parr et Weitao 89]. Nous avons utilisé le logiciel GAUSSIAN [Frisch *et al.*] avec la méthode DFT et les fonctionnelles d'échange et de corrélation B3LYP [Becke 93]. Les orbitales moléculaires sont représentées par des combinaisons linéaires d'orbitales gaussiennes en utilisant la base 6-311G**. La densité électronique générée a été utilisée pour calculer les charges des atomes en utilisant la méthode de partition de stockholder [Rousseau *et al.* 00].

Sur la Fig. 2.5, nous avons représenté les charges atomiques obtenues pour les deux modèles (le tripeptide à gauche et l'enchaînement de 7 acides aminés à droite) en noir. Les courbes rouges représentent les calculs des charges avec l'iodination de la TYR13. Dans le premier modèle (à gauche de la Fig. 2.5), seuls 2 atomes ont une charge différente : il s'agit d'un atome d'hydrogène et de carbone du résidu TYR13 (atomes $n^{\circ}35$ et 38), très proches de l'atome d'iode (atome $n^{\circ}39$). En effet, la charge de l'hydrogène est $0.164C$ et $0.129C$ avec l'iodination. Pour le carbone, sa charge est $-0.064C$ et $-0.032C$ avec la tyrosine iodinisée. Dans le second modèle (à droite de la Fig. 2.5), nous pouvons noter les différences de charges les plus importantes pour 3 atomes. Tout d'abord, le même atome de carbone ($n^{\circ}38$) en contact direct avec l'atome d'iode voit sa charge modifiée dans ce modèle également ($-0.057C$ à $-0.032C$ avec l'iodination). La charge de l'atome d'oxygène ($n^{\circ}55$) du résidu ASN14 est $-0.227C$ et $-0.208C$ avec l'iodination et enfin, la charge de l'atome d'oxygène du résidu CYS26 ($n^{\circ}107$) est $-0.307C$ et $-0.326C$ avec la tyrosine iodinisée. De manière générale, peu de différences sont engendrées par l'iodination de la tyrosine. Une étude approfondie de la modification de la réactivité chimique pourrait être entreprise.

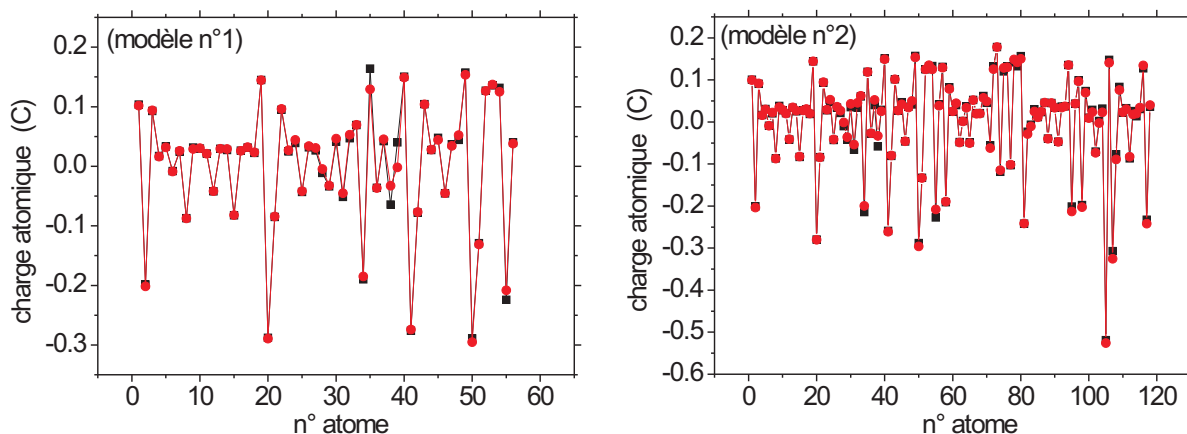


FIG. 2.5 – Variation des charges atomiques pour le modèle $n^{\circ}1$ à gauche et le modèle $n^{\circ}2$ à droite, en noir. Les calculs avec l'iodination de la TYR13 sont représentés en rouge.

Conclusion

Les protéines se comportent comme des systèmes flexibles dont leurs mouvements dépendent de l'environnement et des changements de températures. Il existe une relation entre la structure, la dynamique et la fonction biologique des protéines. En effet, modifier l'un de ces trois critères revient à modifier les deux autres.

La dynamique de petites protéines appartenant à la famille des thionines a été étudiée (4 crambines et 3 viscotoxines). Notre choix s'est porté sur cette famille de protéines car elles possèdent une haute homologie de séquence. Leur intérêt biologique réside dans leur toxicité vis-à-vis des membranes cellulaires. Bien que ces 7 protéines soient homologues, elles ont une toxicité très différente.

Notre objectif a d'abord été d'élucider les relations structure-dynamique de ces thionines en étudiant leur transition dynamique dans la seconde partie de notre thèse. La transition dynamique marque le passage entre un régime harmonique, appelé état R, pour $T < 200K$, et un régime anharmonique, appelé état F, pour $T > 200K$. C'est seulement dans ce régime anharmonique que la fonction biologique s'exprime [Ringe et Petsko 03, Frauenfelder et McMahan 98]. Nous avons simulé l'évolution de l'écart quadratique moyen des fluctuations structurales des protéines ($\frac{MSD}{N}$) et la répartition de ces fluctuations le long de la séquence primaire en fonction de la température, typiquement entre 50 et 300K. Nous avons étudié en détails l'origine des mouvements anharmoniques, les relations séquences-dynamique et indirectement leur relation avec la toxicité. Les relations dynamique-toxicité ont été également explorées par l'étude de l'iodination du résidu TYR13 dans la dernière partie de notre thèse.

A partir de l'ensemble de nos résultats, nous avons pu ébaucher un modèle de la transition dynamique. Les fluctuations structurales des protéines peuvent être décomposées en deux contributions : la première est un terme vibrationnel (oscillations autour d'un minimum local de la surface d'énergie libre), qui augmente linéairement avec la température et la seconde est une constante diffusive (sauts entre minima locaux), avec une évolution non-linéaire. Ces contributions peuvent être séparées et quantifiées par des constantes effectives et/ou des modèles simplifiés de la transition dynamique.

L'étude des fluctuations des angles diédraux du backbone (angles γ) (chapitre 3, partie II) met également clairement en évidence ces deux types de comportement dynamique à l'échelle locale du résidu. On distingue des résidus avec une dynamique "harmonique" et d'autres avec une dynamique "anharmonique". Les premiers fluctuent (du moins pour les angles γ) dans un puits d'énergie libre harmonique, son puits d'énergie libre associé à l'angle diédral du résidu s'élargit avec la température tout en conservant un seul minimum.

Les seconds fluctuent autour d'un seul minimum d'énergie libre à basses températures mais existent dans plusieurs états conformationnels à hautes températures. Les sauts entre minima sont responsables de la transition dynamique des atomes du backbone des protéines. L'anharmonicité de ces résidus consiste donc à des passages entre ces sous-états. Nous avons pu conclure que ces chemins sont reliés à l'eau puisque les puits des protéines déshydratées ont tous un seul minimum.

Les propriétés du solvant jouent un rôle primordial dans cette transition dynamique. En effet, nous avons pu noter que la transition dynamique a lieu à la température de diffusion de l'eau. De plus, le comportement de l'eau interfaciale (différent de l'eau de volume), mis en évidence dans la discussion de la section 3.3.6 de la partie II, apporte des modifications dans la dynamique des protéines. L'eau permet à la protéine des changements de conformations, équivalents à des sauts entre les différents minima des puits d'énergie. Ces minima semblent correspondre aux basculements des chaînes latérales.

Il est difficile de prédire quels types de résidus présentera une forte anharmonicité. Nos calculs montrent que les acides aminés situés en surface de la protéine seront entraînés plus facilement par l'agitation du solvant, par rapport à ceux situés au coeur de la protéine. Plus leur chaîne latérale est grande, plus leur mobilité est forte. Enfin, la charge des résidus joue probablement un rôle important puisque dans le cas des viscotoxines hydratées, la plupart des résidus considérés comme mobiles sont des résidus chargés. Ceci corrobore les observations expérimentales faites sur la parvalbumine [Zanotti *et al.* 99].

Dans le vide, les résidus hydrophobes sont le plus souvent les plus mobiles. Cependant, l'étude des puits d'énergie au chapitre 3 de la partie II montre que les backbones des thionines déshydratées ont un seul minimum d'énergie. Nous avons en effet montré que le backbone des protéines hydratées ne subit aucune transition dynamique ; le MSD évolue linéairement avec la température. Mais les chaînes latérales existent dans différents rotamers. Nous avons pu constater au chapitre 2 de la partie II que les méthyles explorent leurs 3 minima locaux au-delà d'une certaine température de transition qui varie d'un groupement à un autre mais qui se situe au-delà de 100–130K. L'élévation de la température ne modifie probablement pas les hauteurs des barrières d'énergie libre mais la probabilité de sauter cette barrière augmente d'après la loi d'Arrhenius. Ces sauts activent les mouvements des méthyles. Nous avons montré à partir de l'étude de la distribution des barrières d'énergie le long des structures, que la mobilité des résidus était directement corrélée à la hauteur de la barrière d'énergie potentielle entre les minima. Cette mobilité existe toujours dans le cas des protéines hydratées mais leur contribution est masquée par l'augmentation de la mobilité du backbone et des chaînes latérales. Tous les types d'atomes sont susceptibles de bouger dans l'eau. L'étude des protéines dans le vide met ainsi en évidence le rôle des méthyles, les atomes lourds n'ont qu'un comportement harmonique quelque soit la température.

Une seule mutation peut changer le comportement dynamique des protéines. Cette étude a été réalisée sur les crambines, qui ne diffèrent que par une seule mutation.

Comme nous l'avons vu dans le chapitre 2 de la deuxième partie, non seulement la flexibilité globale des crambines est modifiée, mais aussi la répartition de la mobilité le long de sa séquence, alors que la superposition de ces protéines dans le chapitre 5 de la partie I nous indique que les structures secondaires sont conservées. Cependant les angles diédraux mettent en évidence une modification de structure de l'hélice 2 lors du remplacement de PRO22 par SER22. Ceci n'était pas visible dans la superposition des structures.

Ceci s'explique par le fait que la mobilité d'un acide aminé dépend de son environnement dans l'espace. Ainsi des résidus, non liés par des liaisons covalentes, mais proches dans l'espace du résidu muté verront leur mobilité modifiée.

L'analyse des rayons de giration nous a permis de constater que les modifications de dynamique sont liées au type de résidu muté. Les deux crambines avec une proline en 22^{ème} position ont une évolution du rayon de giration semblable et une superposition avec beaucoup plus de similitudes dans la boucle. Il en est de même pour les deux crambines avec une serine sur le résidu 22. Ces deux résidus diffèrent par leur surface d'hydratation mais aussi par leur structure : la proline est constituée d'un cycle spécial qui se referme sur la chaîne latérale, rendant la protéine très peu flexible.

En appliquant un algorithme d'analyse des séries temporelles (DFA, [Peng *et al.* 94]) aux trajectoires des angles diédraux γ , nous avons démontré que le MSD des angles diédraux suit une loi en puissance (Eq. 2.33, partie I), caractérisée par un exposant α . Nous avons pu voir dans l'étude des angles diédraux du backbone (γ) que les exposants sont modifiés sur toute la séquence par une seule mutation et que l'allure des profils d'énergie libre (PMF) peut également changer. Cependant, les parties mobiles restent les mêmes d'une protéine à une autre : elles se situent majoritairement dans la boucle et les feuillets beta.

De notre étude de la transition dynamique d'une famille de protéines, plusieurs conclusions peuvent être tirées :

- L'étude de la dynamique d'un système dépend fortement du temps de simulation et donc de la résolution en énergie des spectromètres neutrons utilisés pour sonder la dynamique. Plus la fenêtre temporelle d'observation ou de simulation est longue, plus la protéine a de chances d'explorer différents sous-états conformationnels et plus son MSD (facteur de Debye-Waller) sera élevé,

- L'augmentation de la flexibilité d'une protéine (mesurée par le $\frac{MSD}{T}$) s'accompagne d'une augmentation des zones mobiles, la dispersion spatiale des fluctuations structurales le long de la structure primaire suit la même évolution avec la température que l'écart quadratique moyen,

- L'étude de la dynamique d'une protéine montre trois types d'évolution possibles avec l'augmentation de la température : l'acide aminé peut être harmonique et voir sa mobilité augmenter linéairement avec la température, il peut également présenter des « sauts » dans sa mobilité en fonction de la température, ou alors une rupture de pente bien prononcée à la température de transition dynamique,

- L'analyse de la mobilité des ions au chapitre 4 de la partie II montre qu'ils subissent eux aussi une transition dynamique. Ces contre-ions peuvent être utilisés comme sonde de la transition dynamique des protéines, c'est pourquoi la détermination d'un procédé expérimental sur la dynamique des ions permettrait d'étudier la dynamique de la protéine,

- La variation du taux d'hydratation, présentée au chapitre 3, partie II, nous a permis de constater qu'à partir d'un taux critique, où la protéine est complètement hydratée, ce taux n'apporte plus de différence dans la transition dynamique des thionines si on l'augmente. Au contraire, en-dessous de ce taux, la dynamique sera complètement différente.

Enfin, dans la dernière partie, nous nous sommes intéressés à la relation dynamique-cytotoxicité. Les viscotoxines perdent toute leur fonction biologique lorsque l'on remplace un atome d'hydrogène de la tyrosine 13 par un atome d'iode. Au chapitre 2 de la partie III, nous avons alors simulé l'iodination de la tyrosine 13. Aucune différence dans la dynamique

globale n'intervient même si des modifications de la distribution le long de la séquence sont assez importantes pour VA2, plus faibles pour VB et quasiment inexistantes pour VA3. Ces différences n'apportent que très peu de conséquences dans la dynamique. La perte de la toxicité des viscotoxines n'est donc pas due à un changement de dynamique ; VA3 étant la protéine la plus toxique et celle pour laquelle l'iodination apporte le moins de modifications de dynamique.

Notre travail pourrait être poursuivi au moins dans deux directions. D'une part, une étude structurale complémentaire de l'ensemble de nos données devrait permettre de mettre en évidence à l'échelle microscopique comment le réseau de liaisons hydrogène labiles de l'eau avec les résidus (en particulier les résidus chargés) permet de créer des sous-états conformationnels pour les atomes du backbone au-delà de $200K$. D'autre part, disposant d'un modèle structural et dynamique de ces protéines, l'étude de leurs interactions avec un modèle de membrane cellulaire pourrait être envisagée. Dans cette perspective, une étude des potentiels électrostatiques et de la possible réactivité des protéines iodinisées et natives avec des modèles de membranes pourrait être simulée en utilisant à la fois les outils de la dynamique moléculaire et des calculs quantiques.

Protocoles dans le vide

A.1 Paramètres et programme utilisés pour l'étape de chauffage

A.1.1 Paramètres

Pour définir les valeurs des paramètres correspondants à l'évaluation des interactions électrostatiques du champ de force, nous nous sommes basés sur les calculs suivants : sachant que le volume de la crambine *P22/L25* vaut 9281.17\AA^3 et en supposant que la protéine peut être représentée par une sphère, son rayon vaut alors 13.037\AA . Nous avons ainsi considéré qu'un rayon de 15\AA était suffisant pour dénombrer toutes les interactions électrostatiques. Il s'agira du paramètre *CUTNB*.

D'après [Brooks *et al.* 83], les autres paramètres sont définis de la façon suivante :

$$CTOFNB = CUTNB - 0.5 = 15 - 0.5 = 14.5\text{\AA} \quad (\text{A.1})$$

$$CTONNB = CTOFNB - 1.0 = 14.5 - 1.0 = 13.5\text{\AA} \quad (\text{A.2})$$

Paramètres	Définitions
ISEED	1938, correspond aux vitesses/conditions initiales
NSTEP	2000, nombre de pas
TIMESTEP	0.001, NSTEP*TIMESTEP=temps en picoseconde
NPRINT	50, fréquence à laquelle est imprimée l'énergie dans le fichier de sortie
IPRFRQ	100, fréquence à laquelle sont calculées les moyennes et fluctuations des énergies
FIRSTT	0, température de départ
FINALT	température de chauffe (variable suivant la température choisie)
TEMINC	incrément de la température (dépend de T_f), T augmente de TEMINC tous les IHTFRQ
IASORS	1, option pour générer ou corriger les vitesses tous les IHTFRQ durant la période de chauffe
IASVEL	1, assigne des vitesses (si > 0 : distribution gaussienne, < 0 : uniforme)
INBFRQ	10, vérifie les interactions non-liées
IHTFRQ	100, nombre de pas où est incrémentée la température de TEMINC
NTRFRQ	100, fréquence pour laquelle les rotations et translations des molécules sont stoppées durant la simulation
CUTNB	15, rayon au-delà duquel on ne prend plus en compte les interactions électrostatiques
CTONNB	14, rayon au-delà duquel on ne prend plus en compte les interactions électrostatiques, deuxième valeur nécessaire pour la fonction 'SWITCH' qui diminue les changements brusques dans le décompte des interactions avec le temps
CTOFNB	14.5, rayon au-delà duquel on ne prend plus en compte les énergies électrostatiques
EPSilon	1, constante diélectrique (valeur par défaut)

TAB. A.1 – Paramètres utilisés dans l'étape de chauffage des différentes structures.

A.1.2 Programme Charmm

```
* CHARMM input file for system heating
* We are heating (0 to 300 K) the crambin 1CCM from NMR average structure
* after minimization
*
! Read in Topology and Parameter files
open unit 1 card read name top_all22_prot.inp
read RTF card unit 1
close unit 1
open unit 1 card read name par_all22_prot.inp
read PARA card unit 1
close unit 1
! Read in PSF and Coordinate file
open read formatted unit 27 name 1ccm_nmr.psf
read psf card unit 27
open read formatted unit 27 name 1ccm_nmr_min_cg.pdb
read coor pdb unit 27
! Specifications for NONBond interactions
NBOND CUTNb 15.0 CTONnb 14.0 CTOFnb 14.5 -
SWITCH VSWItch CDIElectric EPSilon 1.0
! open the restart file
OPEN UNIT 12 WRITe FORMatted NAME 1ccm_heat-nose-300k-min.res
! heat the system from 0 K to 300K
DYNA NOSE QREF 50.0 TREF 300.0 NCYC 5 -
  ISEED 1938 NSTEp 2000 TIMEstep 0.001 -
  NPRInt 50 IPRFrq 100 -
  FIRSTT 0.0 FINALT 300.0 TEMINC 20.0 -
  IASORS 1 IASVEL 1 -
  INBFrq 10 IHTFrq 100 NTRFrq 100 -
  IUNWrit 12
OPEN UNIT 1 WRITe FORMatted NAME 1ccm_heat-nose-300k-min.pdb
WRITe COORDinate PDB UNIT 1
* COORDinates after heating (300K) of 1CCM
*
STOP
```

A.2 Paramètres et programme utilisés pour l'étape d'équilibrage

A.2.1 Paramètres

Paramètres	Définitions
NSTEP	600000
TIMESTEP	0.001
NPRINT	50
IPRFRQ	200
FINALT	température voulue
IASORS	0
IASVEL	1
ISCVEL	0, une valeur unique pour tous les atomes pour changer les vitesses, pour >0 : facteur proportionnel à l'énergie cinétique des atomes
INBFRQ	10
ICHECW	1, vérifie si la température moyenne se situe dans les limites autorisées (si 0 : ne vérifie pas)
TWINDL	-2.0, déviation inférieure tolérée par rapport à FINALT
TWINDH	2.0, déviation supérieure tolérée par rapport à FINALT
IEQFRQ	30, fréquence à laquelle les vitesses sont multipliées
CUTNB	15
CTONNB	14
CTOFNB	14.5
EPSilon	1.0

TAB. A.2 – Paramètres utilisés dans l'étape d'équilibrage des structures chauffées.

A.2.2 Programme Charmm

```

* CHARMM input file for system equilibration 1CCM
* after heating from 0 to 300K
*
! Read in Topology and Parameter files
open unit 1 card read name top_all22_prot.inp
read RTF card unit 1
close unit 1
open unit 1 card read name par_all22_prot.inp
read PARA card unit 1
close unit 1
! Read in PSF and Coordinate file
open read formatted unit 27 name 1ccm_nmr.psf
read psf card unit 27
open read formatted unit 27 name 1ccm_heat-nose-300k-min.pdb

```

```
read coor pdb unit 27
! Specifications for NONBond interactions
NBOND CUTNb 15.0 CTONnb 14.0 CTOFnb 14.5 -
SWITCH VSWItch CDIElectric EPSilon 1.0
! open restart files
OPEN UNIT 11 READ FORMatted NAME 1ccm_heat-nose-300k-min.res
OPEN UNIT 12 WRITe FORMatted NAME 1ccm_equil_nose_300k-min-600ps.res
! equilibrate the molecule for 600 ps at 300 K
DYNAmics NOSE QREF 50.0 TREF 300.0 NCYC 5 REStart -
  NSTEp 600000 TIMEstep 0.001 -
  NPRInt 50 IPRFrq 200 -
  FINALT 300.0 -
  IASORS 0 IASVEL 1 ISCVEL 0 -
  ICHECW 1 TWINDL -2.0 TWINDH 2.0 -
  INBFrq 10 IEQFrq 30 -
  IUNRea 11 IUNWrit 12
OPEN UNIT 1 WRITe FORMatted NAME 1ccm_equil_nose_300k-min-600ps.pdb
WRITe COORdinate PDB UNIT 1
* Equilibration of 1ccm all hydrogens at 300K (600ps)
*
OPEN UNIT 2 WRITe FORMatted NAME 1ccm_equil_nose_300k-min-600ps.crd
WRITe COORdinate cards UNIT 2
* Equilibration of 1ccm all hydrogens at 300K (600ps)
*
STOP
```

A.3 Paramètres et programme utilisés pour l'étape de production

A.3.1 Paramètres

Paramètres	Définitions
NSTEP	dépend de la trajectoire voulue
TIMESTEP	0.001
NPRINT	100
IPRFRQ	200
INBFRQ	20
IUNCrđ	14, sauvegarde des coordonnées (si -1, pas de sauvegarde)
NSAVcrđ	100, sauvegarde des frames
CUTNB	15
CTONNB	14
CTOFNB	14.5
EPSilon	1.0

TABLE A.3 – Paramètres utilisés dans l'étape de production des structures équilibrées à la température T .

A.3.2 Programme Charmm

```

* Production run. Molecular dynamics for crambin
*
! Read in Topology and Parameter files
open unit 1 card read name top_all22_prot.inp
read RTF card unit 1
close unit 1
open unit 1 card read name par_all22_prot.inp
read PARA card unit 1
close unit 1
! Read in PSF and Coordinate file
open read formatted unit 27 name 1ccm_nmr.psf
read psf card unit 27
open read formatted unit 27 name 1ccm_equil_nose_300k-min-600ps.pdb
read coor pdb unit 27
! Specifications for NONBond interactions
NBOND CUTNb 15.0 CTONnb 14.0 CTOFnb 14.5 -
SWITCH VSWItch CDIElectric EPSilon 1.0
! open restart files
OPEN UNIT 11 READ FORMatted NAME 1ccm_equil_nose_300k-min-600ps.res
OPEN UNIT 12 WRITe FORMatted NAME 1ccm_run_300k_nose_1ns.res
! open a trajectory file
OPEN UNIT 14 WRITe UNFORMatted NAME 1ccm_run_300k_nose_1ns.dcd

```

```
open unit 15 write card name 1ccm_run_300k_nose_1ns.crd
write coor cards unit 15
! Production step 20 ps at 70K
DYNAMics NOSE QREF 50.0 TREF 300.0 NCYC 5 REStart -
  NSTEp 1000000 TIMEstep 0.001 -
  NPRInt 100 IPRFrq 200 -
  INBFrq 20 -
  IUNCrD 14 NSAVcrd 100 -
  IUNRea 11 IUNWrit 12
OPEN UNIT 1 WRITe FORMatted NAME 1ccm_run_300k_nose_1ns.pdb
WRITe COORdinate PDB UNIT 12
* COORdinate crambin after production, 300K
*
STOP
```

A.4 Programme Charmm pour le calcul du MSD des protéines dans le vide

```
* RMSD for proteins
*
if @?task eq 0 goto deadformat
if @?pdbfile eq 0 goto iamdead
SET PDBFILE = 1CCM
! get topology and force field for this run
!
!prnlev 0 node 0
wrnlev 0
open unit 1 read form name top_all22_prot.inp
read rtf card unit 1
close unit 1
open unit 1 read form name par_all22_prot.inp
read param card unit 1
close unit 1
prnlev 5 node 0
wrnlev 2
! get psf for the system
!
open unit 1 read form name @pdbfile.psf
read psf card unit 1
close unit 1
label atomrms
! Atom Average rms
set avgnam @pdbfile_no_wat.avgcrd
! Trajectory file of your system without water
!
open unit 30 uniform read name 1ccm_run_300k_nose_1ns.dcd
```

```

open unit 31 form write name @avgnam
open unit 32 form write name @pdbfile_no_wat.avgdmb
coor dyna comp firstu 30 nunit 1
write coor comp card unit 31
write coor comp dumb unit 32
close unit 30
close unit 31
close unit 32
goto extractrms
label doneextrms
stop
label format
! FORMAT
!
calc nunit = @end - @begin +1
open unit 30 uniform read name @pdbfile_e@begin.crd
traj iread 30
traj read
close unit 30
set start ?start
set skip ?skip
set nframes ?nfile
calc nfile = ?nfile * @nunit
calc stop = @skip * @nframes * @nunit + @start - @skip
set unit 30
set current @begin
label dynopen
open unit @unit uniform read name @pdbfile_e@current.crd
incr unit
incr current
if @current le @end goto dynopen
open unit 20 write uniform name @pdbfile_e@begin-@end.crd
merge firstu 30 nunit @nunit skip @skip output 20 nfile @nfile
close unit 20
close unit 30
open unit 30 read uniform name @pdbfile_e@begin-@end.crd
open unit 20 write form name @pdbfile_e@begin-@end.fcrd
dyna format first 30 nunit 1 output 20 -
skip @skip begin @start stop @stop
stop
!
label unformat
! UNFORMAT
set unit 30
open unit @unit uniform write name @pdbfile_e@begin-@end.crd
! open unit @unit uniform write name a12
open unit 20 read form name @pdbfile_e@begin-@end.fcrd

```

```
dyna unformat input 20 output 30
stop
label unmerge
! unmerge
!
calc nunit = @end - @begin +1
open unit 30 uniform read name @pdbfile_e@begin-@end.crd
traj ired 30
traj read
close unit 30
set start ?start
set skip ?skip
calc nframes ?nfile
calc stop = @skip * @nframes + @start - @skip
calc nfile = ?nfile / @nunit
set unit 30
set current @begin
label mergopen
open unit @unit uniform write name @pdbfile_e@current.crd
incr unit
incr current
if @current le @end goto mergopen
open unit 20 uniform read name @pdbfile_e@begin-@end.crd
merge firstu 20 nunit 1 output 30 nfile @nfile -
skip @skip begin @start stop @stop
stop
!*END OF SCRIPT subroutines follow*
label iamdead
system "banner Must specify coordinate/psf filename"
stop
label deadformat
system "banner 'Must '"
system "banner 'specify '"
system "banner 'task='"
system "banner ' stripwat'"
system "banner ' format'"
system "banner ' unformat'"
stop
label extractrms
! fix rms file for plotting
!
system "/bin/rm -f rms_fnam rms.csh awk.file"
system "echo 'BEGIN {header = 1 }' > awk.file"
system "echo '{if ( header == 1 )}' >> awk.file"
system "echo 'if ($1 == \"*\") header = 1;' >> awk.file"
System "echo ' else header = 0; }' >> awk.file"
system "echo 'else print $1, \" \", $2, \" \", [\"$4\"], \" \", $10, \" \", $10*$10; }' >> awk.file"
```

```
system "echo @AVGNAM | awk '{print tolower($1)}' >rms_fnam "  
system "echo ' set fnam='cat rms_fnam'; set fnam=$fnam :r; /bin/rm -f $fnam.rms '  
>> rms.csh"  
system "echo 'awk -f awk.file $fnam.avgcrd > $fnam.rms ' >> rms.csh"  
system "csh rms.csh"  
system "/bin/rm -f rms_fnam rms.csh awk.file"  
goto doneextrms
```


Protocoles des protéines hydratées

B.1 Programme Charmm utilisé pour l'insertion de la protéine dans la boîte d'eau

```
* Generate lccm (simulated annealing structure) in a water box
*
!Input data
!1) box dimensions, xdim, ydim, zdim
!2) identifiers, id1, id2
!3) topology filename
!4) parameter filename
!5) lccm coordinate filename (the corresponding sequence)
!6) waterbox coordinate filename
bomlev -1
!set box dimension
set xdim 48.0
set ydim 34.0
set zdim 25.0
! identifiers
set id1 lccm
set id2 water_box
! 1 : topology
! 2 : parameter
! 3 : cutim
! 4 : cutnb
! 5 : ctonnb
! 6 : ctofnb
! 7 : electrostatic cutoff regime (shift or switch)
! 8 : electrostatic pairing method (atom or group)
! 9 : VDW pairing method (vatom or vgroup)
! VDW cutoff regime is vwsitch
set 1 top_all27_prot_na.rtf
set 2 par_all27_prot_na.prm
set 3 14.0! cutim
set 4 14.0! cutnb
set 5 10.0! ctonnb
set 6 12.0! ctofnb
set 7 shift
set 8 atom
set 9 vatom
! read topology and parameters
open unit 9 read form name @1
read rtf card unit 9
open unit 9 read form name @2
read para card unit 9
! read residue sequence from a pdb file
! the pdb file was obtained from a simulated annealing procedure
```

```
open unit 12 read card name lccm_run_298k.pdb
read sequ pdb unit 12
! now generate the PSF and also the IC table (setup keyword)
generate lccm setup warn
! add disulfide bonds using the patch command
patch disu lccm 3 lccm 40
patch disu lccm 4 lccm 32
patch disu lccm 16 lccm 26
!set variable for later use
set nstrand1 = ?nres
!set variable for later use
calc nstrand2 = ?nres - @nstrand1
autogenerate angle dihedral
! read coordinates from a crd file
open unit 20 read form name lccm_run_298k.crd
read coor card unit 20
! info on size of structure; use this information when
! creating waterbox
coor stat
coor orient sele segid lccm .and. .not. hydrogen show end
coor stat
! read pre-equilibrated water box
open unit 30 read form name wat_48_34_25.crd
read sequence coor unit 30
generate solv first none last none noangle nodihedral
rewind unit 30
read coor card unit 30 append
! delete waters overlapping with the solute
! we are deleting all water molecules around 2.8Å
delete atom sele .byres. (segid solv .and. (.not. hydrogen) .and. -
(((segid lccm) .and. .not. hydr) .around. 2.8)) end
coor stat sele .not. hydrogen end
coor stat sele resn tip3 end
coor stat sele type p* end
! set npho ?nrel
coor stat sele type n* end
coor stat sele type h* end
coor stat sele type o* end
coor stat sele type c* .and. .not. resn cl end
! define term to specify selection of all solute atoms
define solute sele segid lccm end
set dist 0.0
! write coordinates, PSF and IC tables
open unit 20 write form name @id1_@id2_@xdim-@ydim-@zdim.crd
write coor cards unit 20
* generated coordinates, all-hydrogen
* in a cube of water
```

```

* 2.8 Å atom delete
* @1
* @2
*
open unit 20 card write form name @id1_@id2_@xdim-@ydim-@zdim.pdb
write coor pdb unit 20
* generated coordinates, all-hydrogen
* in a cube of water
* 2.8 Å atom delete
* @1
* @2
*
open unit 20 write form name @id1_@id2_@xdim-@ydim-@zdim.psf
write psf cards unit 20
* PSF, all-hydrogen
* in a cube of water
* 2.8 Å atom delete
* @1
* @2
*
ic purge
open unit 20 write form name @id1_@id2_@xdim-@ydim-@zdim.ic
write ic cards unit 20
* IC table, all-hydrogen
* in a cube of water
* 2.8 Å atom delete
* @1
* @2
*
! check total charge
set charge ?cgtot
coor stat sele .not. hydrogen end
coor stat sele resn tip3 end
coor stat sele type p* end
coor stat sele type n* end
coor stat sele type h* end
coor stat sele type o* end
coor stat sele type c* .and. .not. resn cl end
stop

```

B.2 Programme Charmm pour l'insertion d'ions dans la boîte d'eau dans le cas des viscotoxines hydratées

L'insertion de la viscotoxine hydratée se fait avec le programme décrit précédemment. Puis, le programme suivant remplace aléatoirement une molécule d'eau autant de fois que nécessaire jusqu'à avoir neutralisé le système. Pour la viscotoxine A3 (chargée +6), 6 ions Cl^- ont été ajoutés, pour les viscotoxines A2 et B (chargées +5), 5 ions Cl^- ont été ajoutés.

```
* Add either cations or anions randomly
* Random number generator SEED
*
set MND 1! ion exclusion radius
set SEED 1111! RNG
set ANION cla! Add Cl ions randomly in WAT
set I 6! Add 6
set ANID CL! Segid
set PROT vta3! segid of protein to be solvated
! ICL is number of  $\text{Cl}^-$  to add
read sequ @ANION @I
generate @ANID noang nodihe first none last none
! Set all  $\text{Cl}^-$  doordinates to origin.
coor set xdir 0.0 ydir 0.0 zdir 0.0 sele segid @ANID end
! Begin placement loop
Random uniform iseed @SEED
set k 1
LABEL ILP
! Define atoms more than distance @MND away from prot or  $\text{Cl}^-$ 
define xcld sele segid @PROT .or. segid @ANID end
define prox sele xcld .around. @MND end
! Pick just water oxygens
define LIST sele type OH2 .and. .not. PROX end
! Pick a random integer baser on number of atoms in last selection
calc IN int( ?RAND * ?NSEL )
if IN .lt. 1 set in ?NSEL
! Define the specific water O atom for the replacement
define TARG sele LIST .subset. @IN end
! Obtain the water O coordinates
coor stat sele TARG end
! Set the anion coordinates to water O coordinates
coor set xdir ?XAVE ydir ?YAVE zdir ?ZAVE sele atom @ANID @K @ANION end
! Delete the water molecule; byres will include the H atoms
delete atom sele .byres. TARG end sort
! LOOP AND REPEAT
incr k by 1
if @k .le. @I goto ILP
```

B.3 Paramètres et programme utilisés pour l'étape de chauffage

B.3.1 Paramètres

Les valeurs de CUTNB, CTOFNB et CUTNB ne correspondent plus aux équations A.1 et A.2. En effet, en utilisant l'ensemble NPT et la sommation d'Ewald, ces valeurs ont été paramétrées afin d'augmenter au maximum la vitesse de simulation, nous avons donc utilisé les meilleurs paramètres.

Paramètres	Définitions
ISEED	314159
NSTEP	4000
TIMESTEP	0.001
NPRINT	50
IPRFRQ	50
FIRSTT	température de départ (variable suivant la température choisie)
FINALT	température de chauffe (variable suivant la température choisie)
TSTRUC	température d'équilibre voulue
IASORS	1
IMGFRQ	-1, fréquence des images
IASVEL	1
IEQFRQ	0, fréquence de multiplication des vitesses
INBFRQ	-1
NTRFRQ	50
NSAVC	500, fréquence de sauvegarde des données
ISVFRQ	5000, Si NSTEP > ISVFRQ, réécrit un fichier de démarrage (restart)
PCONST	simulation à P constante
PMASS	500, masse du piston (en amu)
PREF	1, pression de référence (en atmosphère)
TMASS	1000, masse du piston thermique ($kcal * ps^2$)
ECHECK	10000, limite maximale pour laquelle l'énergie totale peut changer
CUTNB	12
CTONNB	8
CTOFNB	10
CUTOFF	12
CUTIM	12, distance maximale des atomes inclus dans 'IMAGE'

TAB. B.1 – Paramètres utilisés dans l'étape de chauffage des thionines hydratées.

B.3.2 Programme Charmm

```
* Heat of 1ccm in a box of water with crystal dimensions (48.0 x 34.0 x 25.0)
* Heat between 70K and 100K
*
faster on
! identifiers
set id1 1ccm
set id2 water_box
! cell parameters
set a 48.00
set b 34.00
set c 25.00
! 1 : topology
! 2 : parameter
! 3 : cutim
! 4 : cutnb
! 5 : ctonnb
! 6 : ctofnb
! 7 : electrostatic cutoff regime (shift or switch)
! 8 : electrostatic pairing method (atom or group)
! 9 : VDW pairing method (vatom or vgroup)
! VDW cutoff regime is vwsitch
set 1 top_all27_prot_na.rtf
set 2 par_all27_prot_na.prm
set 3 12.0! cutim
set 4 12.0! cutnb
set 5 8.0! ctonnb
set 6 10.0! ctofnb
set 7 shift
set 8 atom
set 9 vatom
! read topology and parameters
open unit 9 read form name @1
read rtf card unit 9
open unit 9 read form name @2
read para card unit 9
open unit 20 read form name 1ccm_water_box_48.0-34.0-25.0.psf
read psf cards unit 20
open unit 20 read form name 1ccm_water_box_equil_600ps_70k.crd
read coor cards unit 20
crystal define orthorhombic @a @b @c 90. 90. 90.
crystal build noper 0 cutoff @3
!kappa should be optimized based on the value of
!kappa at which the 1st derivative of the energy
!with respect to kappa goes to 0.0
set kap 0.36
```

```

set ord 6
!fft dimensions should correspond to 1 A or less; must
!be power of 2, 3 or 5
update inbfrq -1 imgfrq -1 ihbfrq 0 -
ewald pmewald kappa @kap fftx 64 ffty 32 fftz 32 order @ord -
@7 @8 @9 vswitch cutimg @3 cutnb @4 ctofnb @6 ctonnb @5
!turn on image centering of solvent
image byres xcen 0.0 ycen 0.0 zcen 0.0 -
sele resn tip3 .or. resn sod .or. resn cl end
! harmonic constraints on all 1ccm heavy atoms
cons harm force 5.0 mass sele segid 1ccm end
shake bonh para
set g 0
! open files for restart, trajectory and energies
open unit 31 write form name @id1_@id2_harm_heat_100k.res
open unit 32 write unfo name @id1_@id2_harm_heat_100k.trj
! molecular dynamics, hot start
dynamics cpt leap start nstep 4000 timestep 0.001 iseed 314159 -
  first 70.0 finalt 100.0 tstruc 100.0 -
  inbfrq -1 imgfrq -1 iprfrq 50 nprint 50 -
  iasors 1 iasvel 1 ieqfrq 0 ntrfrq 50 -
  iunrea -1 iunwri 31 iuncrd 32 kunit 33 nsavc 500 isvfrq 5000 -
  pconstant pmass 500.0 pcou 5.0 pref 1.0 isot pgamma 0.0 -
  hoover tref 100.0 reft 100.0 tmass 1000.0 echeck 100000.
! write pdb and crd files
open unit 20 write form name @id1_@id2_100k.crd
write coor cards unit 20
* Heat to 100 K, 1ccm coor, all-hydrogen
* in a @a @b @c box of water
* @1
* @2
*
open unit 20 write form name @id1_@id2_100k.pdb
write coor pdb unit 20
* Heat to 100 K, 1ccm coor, all-hydrogen
* in a @a @b @c box of water
* @1
* @2
*
stop

```

B.4 Paramètres et programme utilisés pour l'étape d'équilibrage

B.4.1 Paramètres

Les paramètres de la simulation sont les mêmes que précédemment sauf pour les cas suivants :

Paramètres	Définitions
NSTEP	600000
TSTRUC	température de la structure
PCONST	simulation à P constante
HOOVER	simulation à T constante
TREF	température de la structure

TAB. B.2 – *Paramètres utilisés dans l'étape d'équilibration des thionines hydratées.*

B.4.2 Programme Charmm

* Equilibration of 1ccm in a box of water with crystal dimensions (48.0 x 34.0 x 25.0)

* Protein in the waterbox at 100K

*

faster on

! identifiers

set id1 1ccm

set id2 water_box

! cell parameters

set a 48.00

set b 34.00

set c 25.00

! 1 : topology

! 2 : parameter

! 3 : cutim

! 4 : cutnb

! 5 : ctonnb

! 6 : ctofnb

! 7 : electrostatic cutoff regime (shift or switch)

! 8 : electrostatic pairing method (atom or group)

! 9 : VDW pairing method (vatom or vgroup)

! VDW cutoff regime is vwsitch

set 1 top_all27_prot_na.rtf

set 2 par_all27_prot_na.prm

set 3 12.0! cutim

set 4 12.0! cutnb

set 5 8.0! ctonnb

set 6 10.0! ctofnb

```

set 7 shift
set 8 atom
set 9 vatom
! read topology and parameters
open unit 9 read form name @1
read rtf card unit 9
open unit 9 read form name @2
read para card unit 9
open unit 20 read form name 1cmm_water_box_48.0-34.0-25.0.psf
read psf cards unit 20
open unit 20 read form name 1ccm_water_box_100k.crd
read coor cards unit 20
crystal define orthorhombic @a @b @c 90. 90. 90.
crystal build noper 0 cutoff @3
!kappa should be optimized based on the value of
!kappa at which the 1st derivative of the energy
!with respect to kappa goes to 0.0
set kap 0.36
set ord 6
!fftx dimensions should correspond to 1 Å or less; must
!be power of 2, 3 or 5
update inbfrq -1 imgfrq -1 ihbfrq 0 -
ewald pmewald kappa @kap fftx 64 ffty 32 fftz 32 order @ord -
@7 @8 @9 vswitch cutimg @3 cutnb @4 ctofnb @6 ctonnb @5
!turn on image centering of solvent
image byres xcen 0.0 ycen 0.0 zcen 0.0 -
sele resn tip3 .or. resn sod .or. resn cl end
! harmonic constraints on all 1ccm heavy atoms
cons harm force 5.0 mass sele segid 1ccm end
shake bonh para
set g 0
! open files for restart, trajectory and energies
open unit 31 read form name @id1_@id2_harm_heat_100k.res
open unit 32 read unfo name @id1_@id2_harm_heat_100k.trj
open unit 31 write form name @id1_@id2_equil_600ps_100k.res
open unit 32 write unfo name @id1_@id2_equil_600ps_100k.trj
! equilibrate the system at 100K during 600ps
dynamics cpt leap start nstep 600000 timestep 0.001 iseed 314159 -
  first 100.0 finalt 100.0 tstruc 100.0 -
  inbfrq -1 imgfrq -1 iprfrq 50 nprint 50 -
  iasors 1 iasvel 1 ieqfrq 0 ntrfrq 50 -
  iunrea -1 iunwri 31 iuncrd 32 kunit 33 nsavc 500 isvfrq 5000 -
  pconstant pmass 500.0 pcou 5.0 pref 1.0 isot pgamma 0.0 -
  hoover tref 100.0 reft 100.0 tmass 1000.0 echeck 100000.
open unit 20 write form name @id1_@id2_equil_600ps_100k.crd
write coor cards unit 20
* Equilibrate at 100 K, 1ccm coor, all-hydrogen

```

```
* in a @a @b @c box of water
* @1
* @2
*
open unit 20 write form name @id1_@id2_equil_600ps_100k.pdb
write coor pdb unit 20
* Equilibrate at 100 K, 1ccm coor, all-hydrogen
* in a @a @b @c box of water
* @1
* @2
*
stop
```

B.5 Programme utilisé dans l'étape de production

Ce programme est utilisé une première fois pendant 100ps en éliminant les contraintes harmoniques sur la protéine. De cette façon, la protéine pourra se stabiliser dans les nouvelles conditions imposées (température, volume, etc) et les structures générées par cette étape seront le point de départ des trajectoires utilisées dans les calculs de MSD. Puis ce même programme sera utilisé dans la production des trajectoires.

```
* Production step of 1ccm in a box of water with crystal dimensions (48.0 x 34.0 x 25.0)
*
faster on
! identifiers
set id1 1ccm
set id2 water_box
! cell parameters
set a 48.00
set b 34.00
set c 25.00
! 1 : topology
! 2 : parameter
! 3 : cutim
! 4 : cutnb
! 5 : ctonnb
! 6 : ctofnb
! 7 : electrostatic cutoff regime (shift or switch)
! 8 : electrostatic pairing method (atom or group)
! 9 : VDW pairing method (vatom or vgroup)
! VDW cutoff regime is vwsitch
set 1 top_all27_prot_na.rtf
set 2 par_all27_prot_na.prm
set 3 12.0! cutim
set 4 12.0! cutnb
set 5 8.0! ctonnb
```

```

set 6 10.0! ctofnb
set 7 shift
set 8 atom
set 9 vatom
! read topology and parameters
open unit 9 read form name @1
read rtf card unit 9
open unit 9 read form name @2
read para card unit 9
open unit 20 read form name lccm_water_box_48.0-34.0-25.0.psf
read psf cards unit 20
open unit 20 read form name lccm_water_box_equil_600ps_100k.crd
read coor cards unit 20
crystal define orthorhombic @a @b @c 90. 90. 90.
crystal build noper 0 cutoff @3
!kappa should be optimized based on the value of
!kappa at which the 1st derivative of the energy
!with respect to kappa goes to 0.0
set kap 0.36
set ord 6
!fftx dimensions should correspond to 1 Å or less; must
!be power of 2, 3 or 5
update inbfrq -1 imgfrq -1 ihbfrq 0 -
ewald pmewald kappa @kap fftx 64 ffty 32 fftz 32 order @ord -
@7 @8 @9 vswitch cutimg @3 cutnb @4 ctofnb @6 ctonnb @5
!turn on image centering of solvent
image byres xcen 0.0 ycen 0.0 zcen 0.0 -
sele resn tip3 .or. resn sod .or. resn cl end
shake bonh para
set g 0
! open files for restart, trajectory and energies
open unit 30 read formatted name lccm_water_box_equil_600ps_100k.res
open unit 31 write form name @id1_@id2_produc_100k_100ps.res
open unit 32 write unfo name @id1_@id2_produc_100k_100ps.trj
dynamics cpt leap start nstep 100000 timestep 0.001 iseed 314159 -
  first 100.0 finalt 100.0 tstruc 100.0 -
  inbfrq -1 imgfrq -1 iprfrq 50 nprint 50 -
  iasors 1 iasvel 1 ieqfrq 0 ntrfrq 50 -
  iunrea -1 iunwri 31 iuncrd 32 kunit 33 nsavc 500 isvfrq 5000 -
  pconstant pmass 500.0 pcou 5.0 pref 1.0 isot pgamma 0.0 -
  hoover tref 100.0 reft 100.0 tmass 1000.0 echeck 100000.
open unit 20 write form name @id1_@id2_produc_100k_100ps.crd
write coor cards unit 20
* Poduction to 100K, lccm coor, all-hydrogen
* in a @a @b @c box of water
* @1
* @2

```

```
*
open unit 20 write form name @id1_@id2_produc_100k_100ps.pdb
write coor pdb unit 20
* Equilibration to 100K, 1ccm coor, all-hydrogen
* in a @a @b @c box of water
* @1
* @2
*
stop
```

B.6 Programme Charmm pour le calcul du MSD des protéines hydratées

B.6.1 Réorientation des trajectoires

```
* Remove overall protein translation/rotation. Make one trajectory with solvent
* recentered around protein, and one with solvent removed.
*
bomlev -1 ! all files may not exist yet
* Open&read rtf,param,psf,crd for 1ccm
* Open "raw" trajectory on unit 51, recentered/oriented trajectory on unit 61
*
read rtf card name top_all22_prot.inp
read para card name par_all22_prot.inp
read psf card name 1ccm_water_box_48.0-34.0-25.0.psf
read coor card name 1ccm_water_box_produc_300k_100ps.crd
define SOLUTE sele segid 1ccm end
set MXTIME 81
set NF 1
set SG 1ccm
set xtla 48.0
set xtlb 34.0
set xtlc 25.0
! en read form unit 50 name grx1_solv.frm
open read uniform unit 51 name 1ccm_water_box_produc_100k_1ns.trj
open read uniform unit 61 name 1ccm_water_box_produc_100k_1ns_rc.trj
! TO USE THIS TRAJECTORY : DELETE ATOM SELE .NOT. SOLUTE END
open read uniform unit 62 name 1ccm_water_box_produc_100k_1ns_orient.trj
crystal define orthorhombic @xtla @xtlb @xtlc 90.0 90.0 90.0
crystal build cutoff 20.0
image byres sele .not. SOLUTE end
! open write uniform unit 51 name 1ccm_solv.trj
! dynamics unformat inputu 50 outputu 51
open read uniform unit 51 name 1ccm_water_box_produc_300k_1ns.trj
open write uniform unit 88 name 1ccm_water_box_produc_300k_1ns_rc.trj
merge firstu 51 nunit @NF output 88 sele all end -
```

```

recenter orient sele segid 1ccm end
open write uniform unit 88 name 1ccm_water_box_produc_300k_1ns_orient.trj
merge firstu 51 nunit @NF output 88 sele SOLUTE end orient sele SOLUTE end

```

B.6.2 Calcul du MSD

Ce programme génère la structure moyenne calculée sur une trajectoire réorientée. Le fichier de coordonnées obtenu contient les valeurs du MSD de la protéine.

```

* Calculation of the RMS
*
open unit 10 read form name top_all22_prot.inp
read rtf card unit 10
close unit 10
open unit 10 read form name par_all22_prot.inp
read para card unit 10
close unit 10
faster on
open unit 20 read form name 1ccm_water_box_48.0-34.0-25.0.psf
read psf cards unit 20
close unit 20
open unit 20 read form name 1ccm_water_box_produc_300k_100ps.crd
read coor cards unit 20 comp resi
close unit 20
define SL1 select segid 1ccm end
open read uniform unit 22 name 1ccm_water_box_produc_300k_1ns.trj
open write uniform unit 23 name 1ccm_water_box_produc_300k_1ns_orient.trj
merge firstu 22 nunit 1 output 23 sele all end orient comp select SL1 end
close unit 23
close unit 22
open read uniform unit 22 name 1ccm_water_box_produc_300k_1ns_orient.trj
coor dyna firstu 22 nunit 1
open unit 31 form write name 1ccm_water_box_300k_1ns_orient_average.crd
write coor card unit 31
close unit 31
close unit 22

```


Simulation de recuit

C.1 Calibrage de la température maximale de chauffage

Dans le tableau suivant, nous répertorions les valeurs de l'énergie \bar{E} , du *RMS* et du ratio *R* nécessaires pour calibrer la température maximale de chauffe.

	300K	600K	700K	800K	900K
$\bar{E}(kcal/mol)$	-1454.671	-448.232	128.031	577.551	795.762
$RMS(\hat{A})$	0.654055	0.927663	1.059056	1.045632	1.046001
<i>R</i>	$2.656 \cdot 10^{-3}$	$8.138 \cdot 10^{-2}$	$7.655 \cdot 10^{-1}$	$2.176 \cdot 10^{-1}$	$1.896 \cdot 10^{-1}$

TAB. C.1 – Calibrage de la température de chauffage.

Chauffer le système jusqu'à 600K paraît être la meilleure solution. En effet avec une température supérieure, les critères de stabilité sur le RMS et R ne sont plus respectés et l'énergie devient positive.

C.2 Résultats

Dans les tableaux suivants, nous répertorions les valeurs des paramètres utilisés dans la simulation de recuit des protéines. Ces valeurs nous permettent de nous assurer de la stabilité des structures.

C.2.1 Recuit simulé des viscotoxines

Les résultats correspondent aux critères requis. Nous avons obtenu les structures équilibrées des viscotoxines qui seront utilisées dans la suite des simulations.

VA3	550K	500K	400K	298K	
$\bar{E}(kcal/mol)$	-669.793	-870.760	-1326.499	-1773.383	
$RMS(\hat{A})$	0.881341	0.877128	0.880664	0.894037	
<i>R</i>	$3.584 \cdot 10^{-2}$	$2.231 \cdot 10^{-2}$	$5.768 \cdot 10^{-3}$	$1.267 \cdot 10^{-3}$	
VA2	600K	550K	500K	400K	298K
$\bar{E}(kcal/mol)$	-850.765	-1132.903	-1334.913	-1770.664	-2193.595
$RMS(\hat{A})$	0.832925	0.830484	0.832012	0.828914	0.833524
<i>R</i>	$4.602 \cdot 10^{-2}$	$1.990 \cdot 10^{-2}$	$1.155 \cdot 10^{-2}$	$3.246 \cdot 10^{-3}$	$6.671 \cdot 10^{-4}$
VB	600K	500K	400K	298K	
$\bar{E}(kcal/mol)$	-1321.9013	-1741.1819	-2161.2717	-2583.6843	
$RMS(\hat{A})$	0.922538	0.922133	0.901834	0.902279	
<i>R</i>	$1.458 \cdot 10^{-2}$	$8.923 \cdot 10^{-3}$	$5.552 \cdot 10^{-3}$	$3.252 \cdot 10^{-3}$	

TAB. C.2 – Recuit simulé des viscotoxines.

C.2.2 Recuit simulé des crambines

Les critères fixés dans la simulation de recuit ont été respectés pour les crambines, ces structures équilibrées seront le point de départ de nos simulations.

P22/L25	600K	550K	500K	400K	298K
$\bar{E}(kcal/mol)$	-352.520	-625.243	-853.872	-1276.644	-1709.094
$RMS(\text{\AA})$	0.913199	0.894187	0.898257	0.901924	0.902745
R	$8.810 \cdot 10^{-2}$	$3.302 \cdot 10^{-2}$	$1.852 \cdot 10^{-2}$	$5.503 \cdot 10^{-3}$	$1.264 \cdot 10^{-3}$
S22/I25	600K	550K	500K	400K	298K
$\bar{E}(kcal/mol)$	-267.672	-474.657	-702.612	-1111.653	-1519.932
$RMS(\text{\AA})$	1.071141	1.081881	1.057030	1.070655	1.088989
R	$9.090 \cdot 10^{-2}$	$4.533 \cdot 10^{-2}$	$1.842 \cdot 10^{-2}$	$2.890 \cdot 10^{-3}$	$1.753 \cdot 10^{-3}$
P22/I25	600K	550K	500K	400K	298K
$\bar{E}(kcal/mol)$	-265.991	-506.580	-717.857	-1155.373	-1557.129
$RMS(\text{\AA})$	0.915430	0.912468	0.909145	0.915274	0.917983
R	$1.076 \cdot 10^{-1}$	$3.806 \cdot 10^{-2}$	$2.148 \cdot 10^{-2}$	$4.637 \cdot 10^{-3}$	$8.485 \cdot 10^{-4}$
S22/L25	600K	550K	500K	400K	298K
$\bar{E}(kcal/mol)$	-238.953	-467.492	-690.372	-1106.521	-1497.522
$RMS(\text{\AA})$	0.913630	0.952718	0.953338	0.969452	0.964449
R	$1.238 \cdot 10^{-1}$	$3.310 \cdot 10^{-2}$	$1.695 \cdot 10^{-2}$	$4.074 \cdot 10^{-3}$	$9.282 \cdot 10^{-4}$

TAB. C.3 – Recuits simulés des crambines.

Protéines dans le vide : Séparation des contributions en type d'atomes

Sur la Fig. D.1, nous représentons la variation de l'écart quadratique moyen des autres thionines pour toute la protéine (courbe noire), de tous les atomes d'hydrogène (courbe rouge) et de tous les atomes qui ne sont pas des hydrogènes (courbe bleue). Cette représentation nous permet de constater la forte contribution des hydrogènes à la flexibilité de la protéine dans le vide.

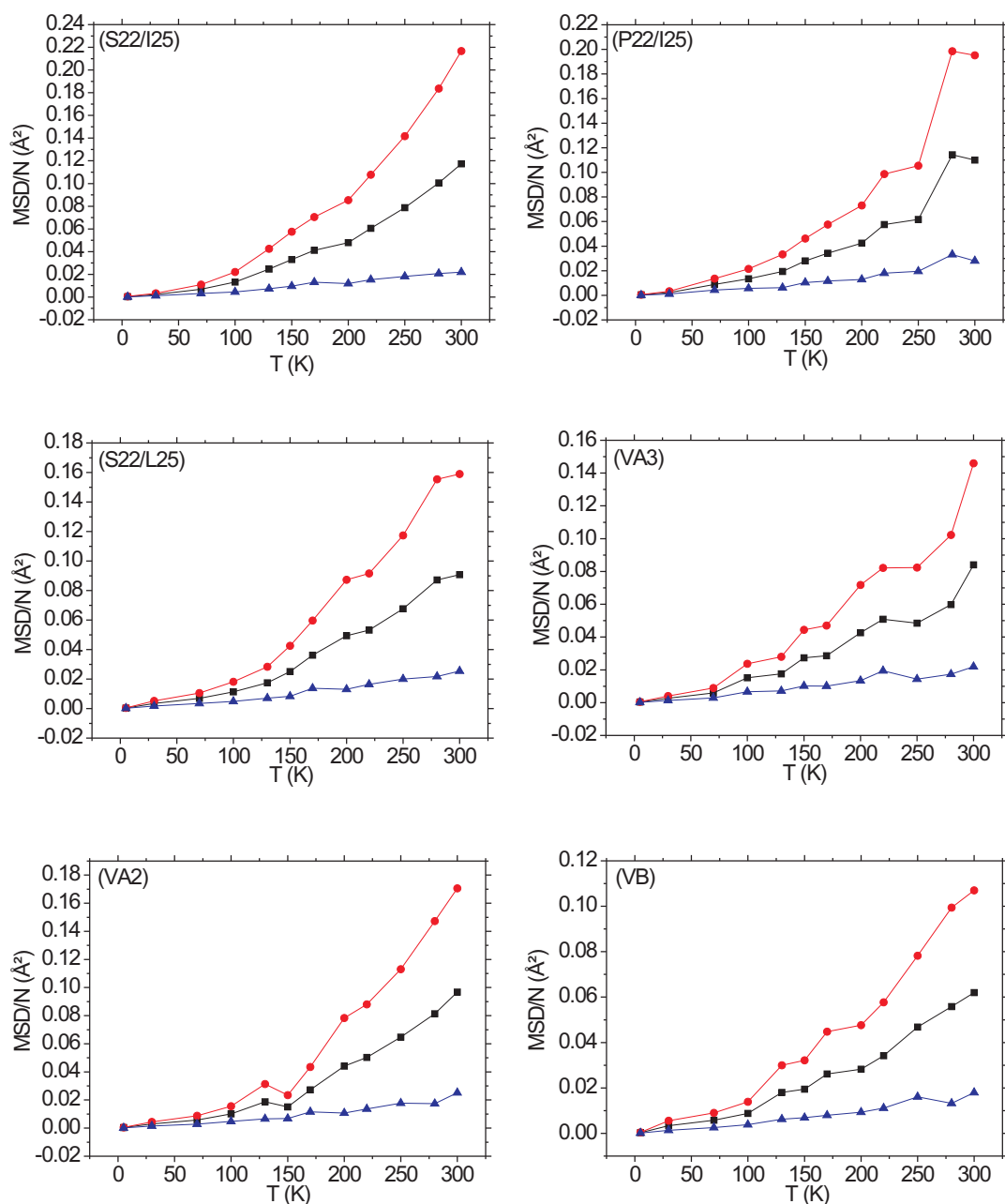


FIG. D.1 – Ecart quadratique moyen de la protéine globale (courbe noire) comparé aux différents types d'atomes : pour tous les hydrogènes (courbe rouge) et tous les atomes qui ne sont pas des hydrogènes (courbe bleue) dans le vide, sur une trajectoire de 1ns.

Protéines dans le vide : Dépendance en temps de simulations

Sur la Fig. E.1, nous avons représenté les variations de l'écart quadratique moyen des autres thionines étudiées en fonction du temps de simulation. Nous avons simulé 4 trajectoires différentes : 50ps, 100ps, 200ps et 1ns. Le MSD augmente avec le temps de production pour les autres thionines également. Plus la trajectoire sera longue, plus la simulation prendra en compte des mouvements que subissent les atomes.

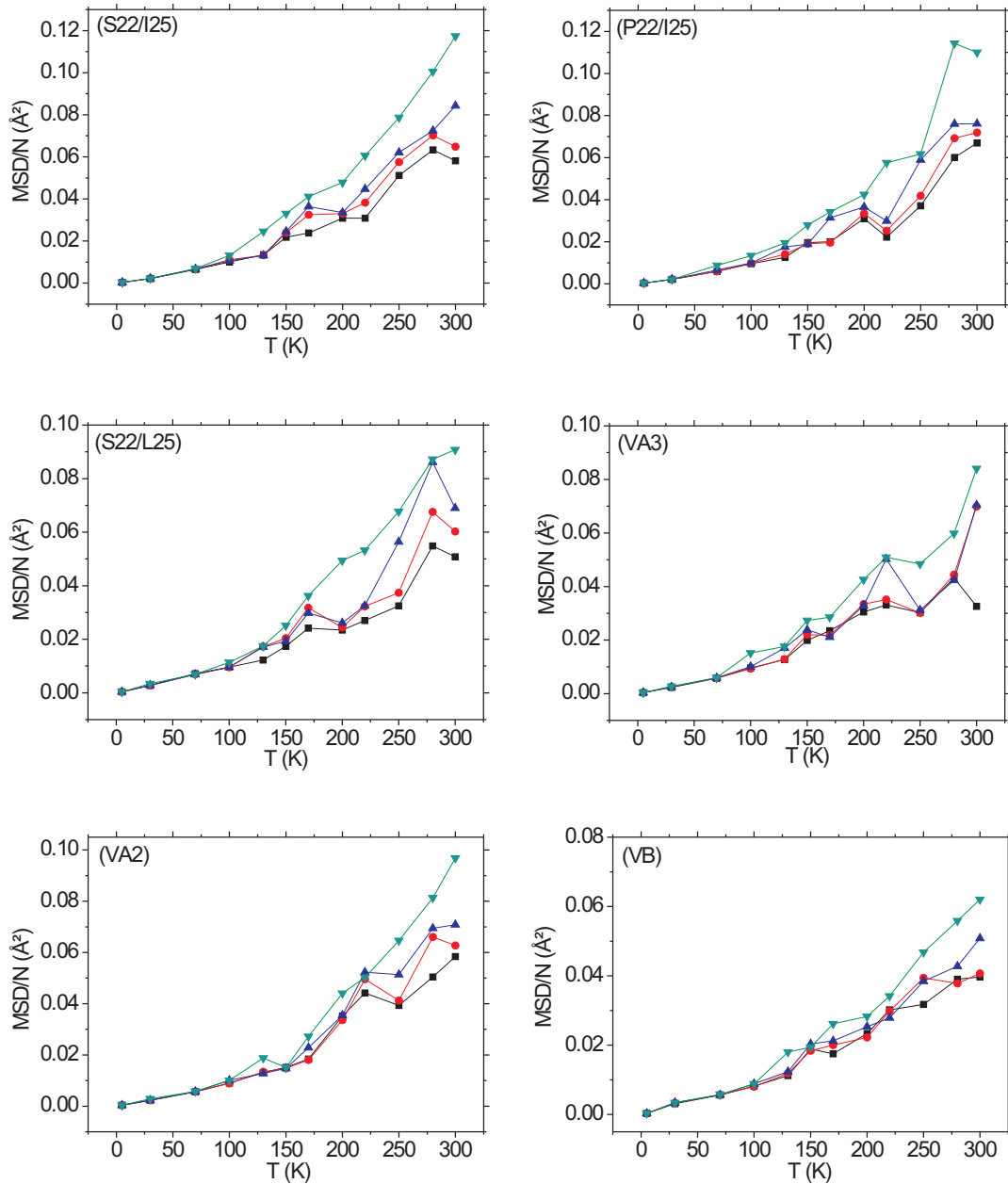


FIG. E.1 – Variation de l'écart quadratique moyen selon le temps de production : 50ps (courbe noire), 100ps (courbe rouge), 200ps (courbe bleue) et 1ns (courbe verte) pour les thionines dans le vide.

Dispersion des fluctuations des protéines déshydratées

Sur la Fig. F.1 sont représentées les variations des dispersions $\frac{\sigma^2}{T^2}$ des protéines dans le vide (courbe noire) et de leur écart quadratique moyen $\frac{MSD}{NT}$ (courbe bleue). Nous constatons que la dispersion des fluctuations subit la même évolution que la flexibilité de la protéine. Plus la température augmente, plus la protéine présente de zones mobiles.

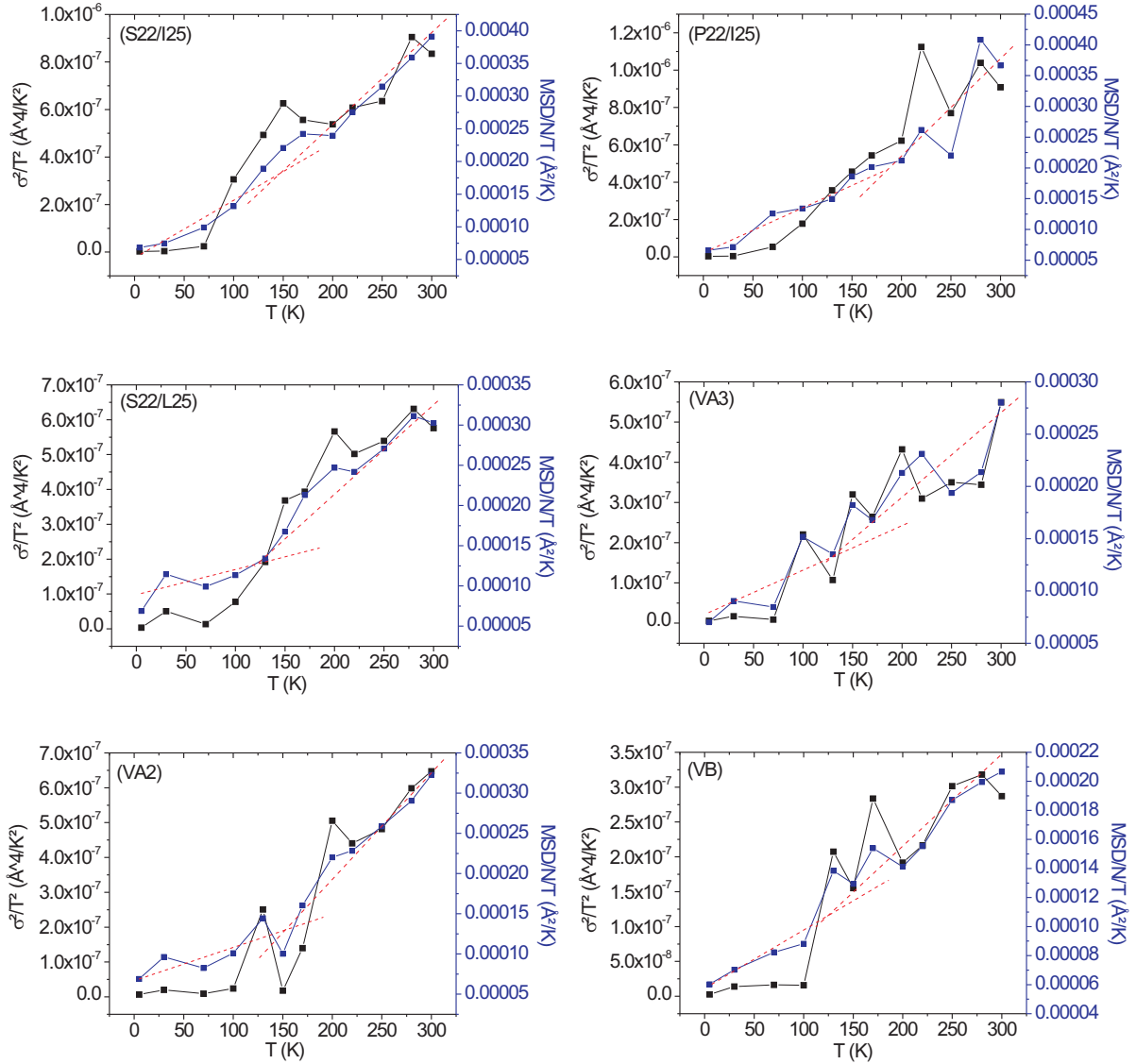


FIG. F.1 – Evolution de la dispersion des fluctuations (courbe noire) comparée à la variation du $\frac{MSD}{NT}$ des protéines (courbe bleue) en fonction de la température, dans le vide sur une trajectoire de 1ns. Les lignes rouges en traits interrompus correspondent aux régressions linéaires des états R et F.

**Augmentation de la mobilité des
atomes dans le vide avec la
température**

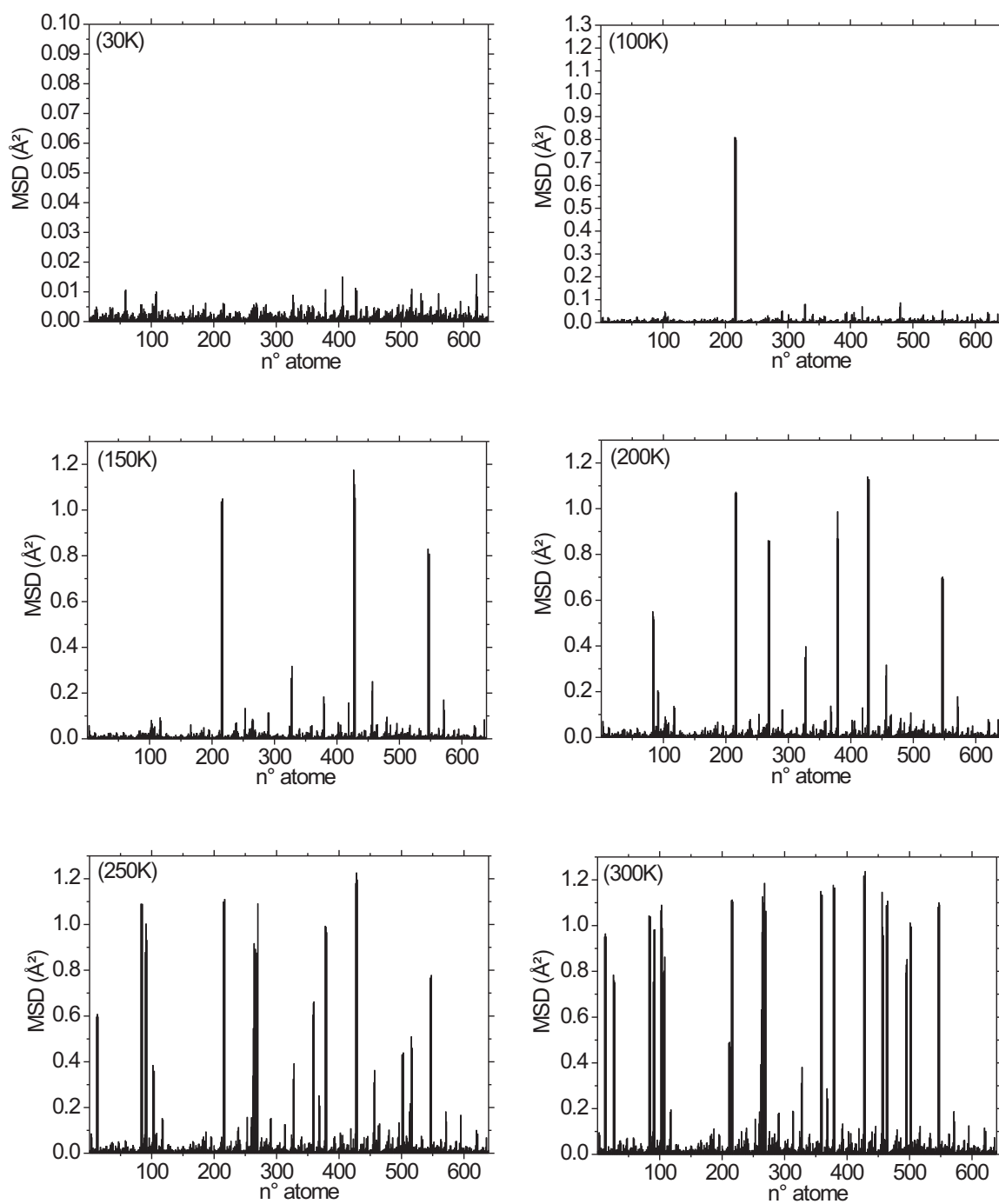


FIG. G.1 – Variation de la mobilité des atomes de la crambine S22/I25 en fonction de la température dans le vide, sur une trajectoire de 1ns.

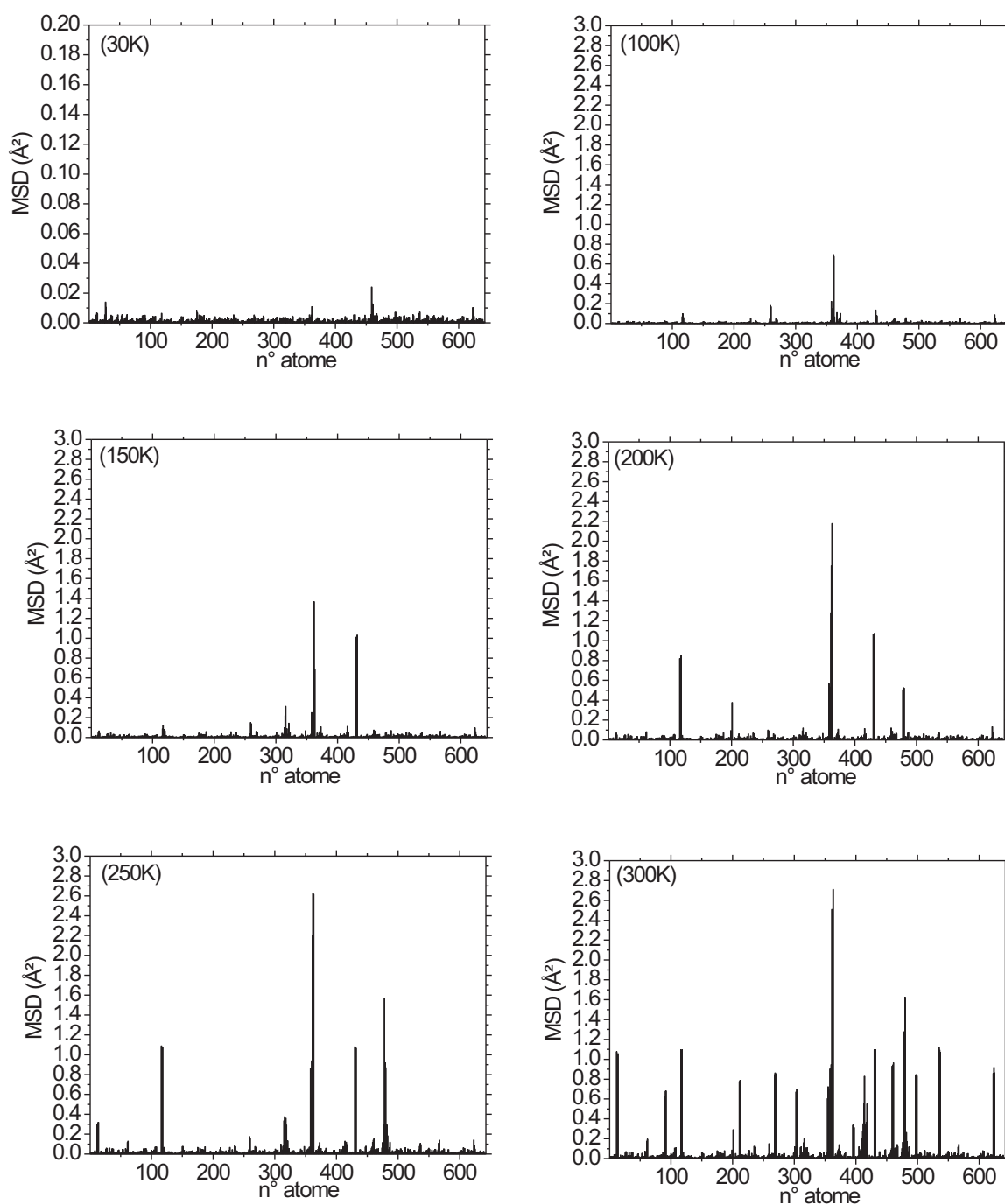


FIG. G.2 – Variation de la mobilité des atomes de la crambine P22/I25 en fonction de la température dans le vide, sur une trajectoire de 1ns.

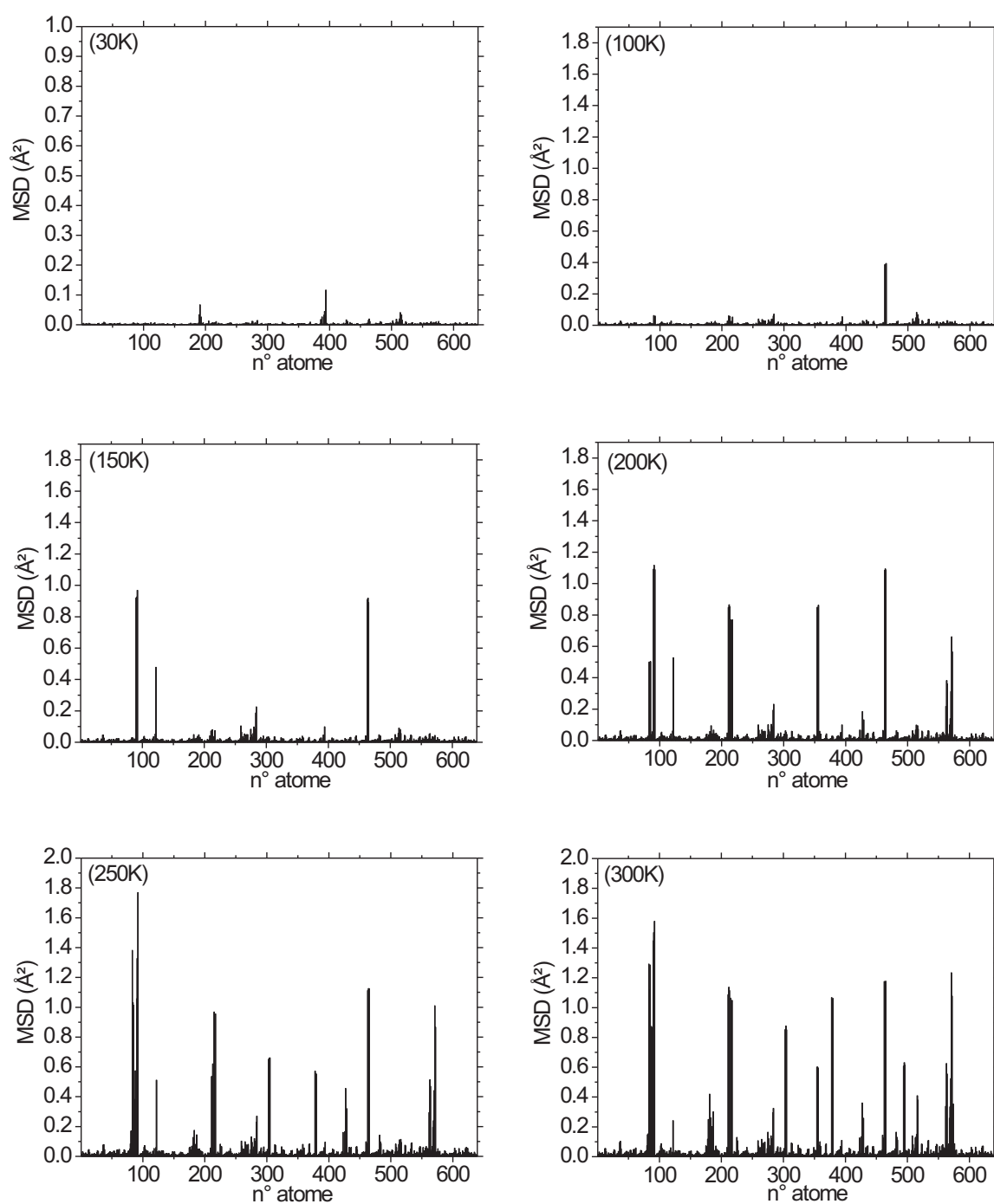


FIG. G.3 – Variation de la mobilité des atomes de la crambine S22/L25 en fonction de la température dans le vide, sur une trajectoire de 1ns.

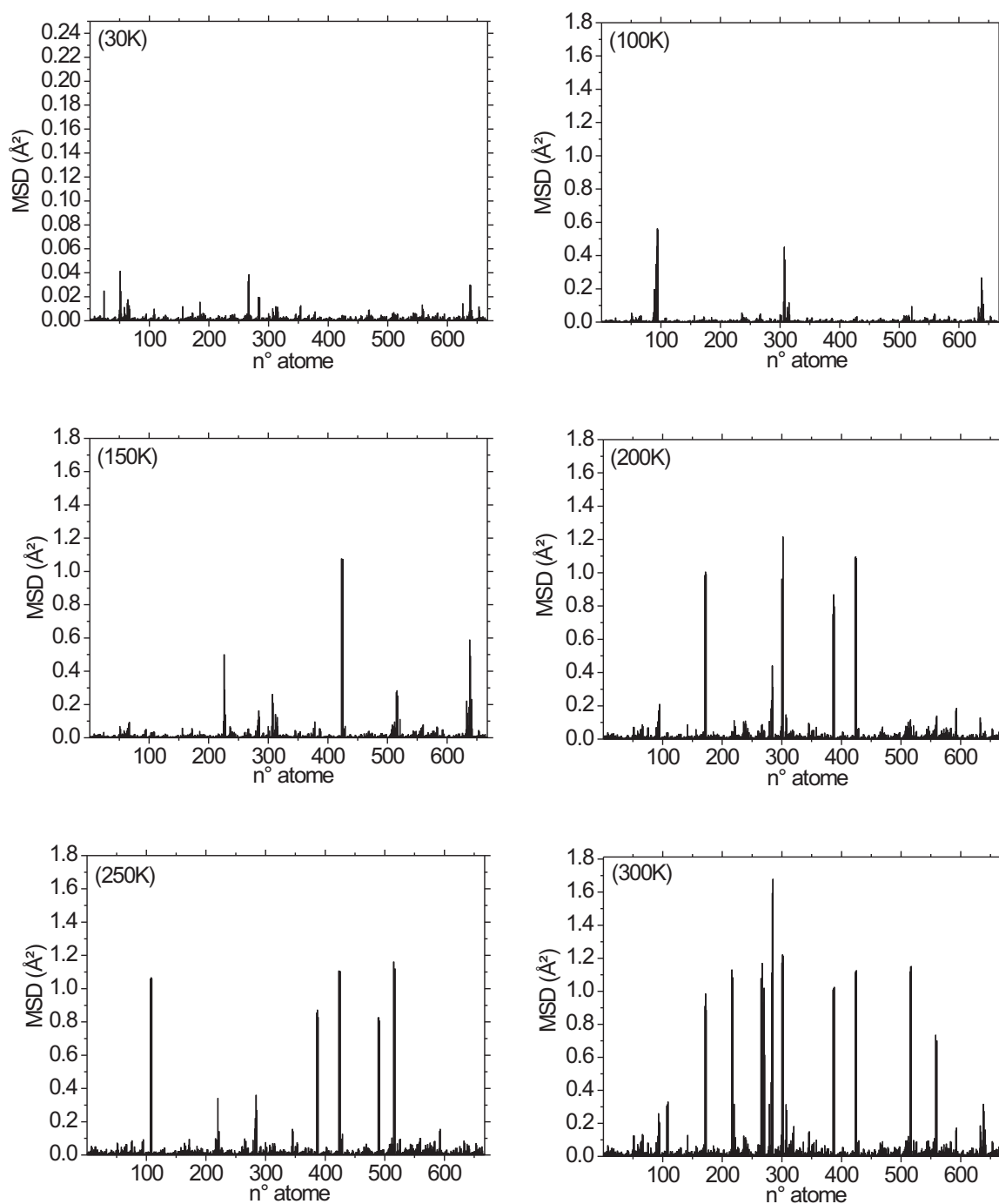


FIG. G.4 – Variation de la mobilité des atomes de la viscotoxine A3 en fonction de la température dans le vide, sur une trajectoire de 1ns.

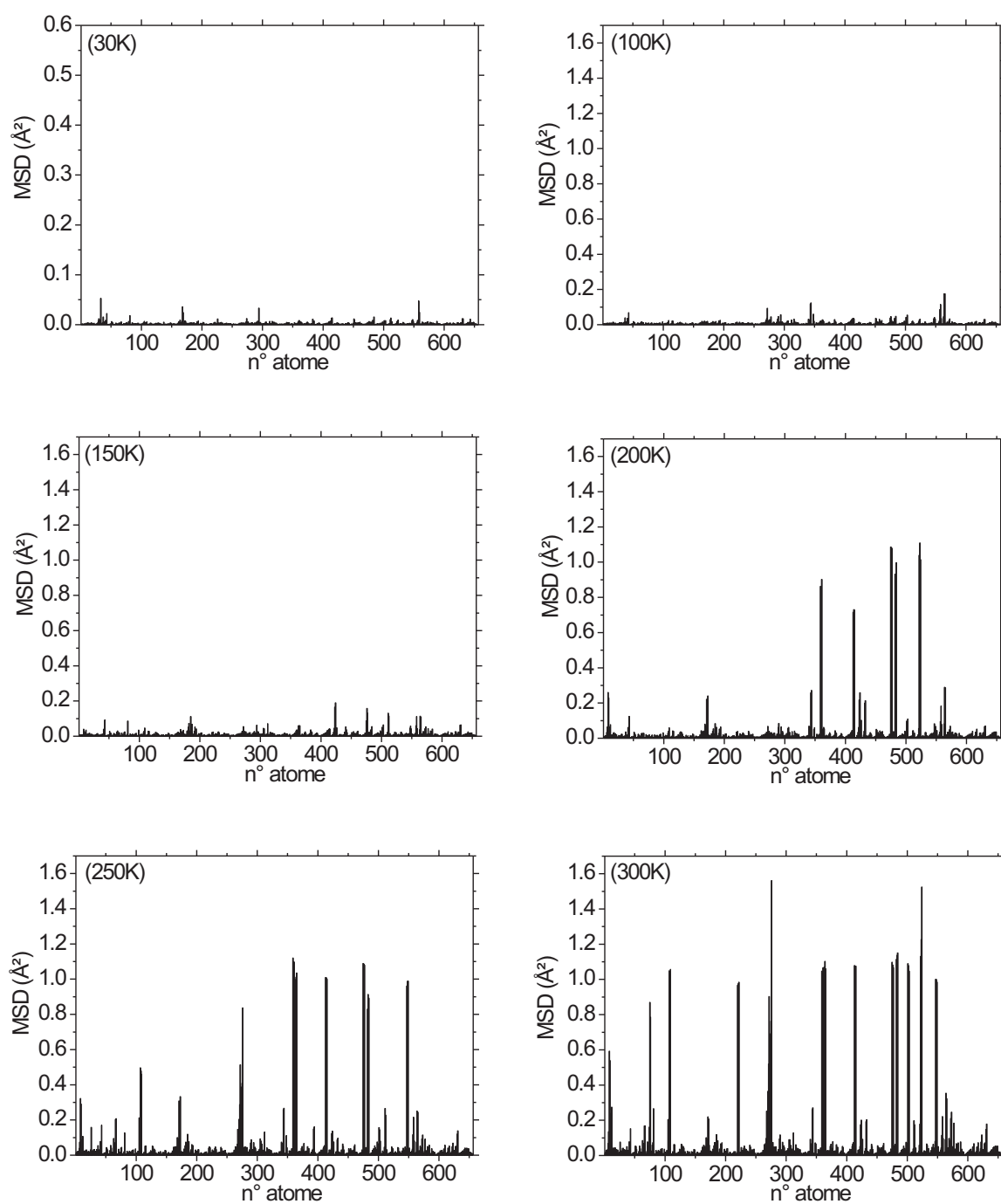


FIG. G.5 – Variation de la mobilité des atomes de la viscotoxine A2 en fonction de la température dans le vide, sur une trajectoire de 1ns.

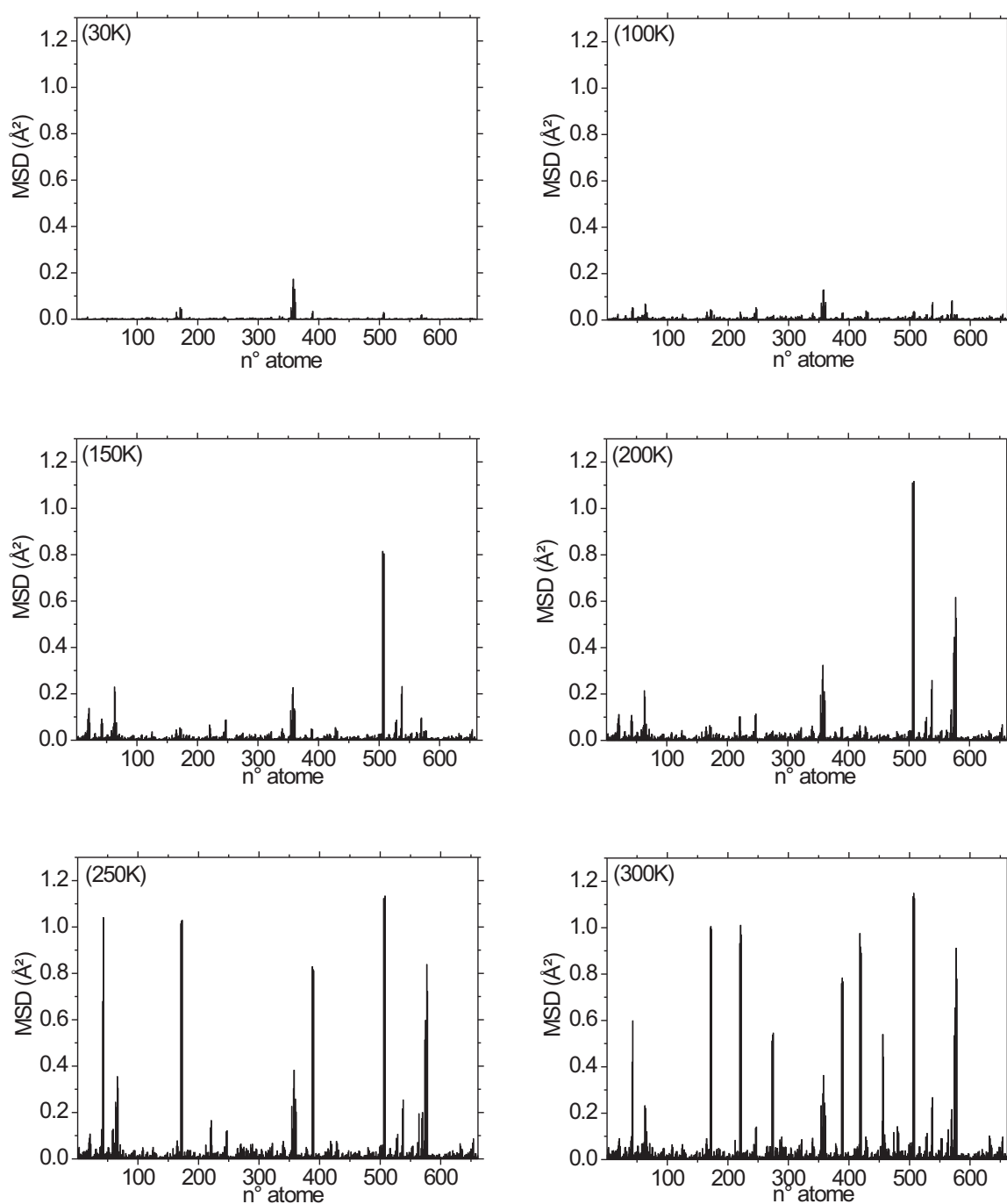


FIG. G.6 – Variation de la mobilité des atomes de la viscotoxine B en fonction de la température dans le vide, sur une trajectoire de 1ns.

**Dépendance de la mobilité avec le
temps de production des protéines
dans le vide**

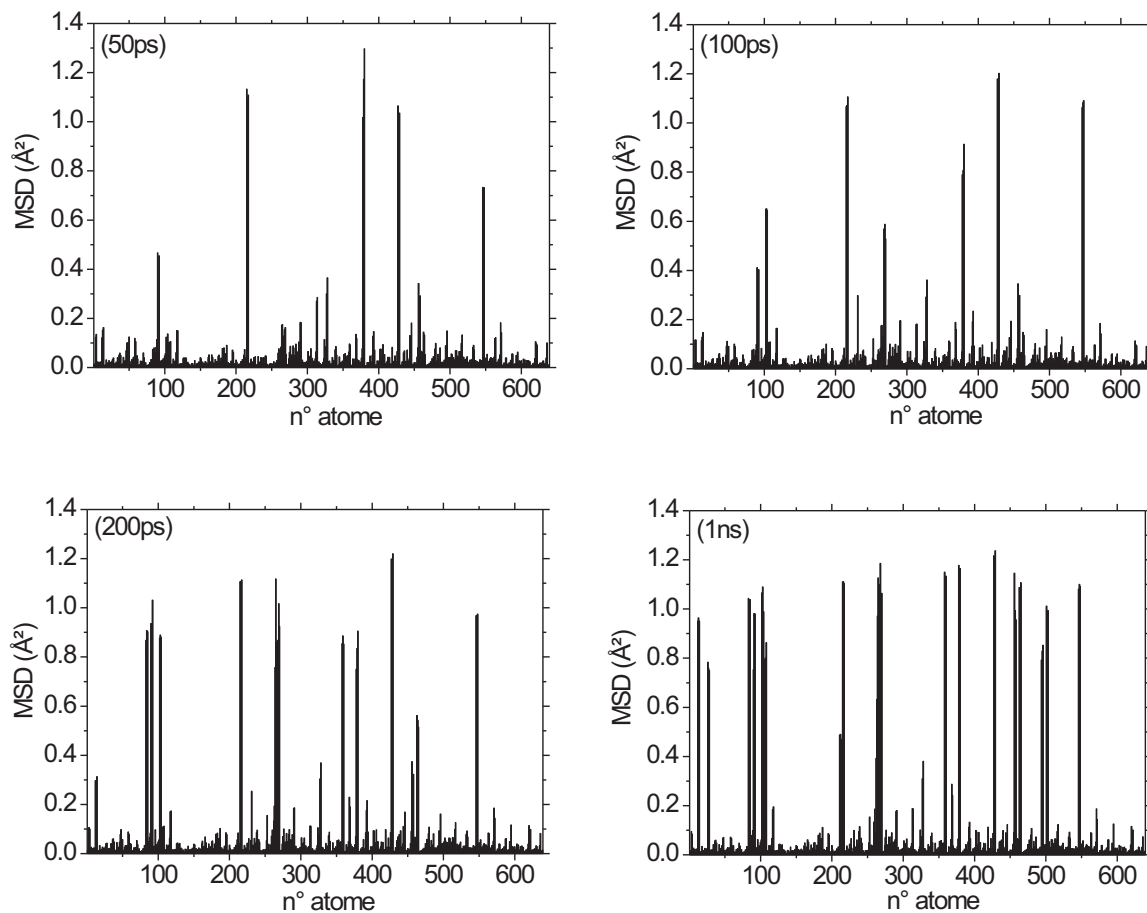


FIG. H.1 – Variation de la mobilité des atomes de la crambine S22/I25 en fonction du temps de production, à 300K.

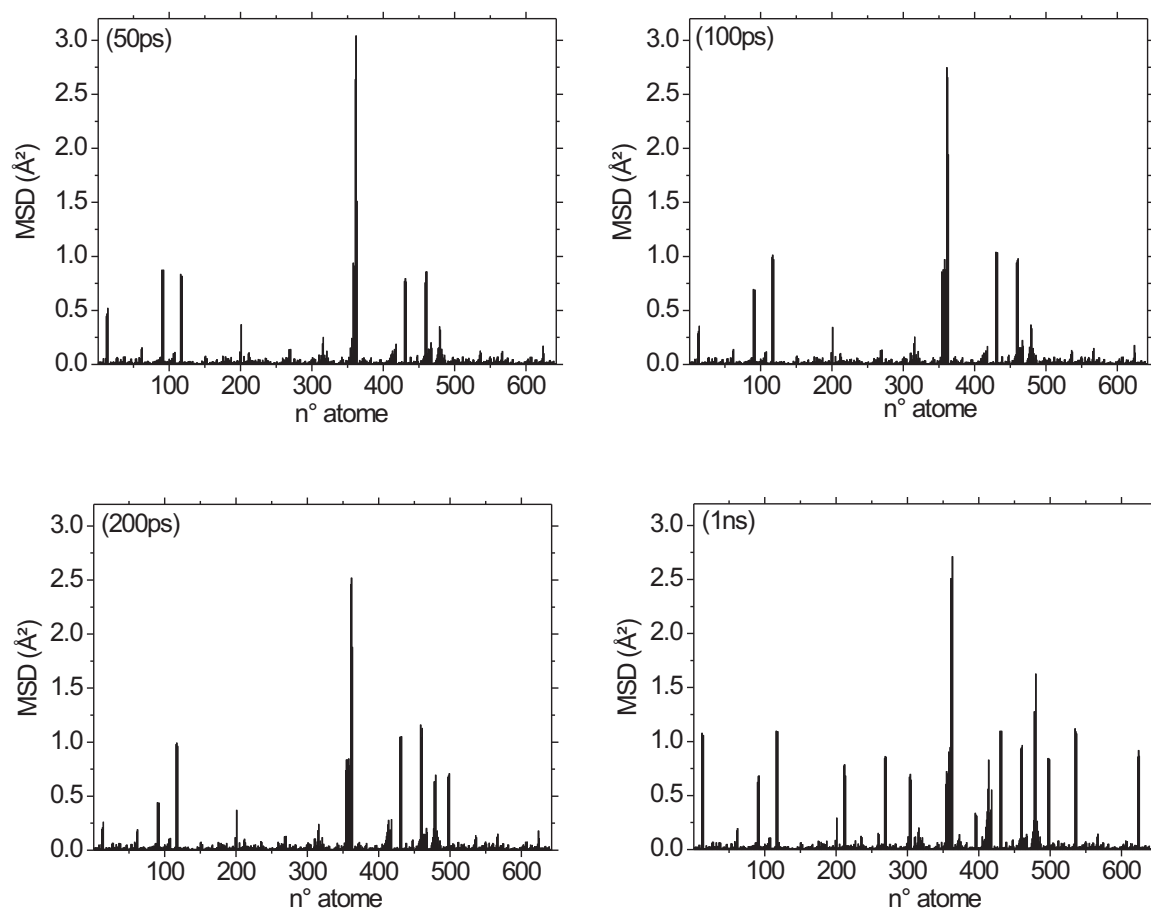


FIG. H.2 – Variation de la mobilité des atomes de la crambine P22/I25 en fonction du temps de production, à 300K.

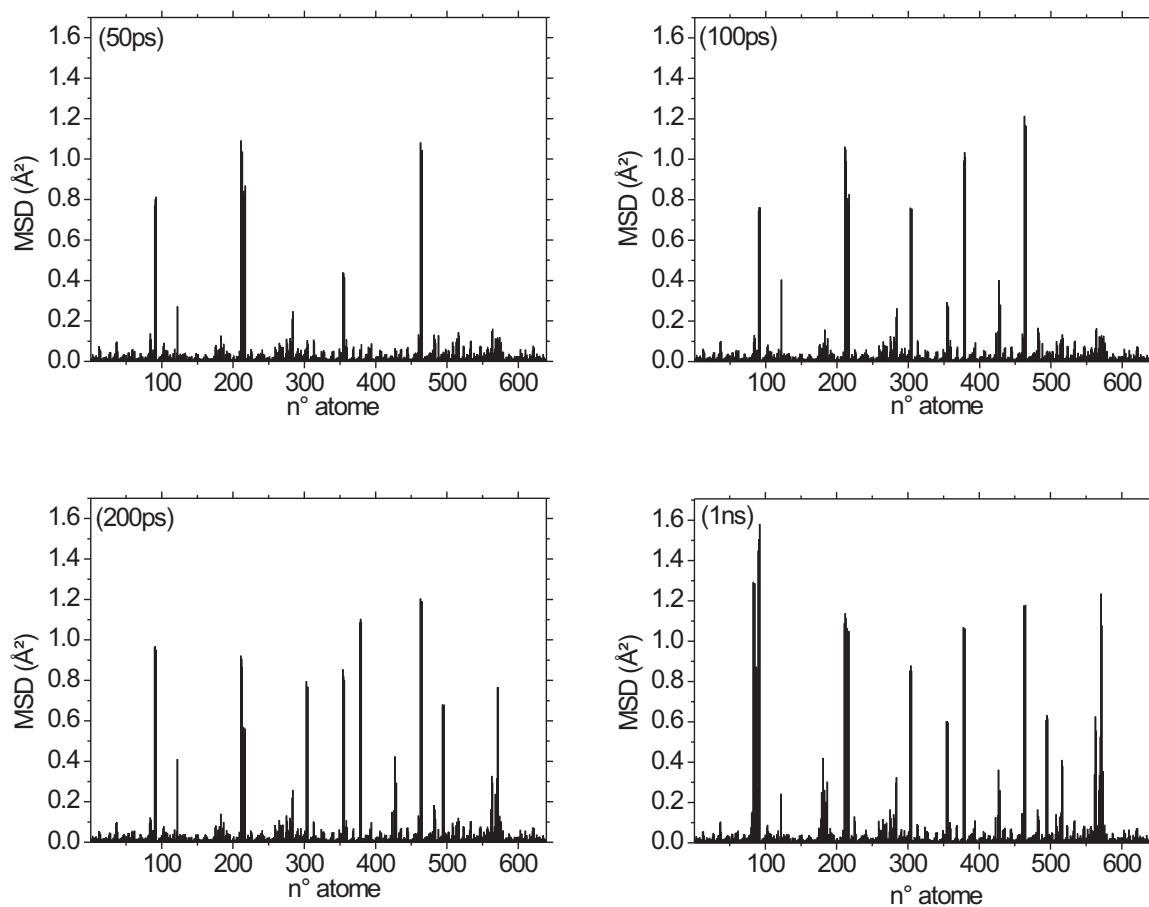


FIG. H.3 – Variation de la mobilité des atomes de la crambine S22/L25 en fonction du temps de production, à 300K.

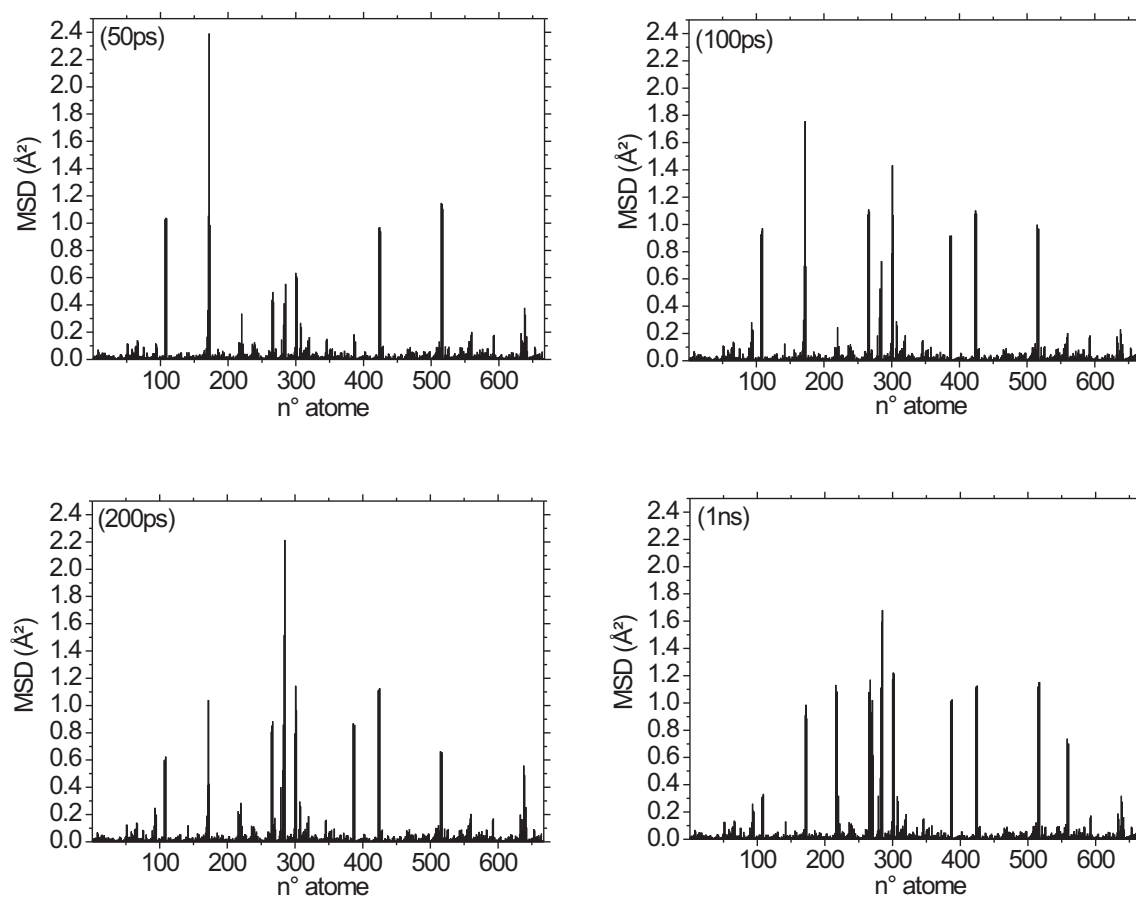


FIG. H.4 – Variation de la mobilité des atomes de la viscotoxine A3 en fonction du temps de production, à 300K.

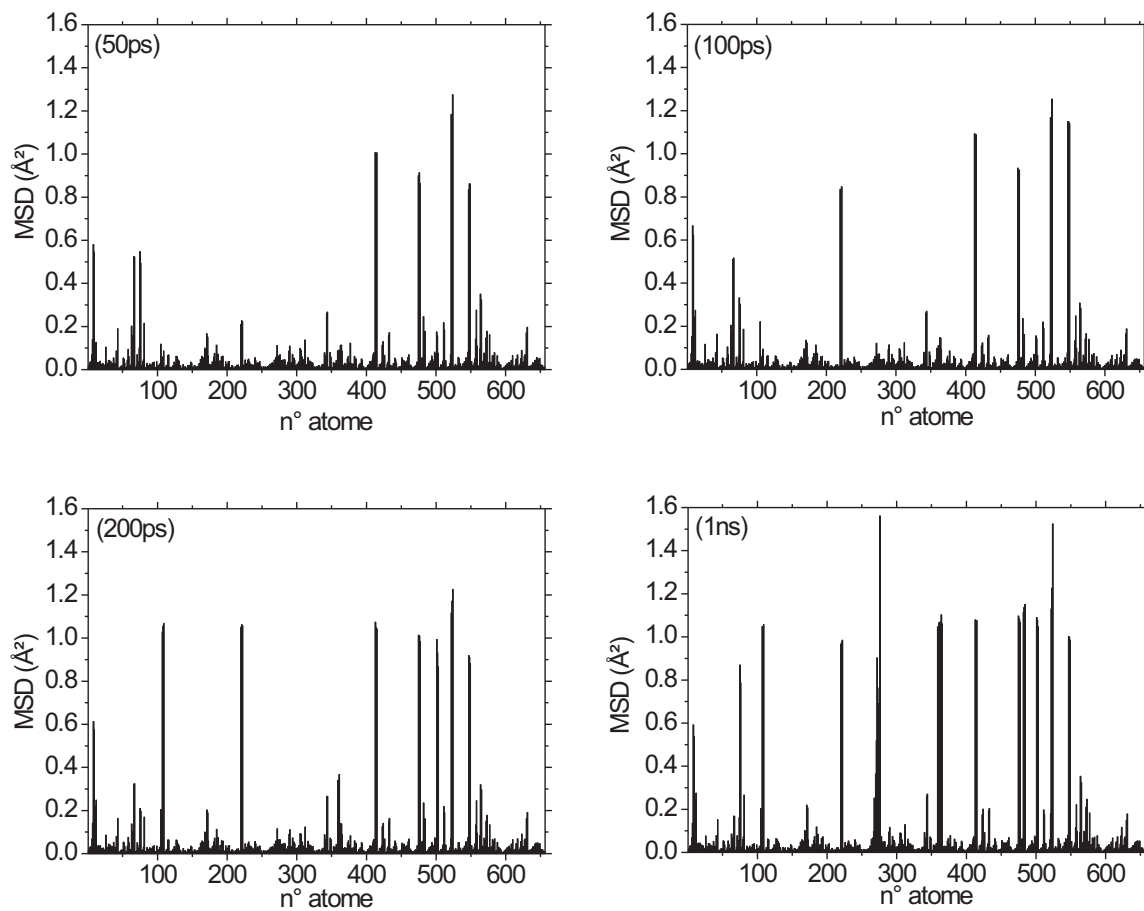


FIG. H.5 – Variation de la mobilité des atomes de la viscotoxine A2 en fonction du temps de production, à 300K.

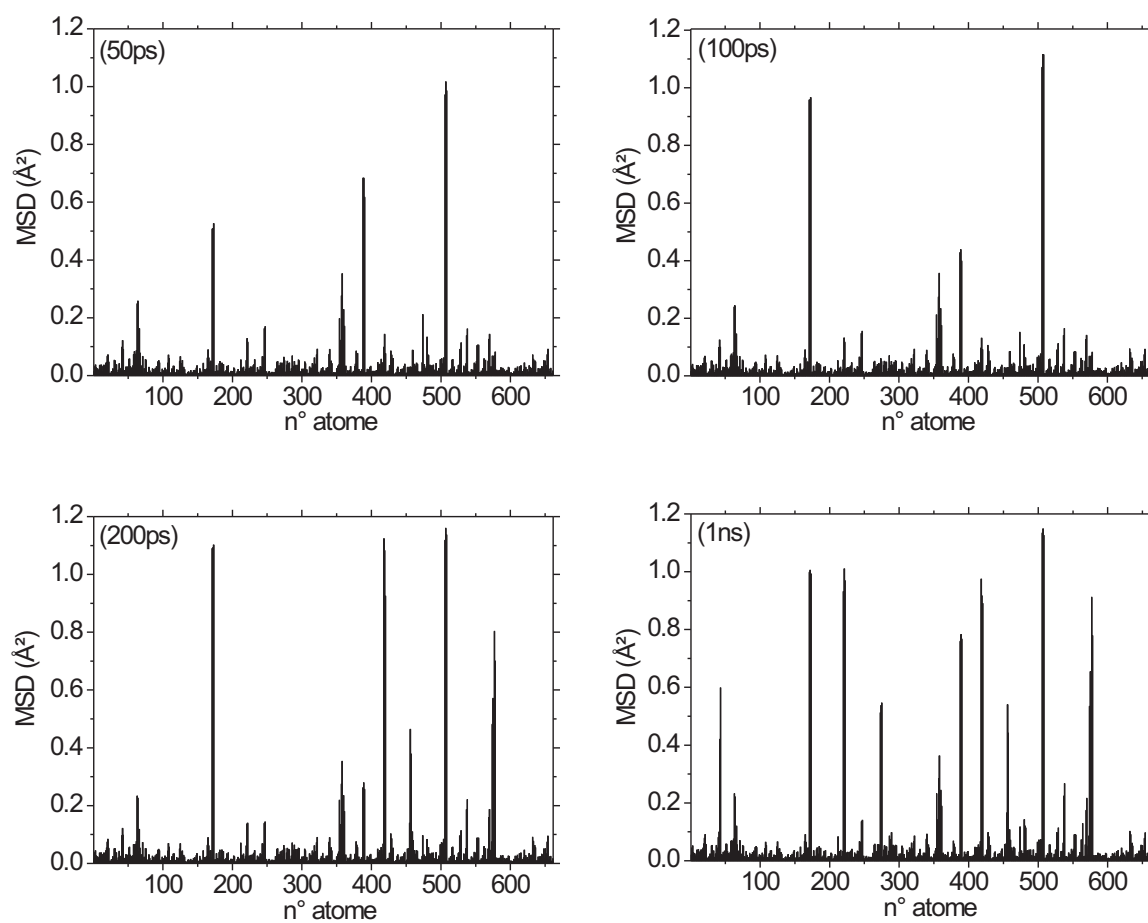


FIG. H.6 – Variation de la mobilité des atomes de la viscotoxine B en fonction du temps de production, à 300K.

Mobilité des résidus dans le vide

Sur la Fig. I.1, nous avons représenté le nombre de résidus suivant leur valeur du $\frac{MSD}{T}$. Ces histogrammes mettent en évidence le début de la flexibilité des thionines à environ 100/130K. C'est à cette température (100K : histogramme bleu-vert, 130K : histogramme fushia) que le nombre de résidus s'étale sur des valeurs du $\frac{MSD}{T}$ plus grandes.

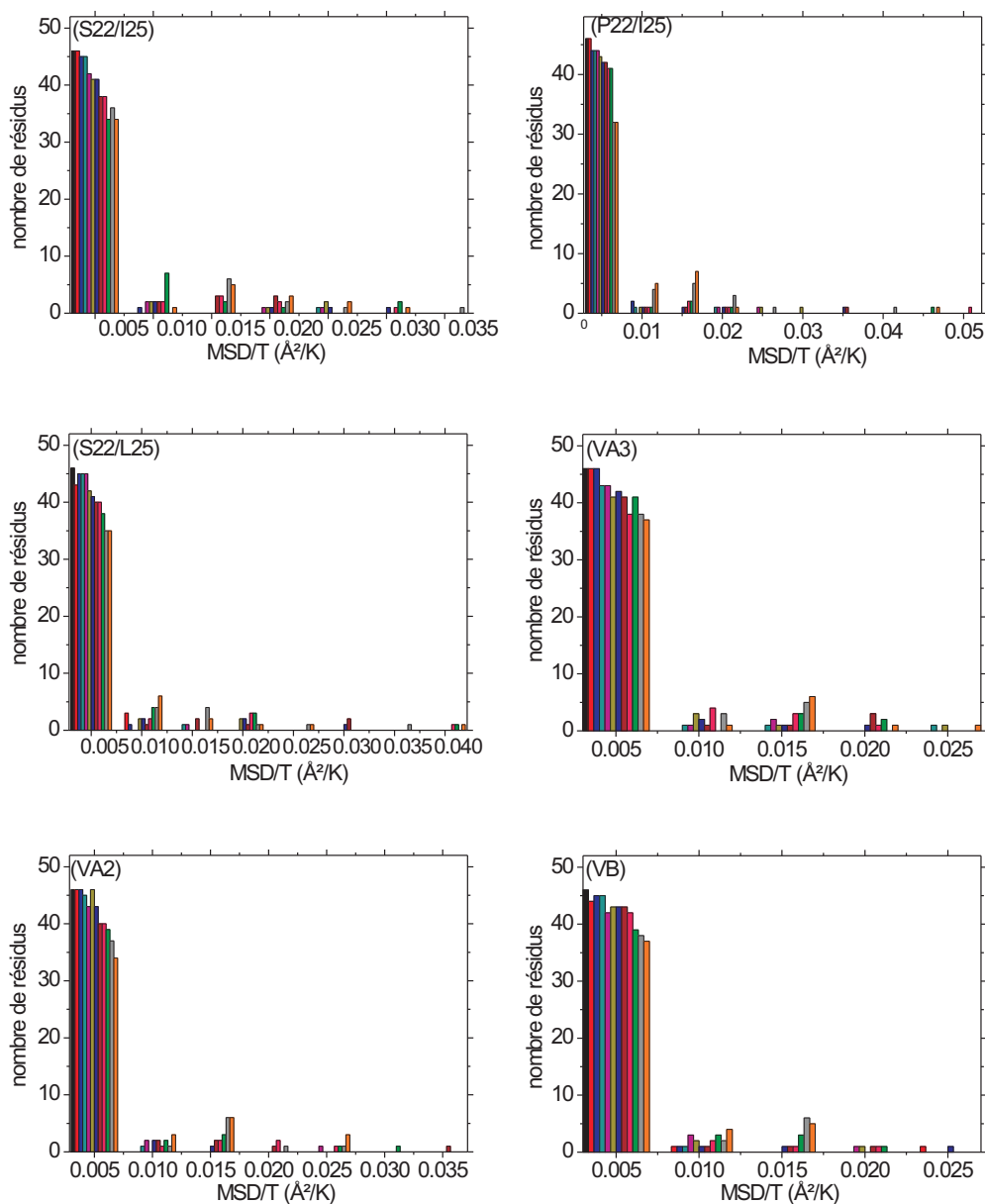


FIG. I.1 – Variation du nombre de résidus suivant leur valeur du $\frac{MSD}{T}$ pour les crambines S22/I25, P22/I25, S22/L25 et les 3 viscotoxines déshydratées, sur une trajectoire de 1ns. Les histogrammes correspondent aux températures suivantes : noir = 5K, rouge = 30K, bleu = 70K, bleu-vert = 100K, fushia = 130K, kaki = 150K, bleu marine = 170K, marron = 200K, rose = 220K, vert = 250K, gris = 280K, orange = 300K.

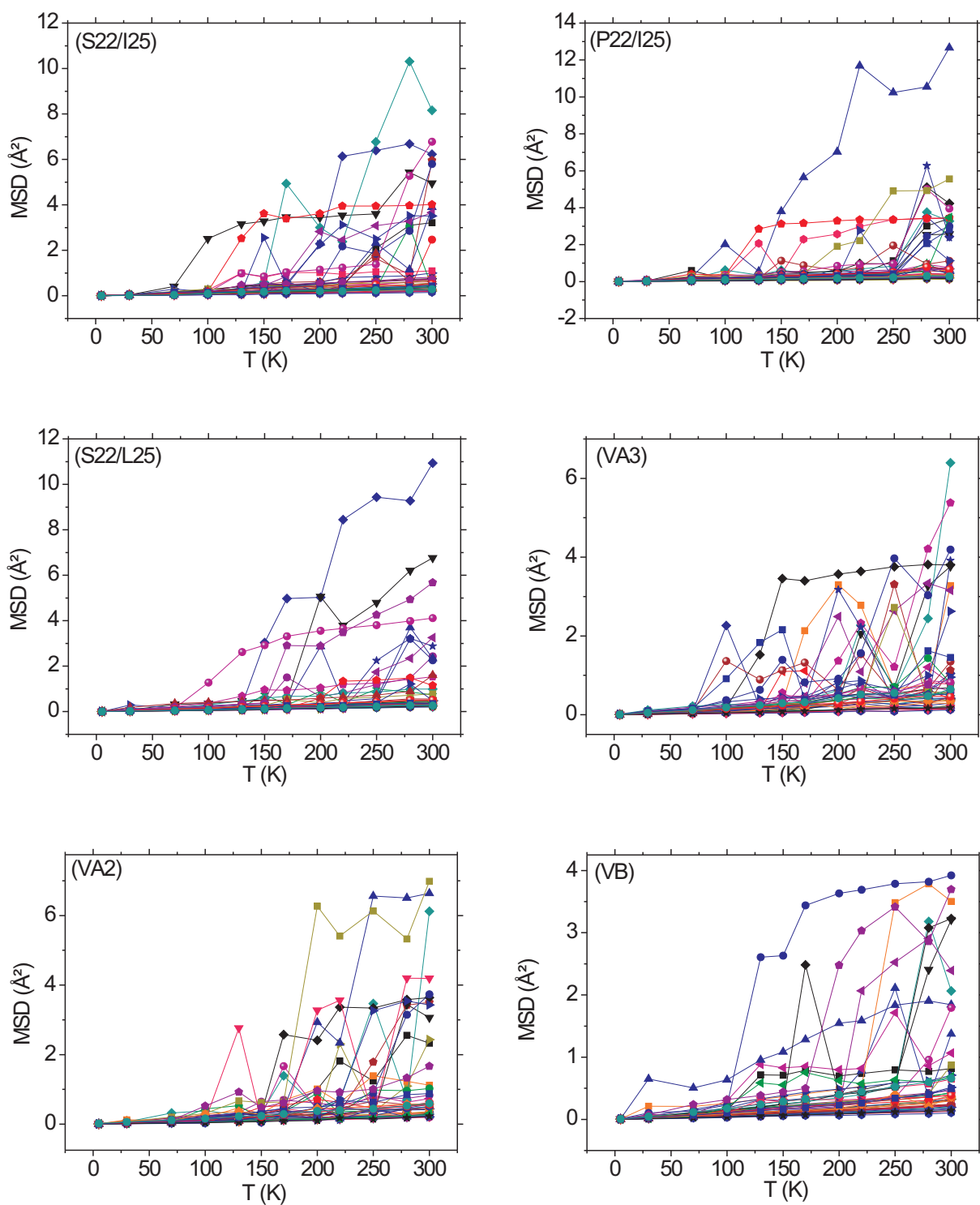


FIG. I.2 – Variation du MSD de chaque acide aminé avec la température pour les autres thionines dans le vide, sur une trajectoire de 1ns.

Modèle de Doster

J.1 Ajustement des paramètres des thionines hydratées

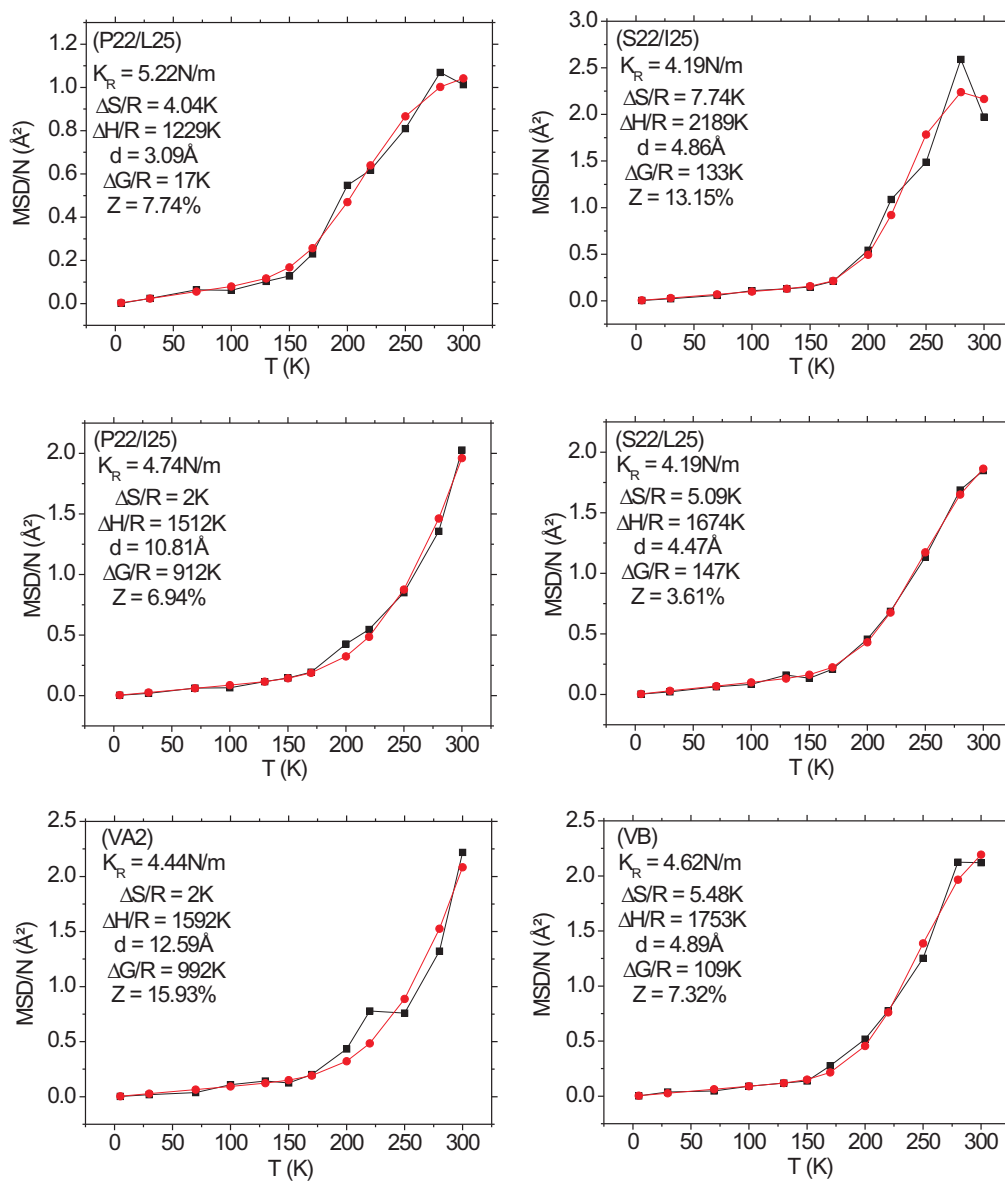


FIG. J.1 – Ajustement des paramètres du modèle de Doster pour les protéines hydratées. La courbe noire est l'écart quadratique moyen calculé par nos simulations de dynamique moléculaire et la courbe rouge est le modèle de Doster.

J.2 Deuxième ajustement pour VA3

Un deuxième ajustement a été calculé pour la viscotoxine A3, cf Fig. J.2. Dans ce calcul, tous les paramètres sont ajustés et les données sont non-pondérées. Nous obtenons également des résultats en accord, avec un pourcentage d'erreur de 6.29%.

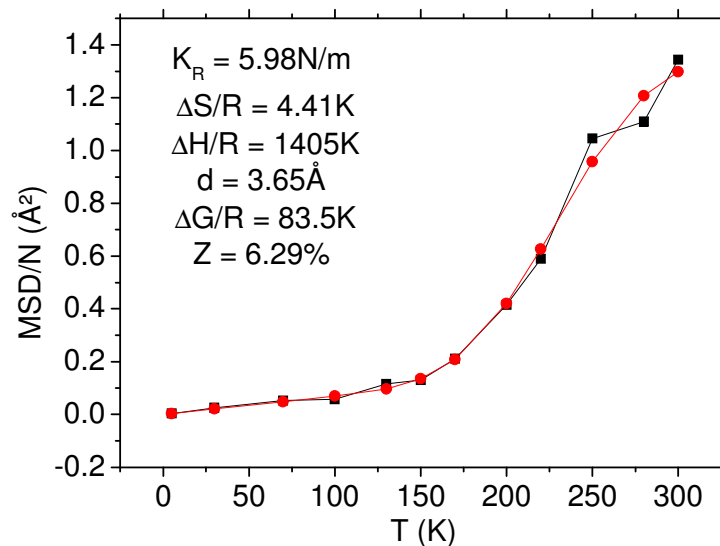


FIG. J.2 – Second ajustement des paramètres du modèle de Doster pour la viscotoxine VA3 hydratée. La courbe noire est l'écart quadratique moyen calculé par nos simulations de dynamique moléculaire et la courbe rouge est le modèle de Doster.

Distribution des barrières d'énergie des méthyles

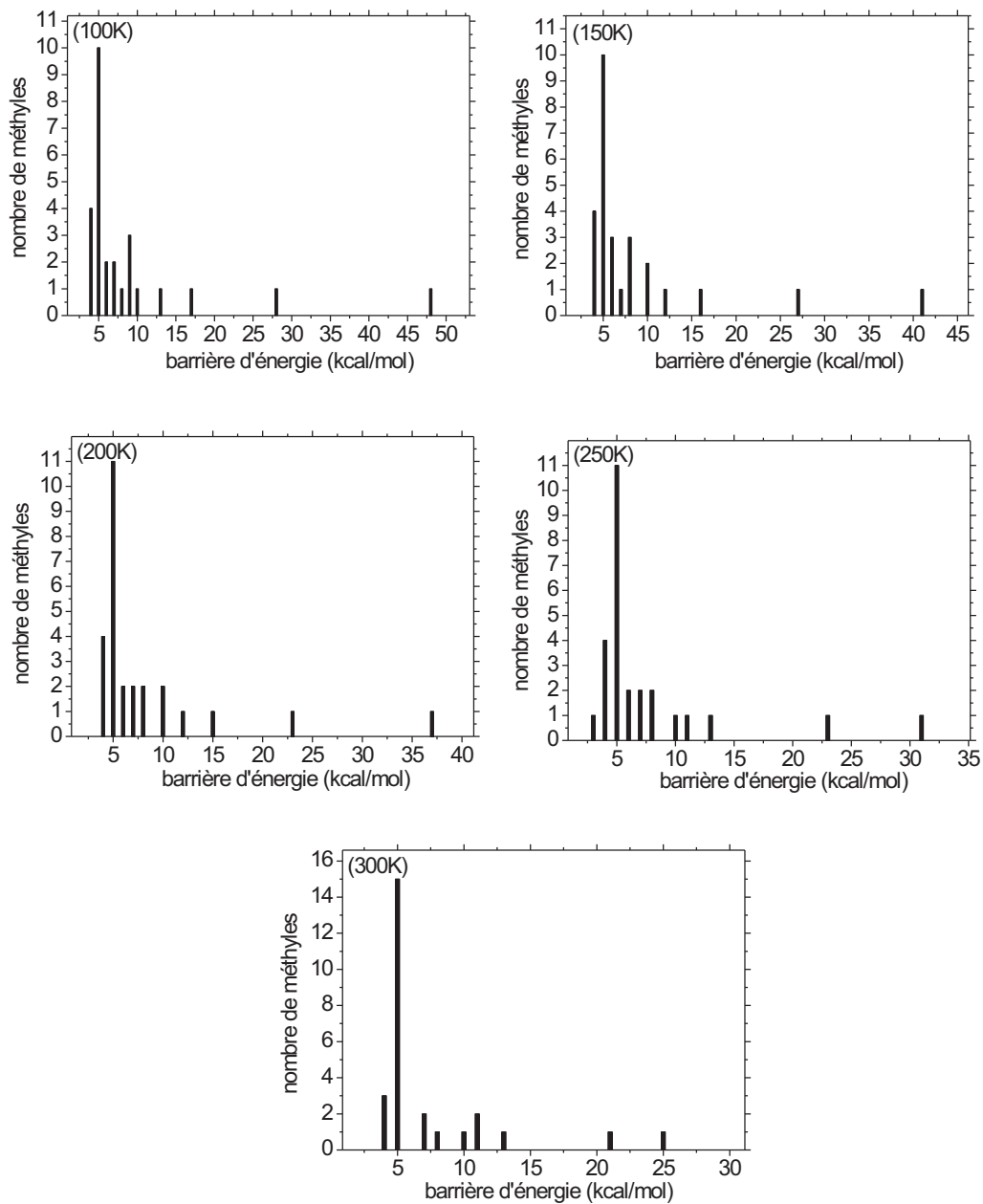


FIG. K.1 – Distribution des barrières d'énergie des méthyles pour la crambine P22/L25 à différentes températures.

**Variation de la mobilité des protéines
hydratées en fonction des 5 conditions
initiales utilisées sur des trajectoires
de 1ns**

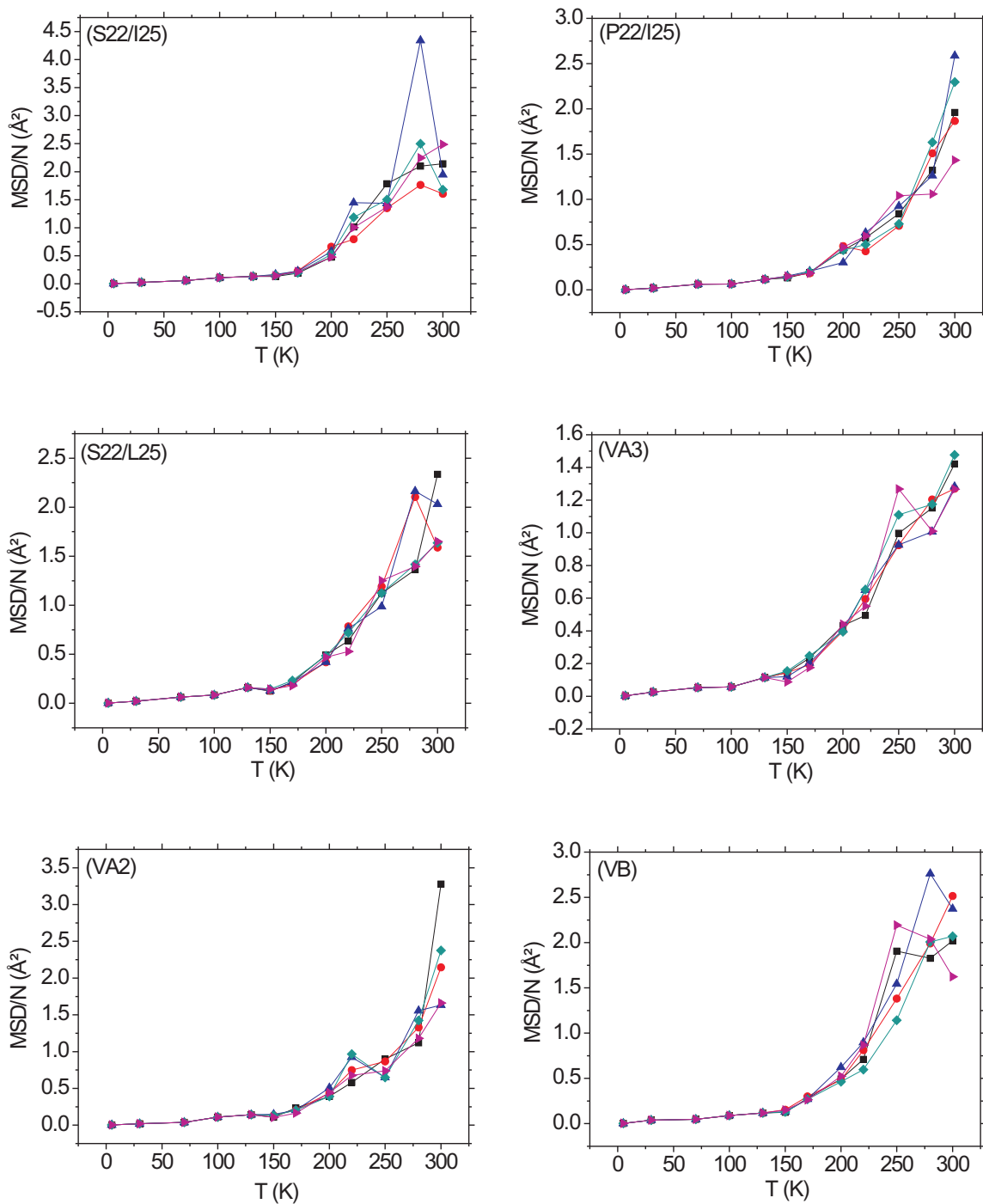


FIG. L.1 – Variation de l'écart quadratique moyen des thionines hydratées sur une trajectoire de 1ns. Les carrés noirs correspondent à la 1ère condition initiale, les ronds rouges à la 2nde, les triangles bleus à la 3ième, les losanges verts à la 4ième et les triangles roses à la 5ième.

Protéines hydratées : Séparation des contributions en type d'atomes

La Fig. M.1 nous permet de constater que les hydrogènes (courbe noire) restent majoritaires dans la mobilité de la protéine hydratée (courbe bleue), bien que d'autres types d'atomes soient susceptibles de subir des mouvements de grandes amplitudes.

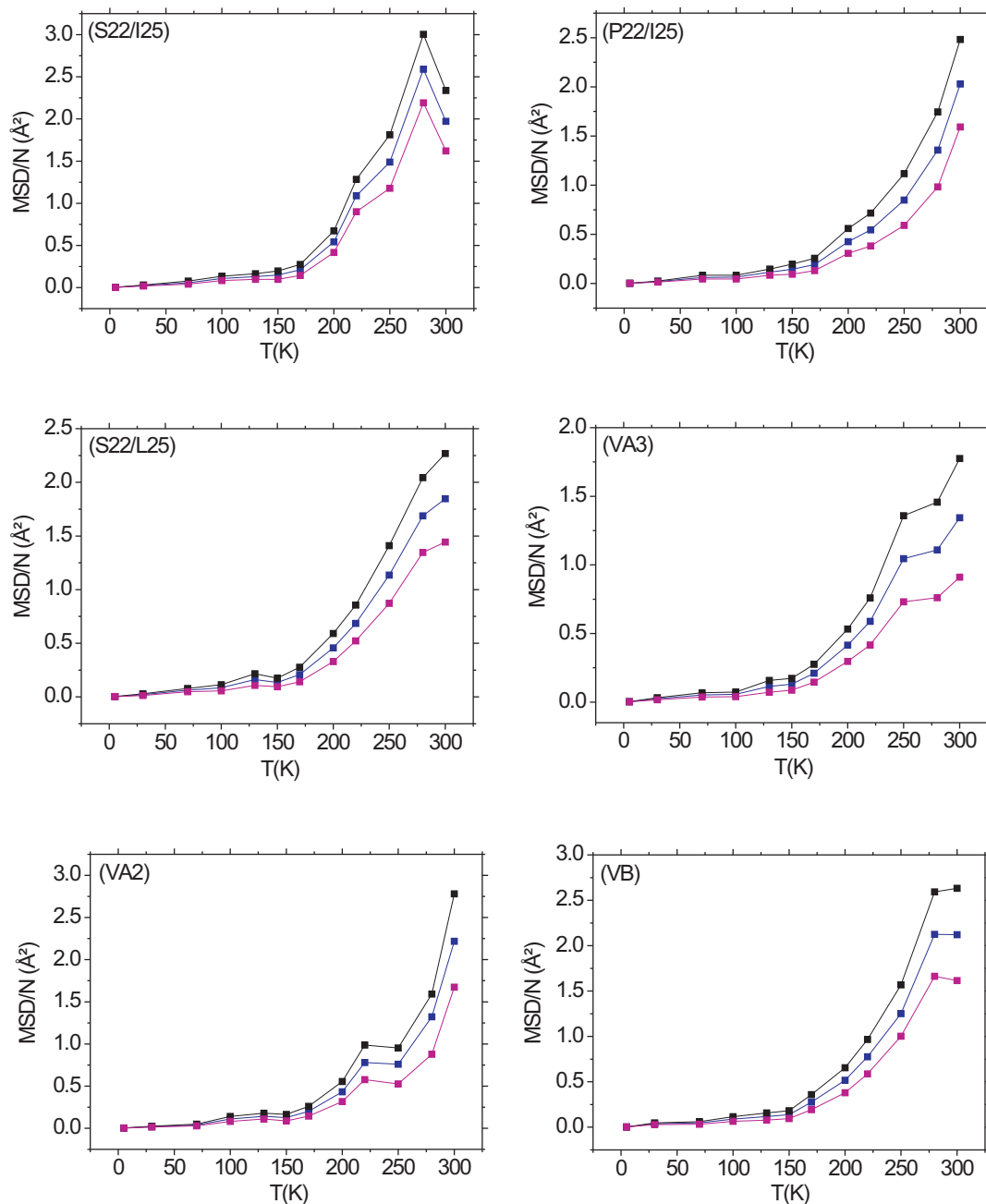


FIG. M.1 – Ecart quadratique moyen des protéines hydratées (courbe bleue) comparé aux différents types d'atomes : pour tous les atomes d'hydrogènes (courbe noire) et tous les atomes qui ne sont pas des hydrogènes (courbe rose), sur une trajectoire de 1ns.

Protéines hydratées : Dépendance en temps de simulations

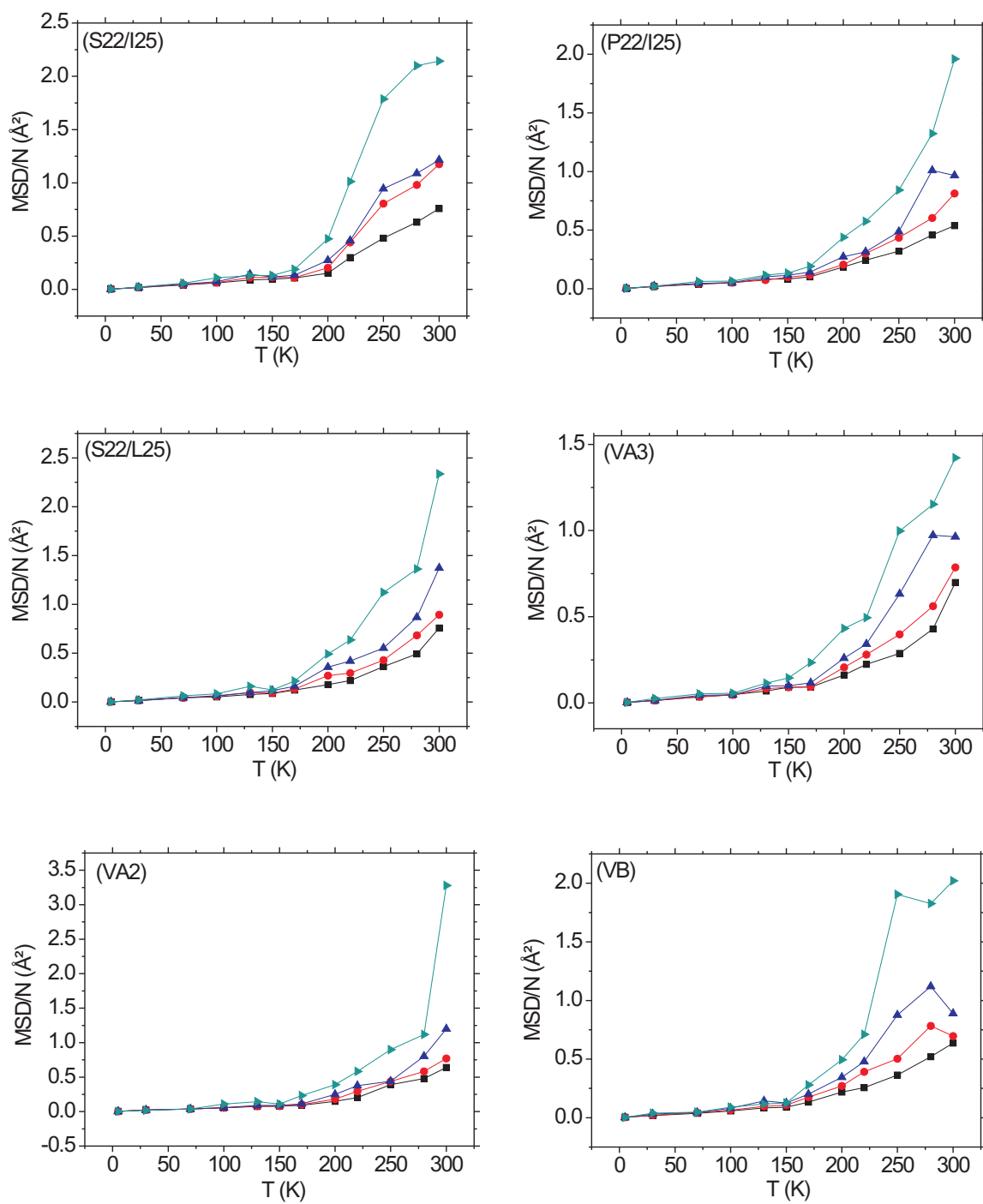


FIG. N.1 – Variation de l'écart quadratique moyen selon le temps de production : 50ps (courbe noire), 100ps (courbe rouge), 200ps (courbe bleue) et 1ns (courbe verte) pour les thionines hydratées.

Dispersion des fluctuations des protéines hydratées

Sur la Fig. O.1 sont représentées les variations des dispersions $\frac{\sigma^2}{T^2}$ des protéines hydratées (courbe noire) et de leur écart quadratique moyen $\frac{MSD}{NT}$ (courbe bleue). Nous constatons que la dispersion des fluctuations subit la même évolution que la mobilité de la protéine, également dans le cas des protéines hydratées. L'augmentation de la température implique une augmentation de zones mobiles.

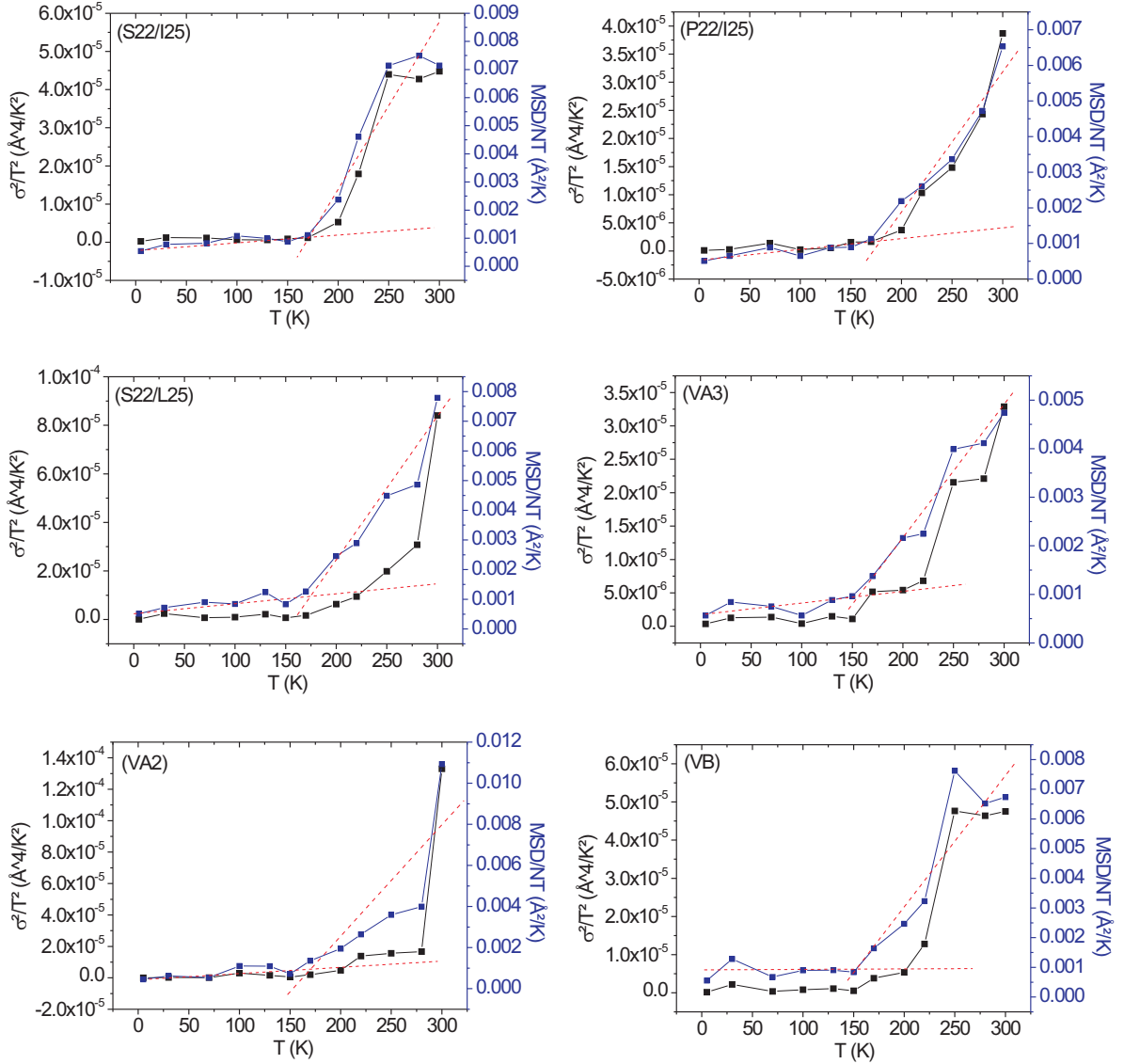


FIG. O.1 – Evolution de la dispersion des fluctuations (courbe noire) comparée à la variation du $\frac{MSD}{NT}$ des protéines hydratées (courbe bleue) en fonction de la température, sur une trajectoire de 1ns. Les lignes rouges en traits interrompus correspondent aux régressions linéaires des états R et F.

**Augmentation de la mobilité des
atomes des protéines hydratées avec
la température**

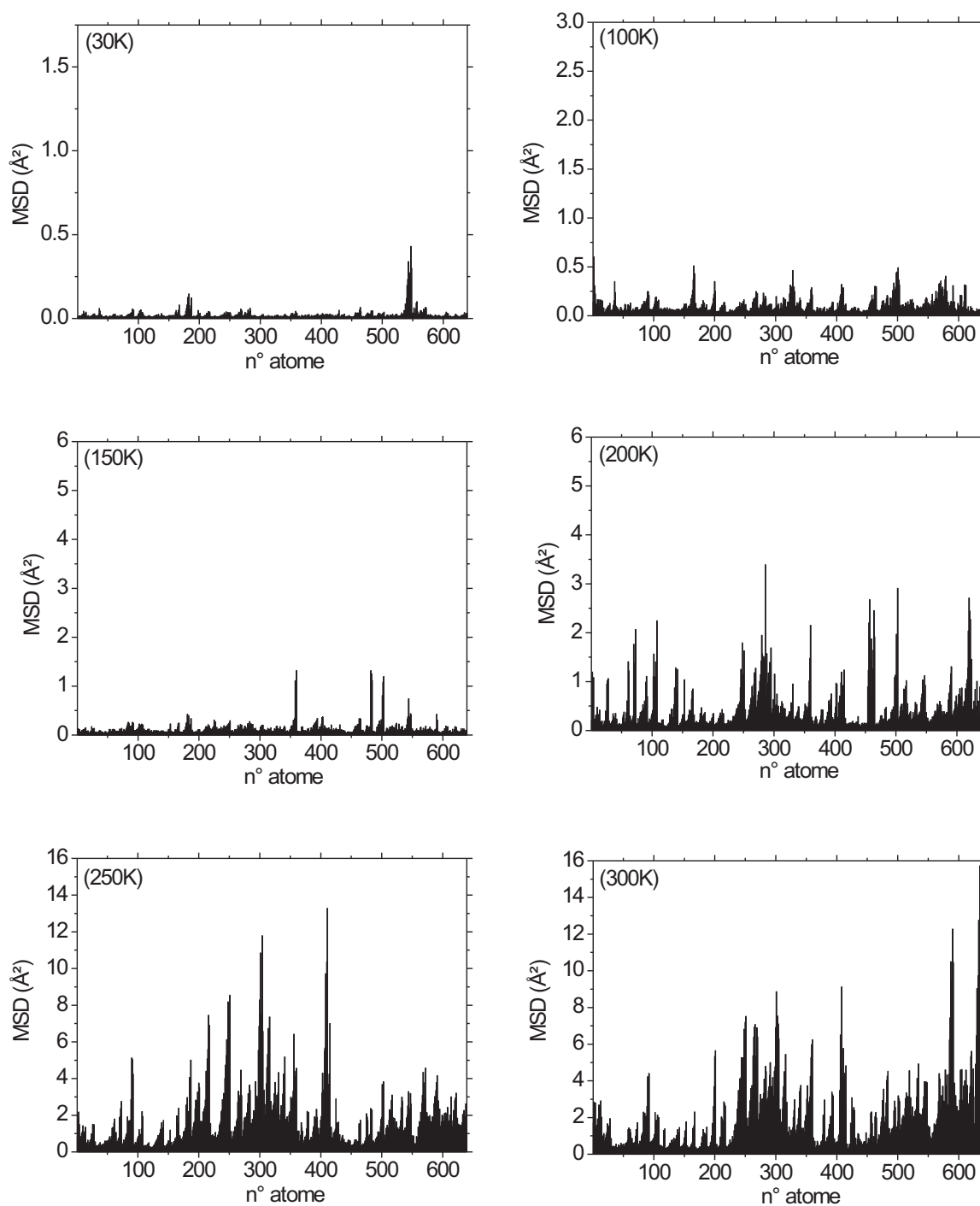


FIG. P.1 – Variation de la mobilité des atomes de la crambine S22/I25 hydratée en fonction de la température, sur une trajectoire de 1ns.

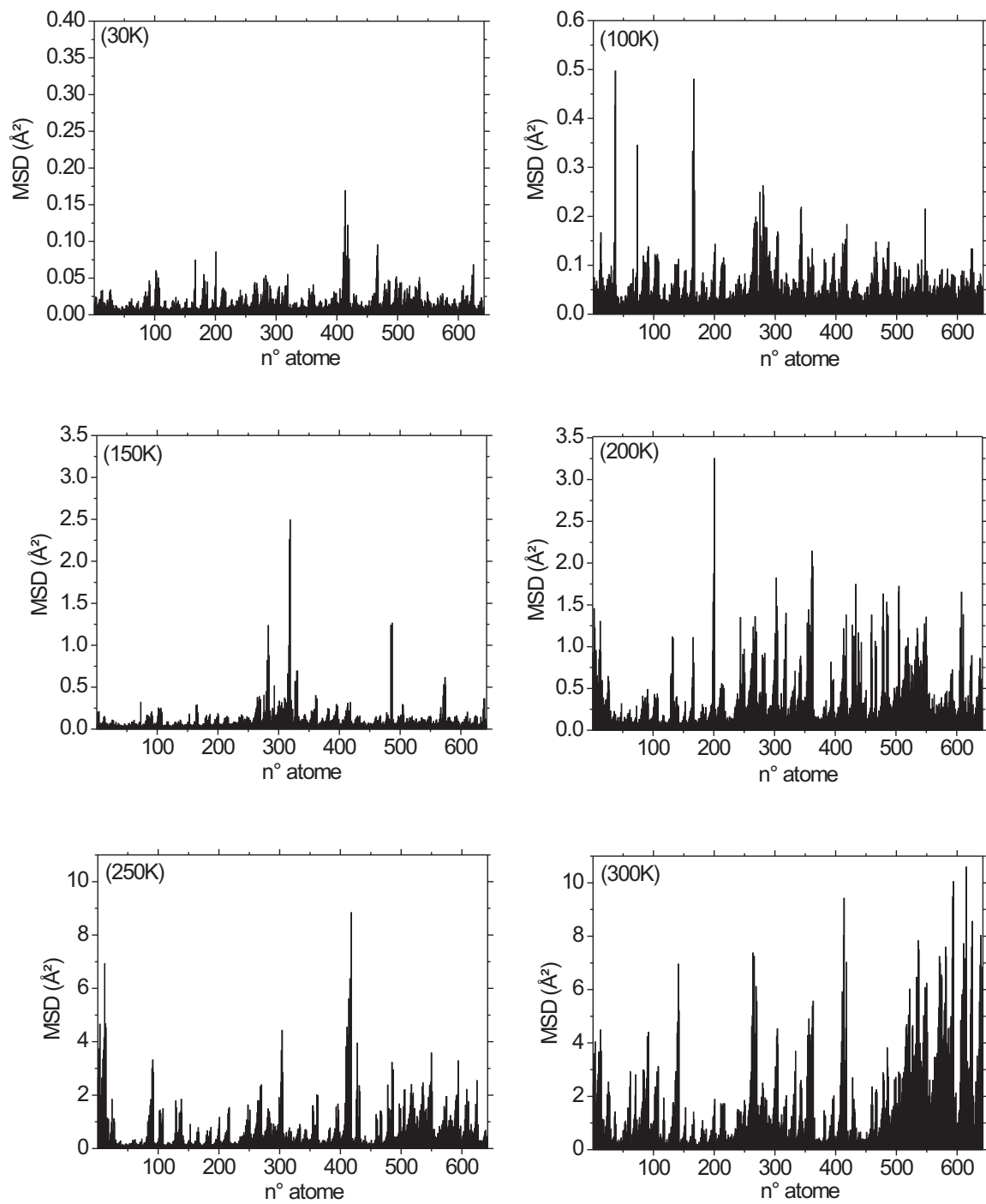


FIG. P.2 – Variation de la mobilité des atomes de la crambine P22/I25 hydratée en fonction de la température, sur une trajectoire de 1ns.

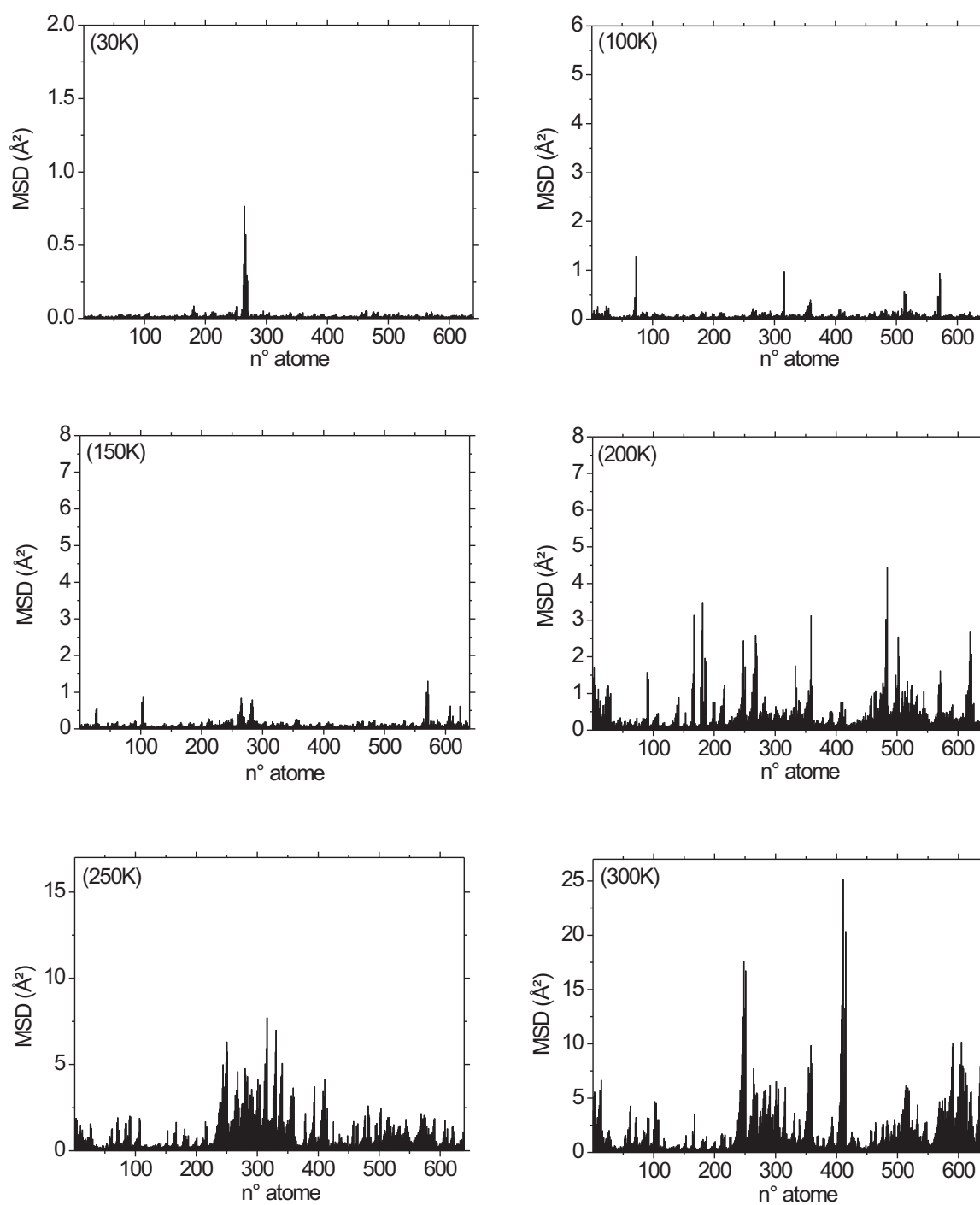


FIG. P.3 – Variation de la mobilité des atomes de la crambine S22/L25 hydratée en fonction de la température, sur une trajectoire de 1ns.

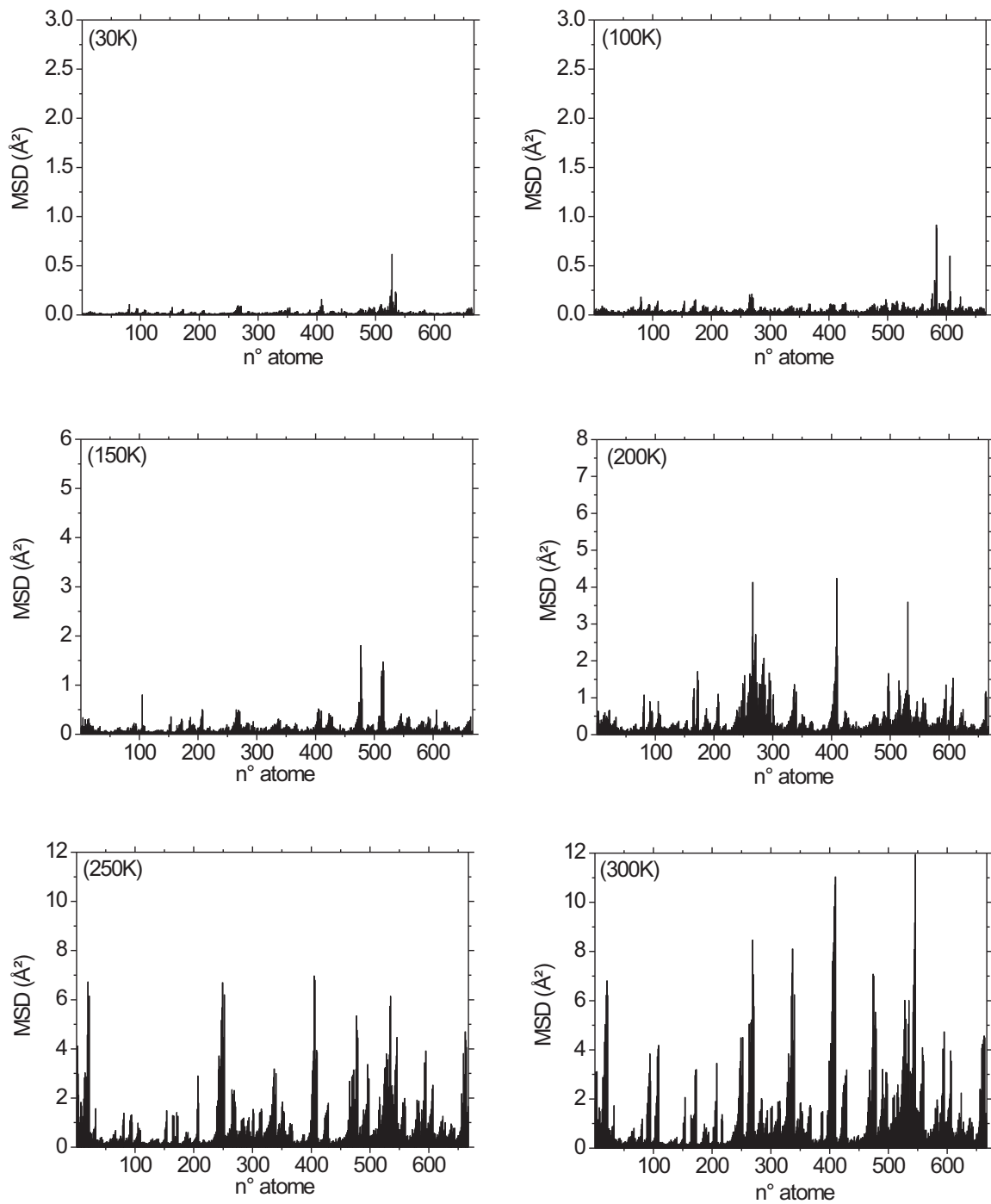


FIG. P.4 – Variation de la mobilité des atomes de la viscotoxine A3 hydratée en fonction de la température, sur une trajectoire de 1ns.

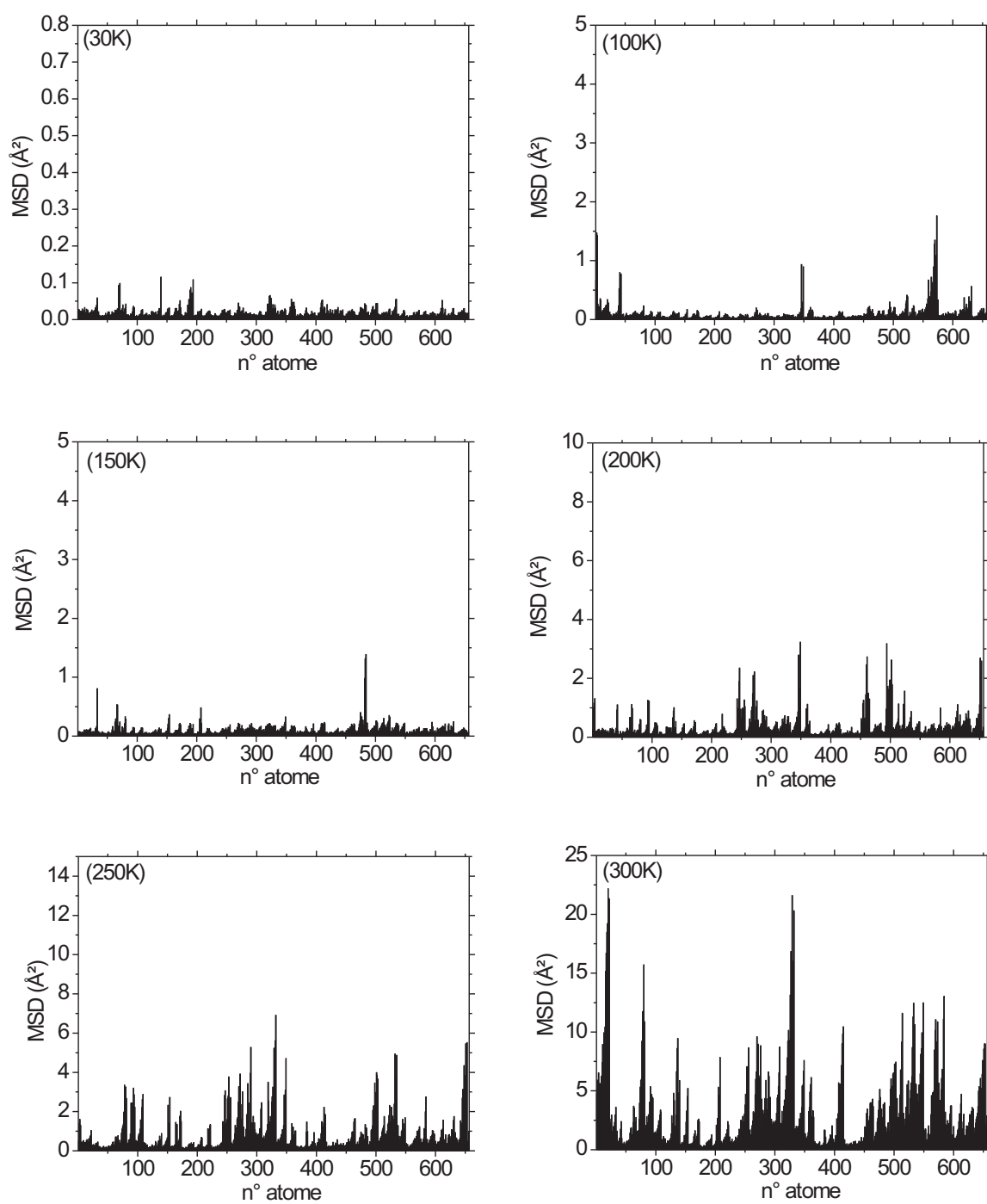


FIG. P.5 – Variation de la mobilité des atomes de la viscotoxine A2 hydratée en fonction de la température, sur une trajectoire de 1ns.

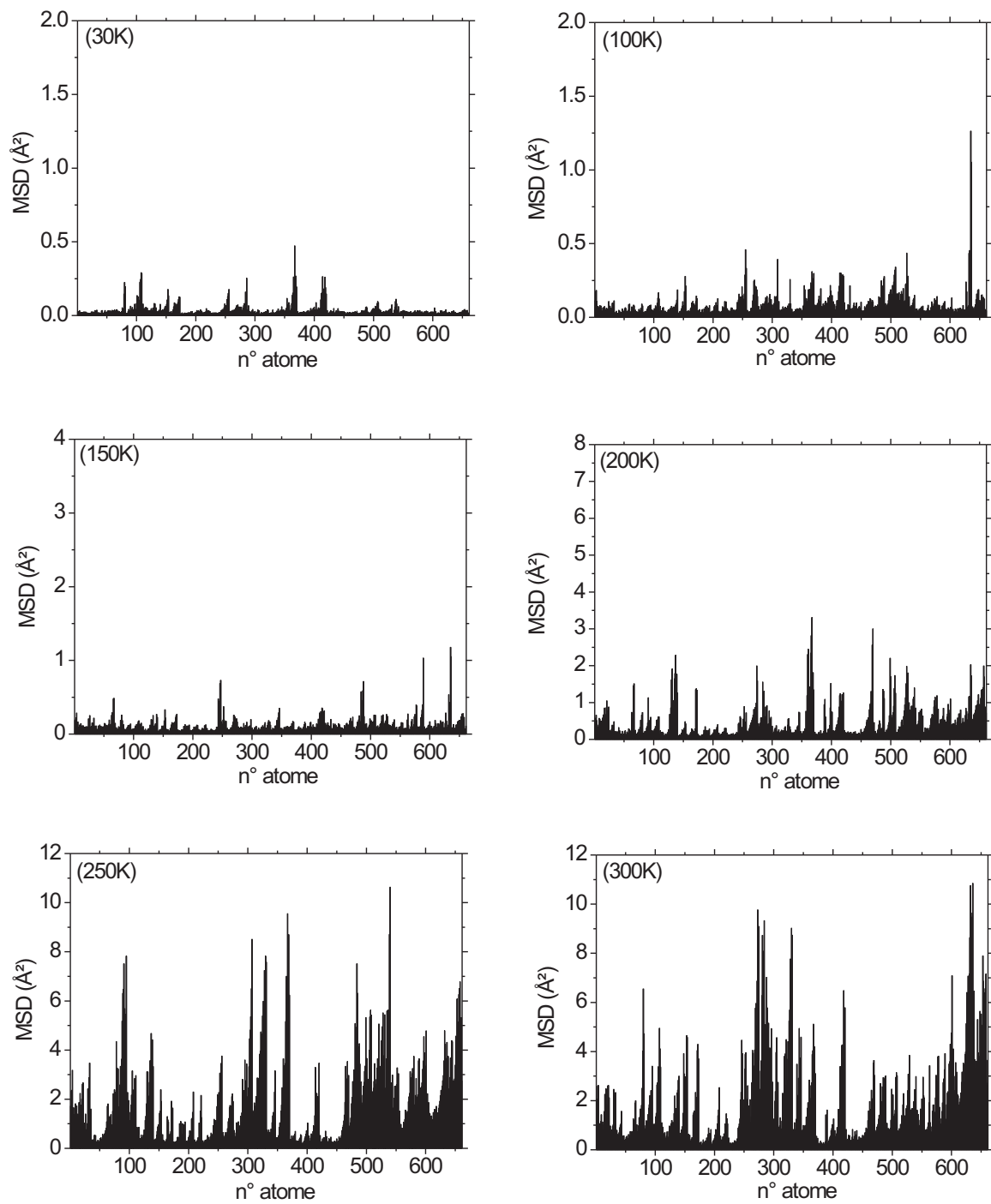


FIG. P.6 – Variation de la mobilité des atomes de la viscotoxine B hydratée en fonction de la température, sur une trajectoire de 1ns.

**Dépendance de la mobilité avec le
temps de production des protéines
hydratées**

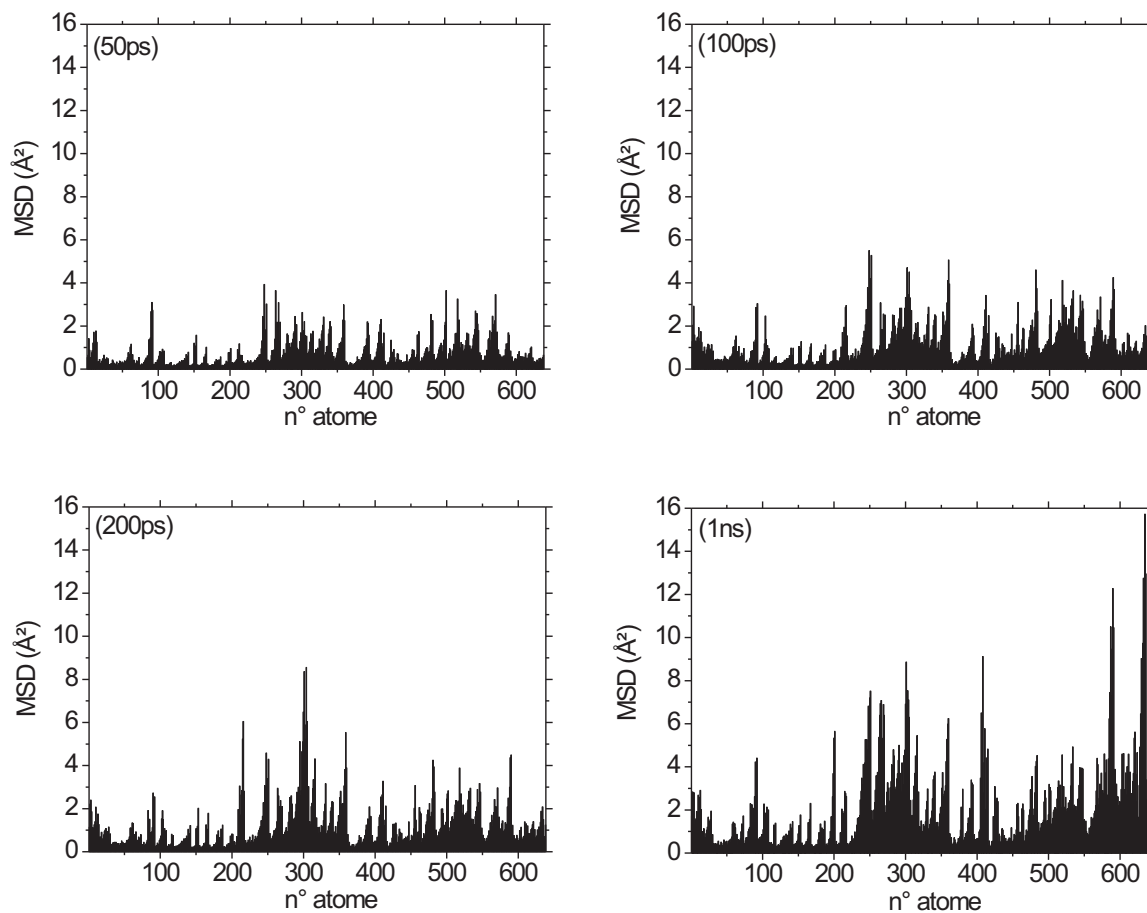


FIG. Q.1 – Variation de la mobilité des atomes de la crambine S22/I25 en fonction du temps de production, à 300K.

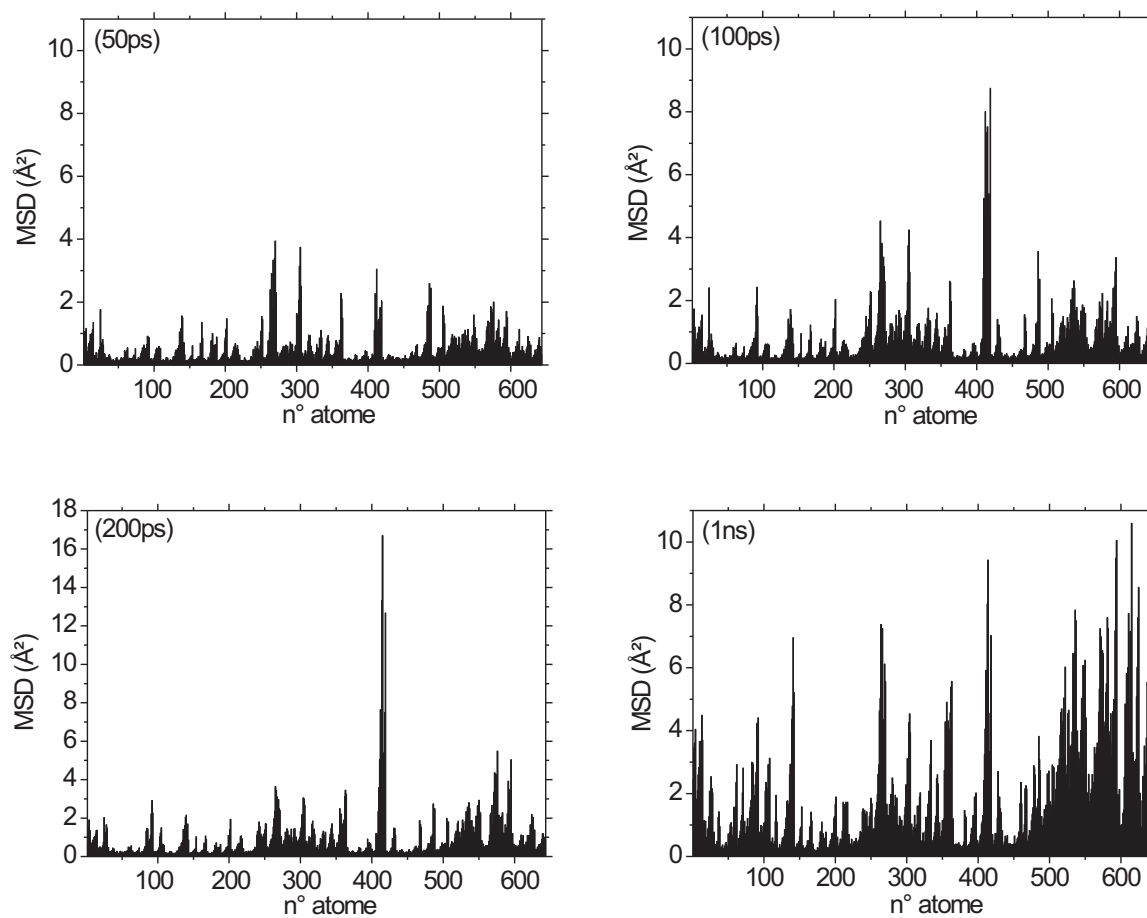


FIG. Q.2 – Variation de la mobilité des atomes de la crambine P22/I25 en fonction du temps de production, à 300K.

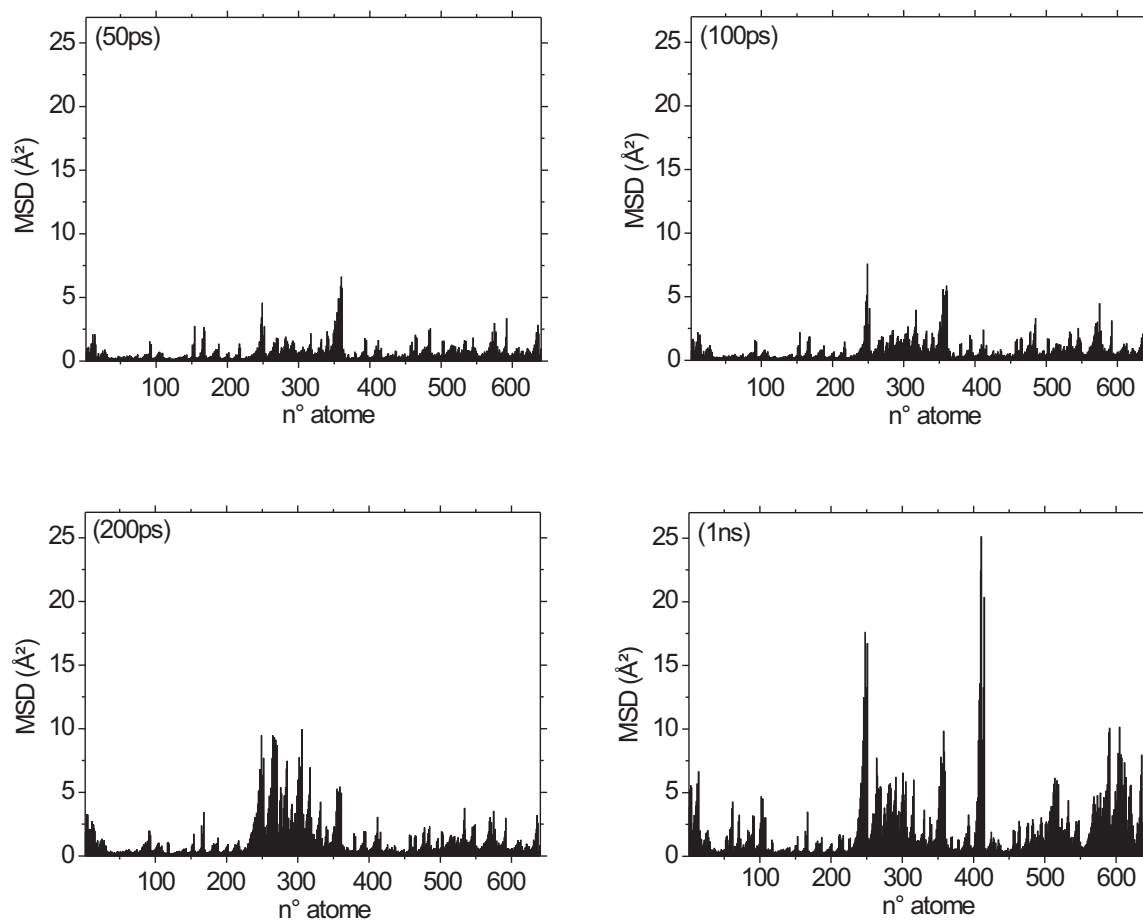


FIG. Q.3 – Variation de la mobilité des atomes de la crambine S22/L25 en fonction du temps de production, à 300K.

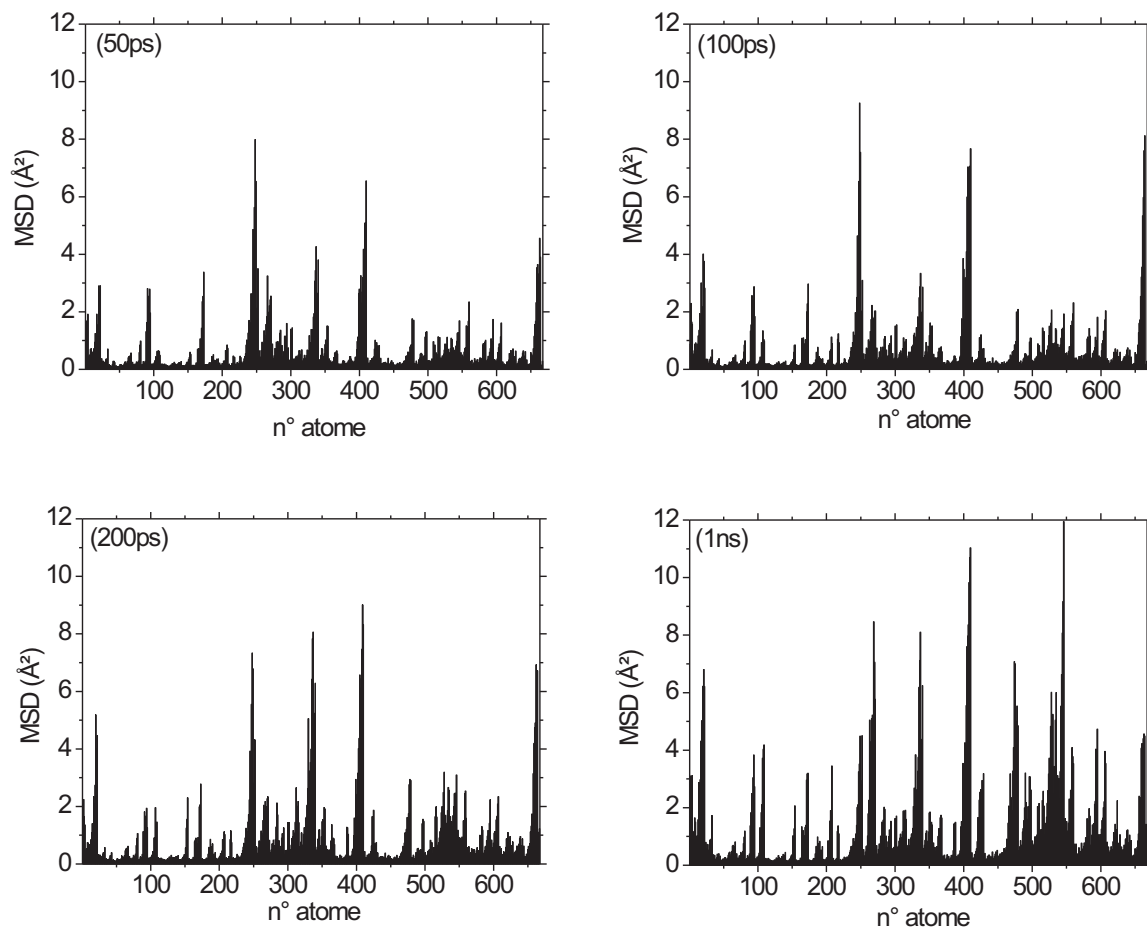


FIG. Q.4 – Variation de la mobilité des atomes de la viscotoxine A3 en fonction du temps de production, à 300K.

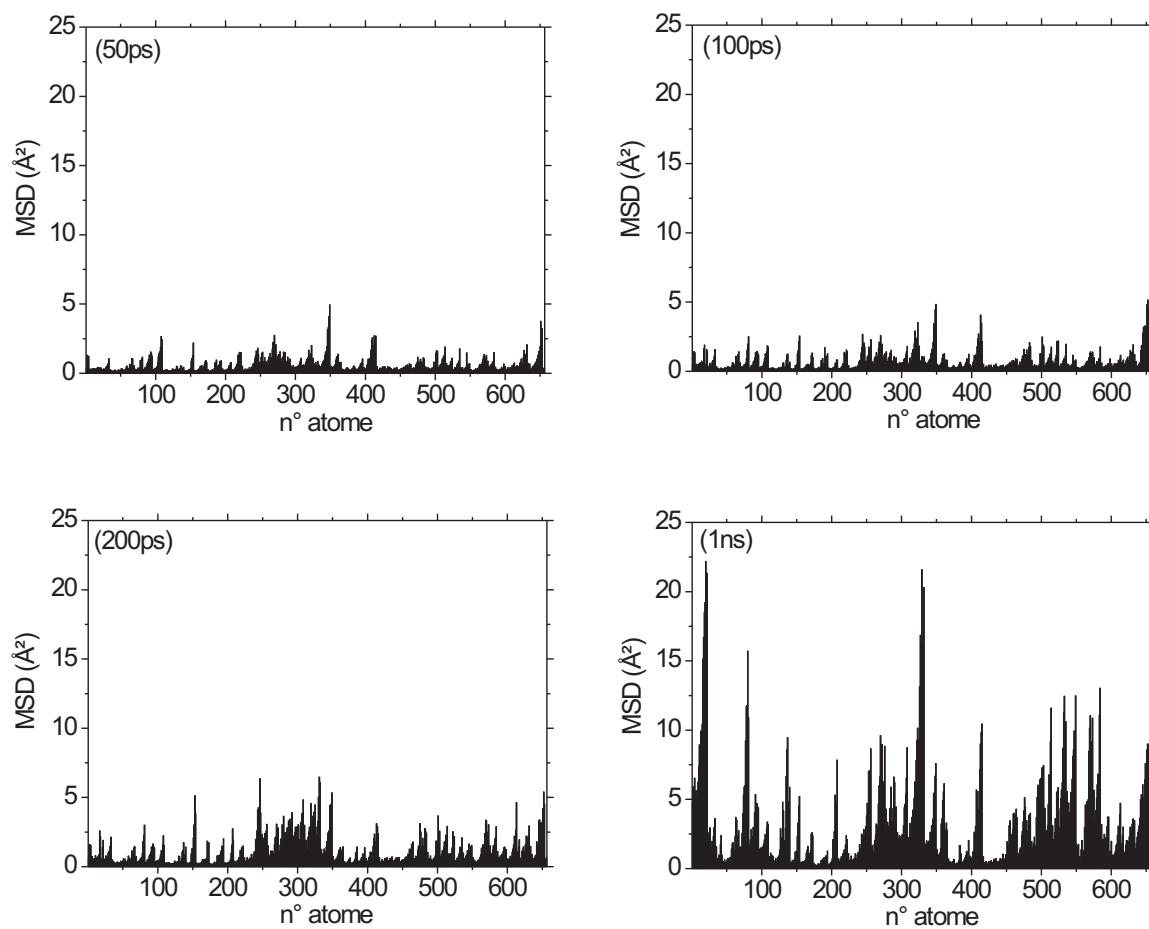


FIG. Q.5 – Variation de la mobilité des atomes de la viscotoxine A2 en fonction du temps de production, à 300K.

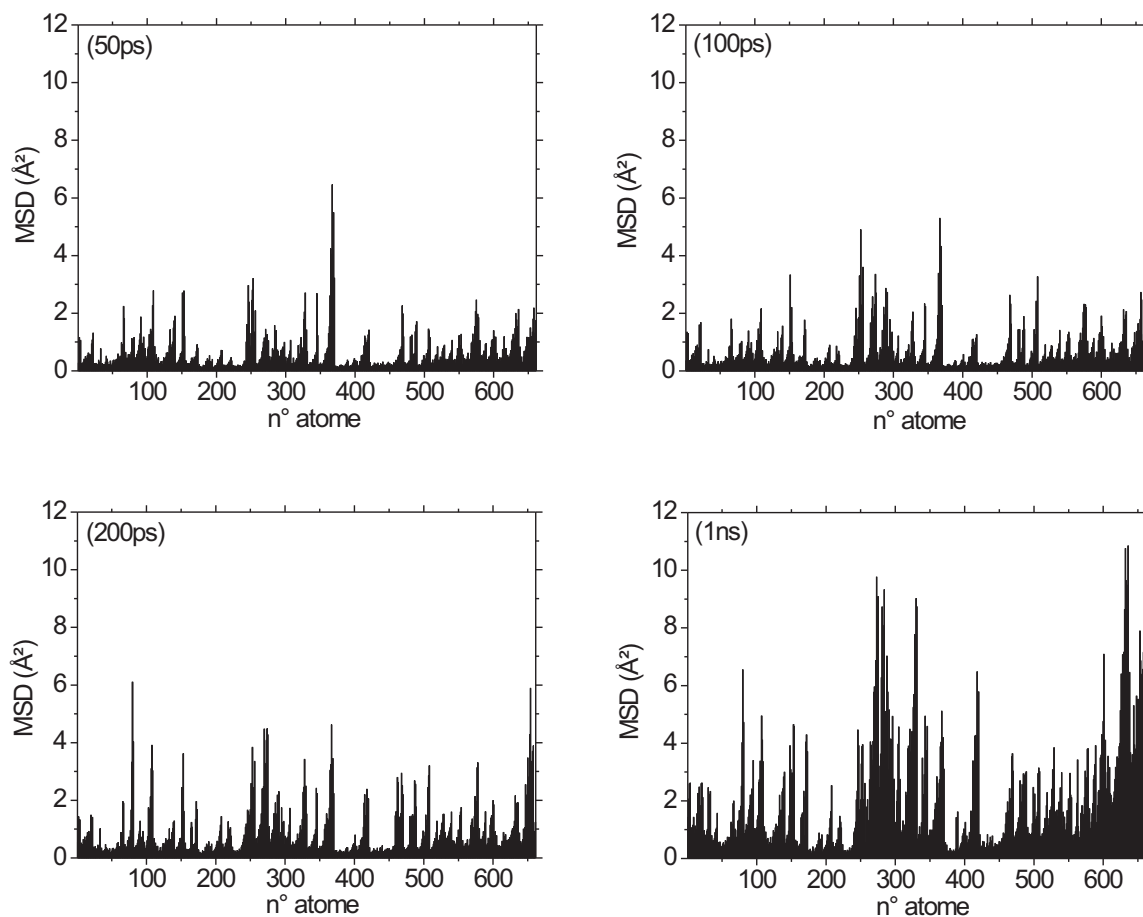


FIG. Q.6 – Variation de la mobilité des atomes de la viscotoxine B en fonction du temps de production, à 300K.

Mobilité des résidus des protéines hydratées

Dans le cas des protéines hydratées, les histogrammes représentés sur la Fig. R.1 nous indiquent que la transition dynamique débute vers 200/220K (histogrammes marron et rose respectivement). En effet, c'est à ces températures que les valeurs du $\frac{MSD}{T}$ des résidus augmentent.

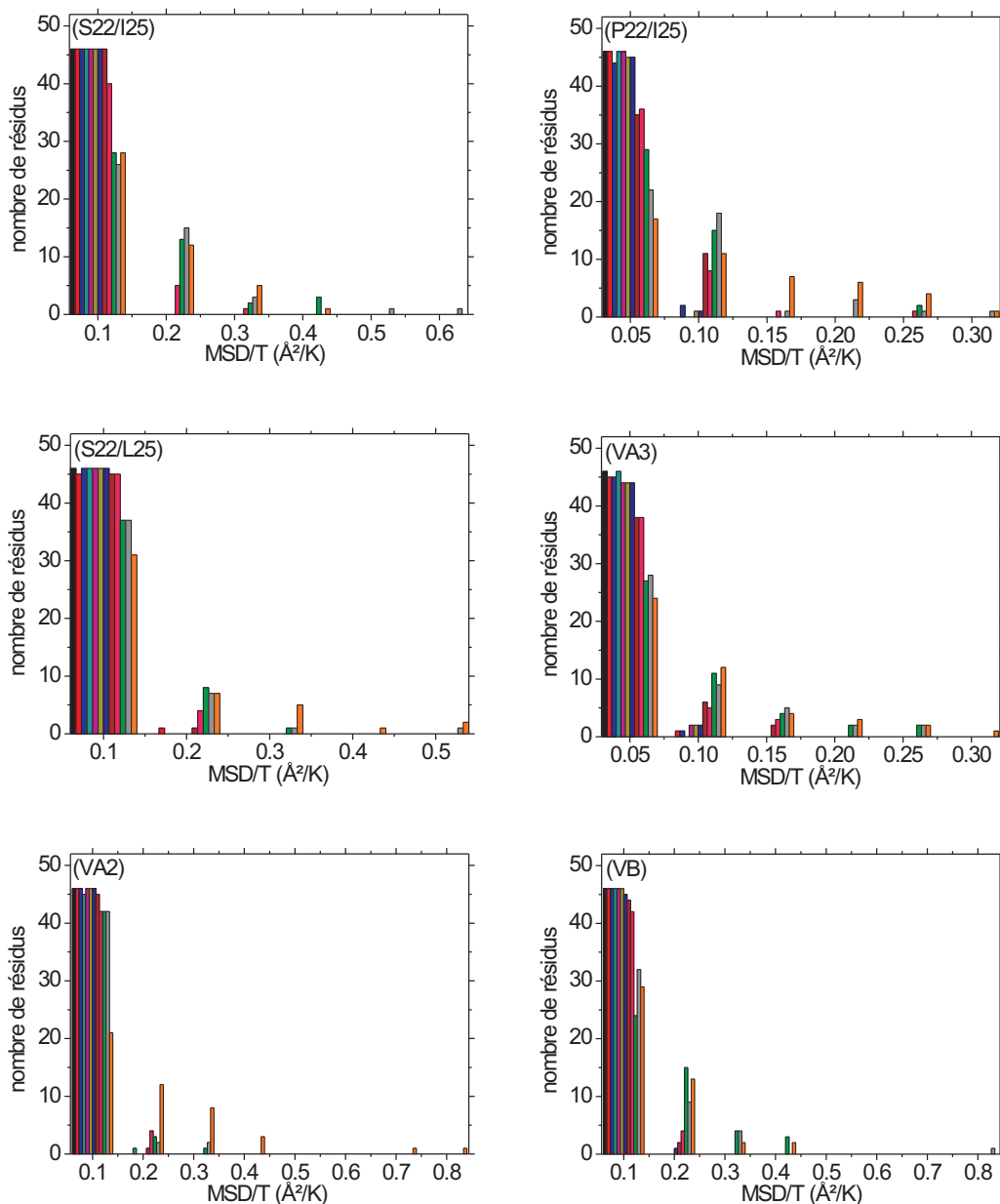


FIG. R.1 – Variation du nombre de résidus suivant leur valeur du $\frac{MSD}{T}$ pour les crambines S22/I25, P22/I25, S22/L25 et les 3 viscotoxines hydratées, sur une trajectoire de 1ns. Les histogrammes correspondent aux températures suivantes : noir = 5K, rouge = 30K, bleu = 70K, bleu-vert = 100K, fushia = 130K, kaki = 150K, bleu marine = 170K, marron = 200K, rose = 220K, vert = 250K, gris = 280K, orange = 300K.

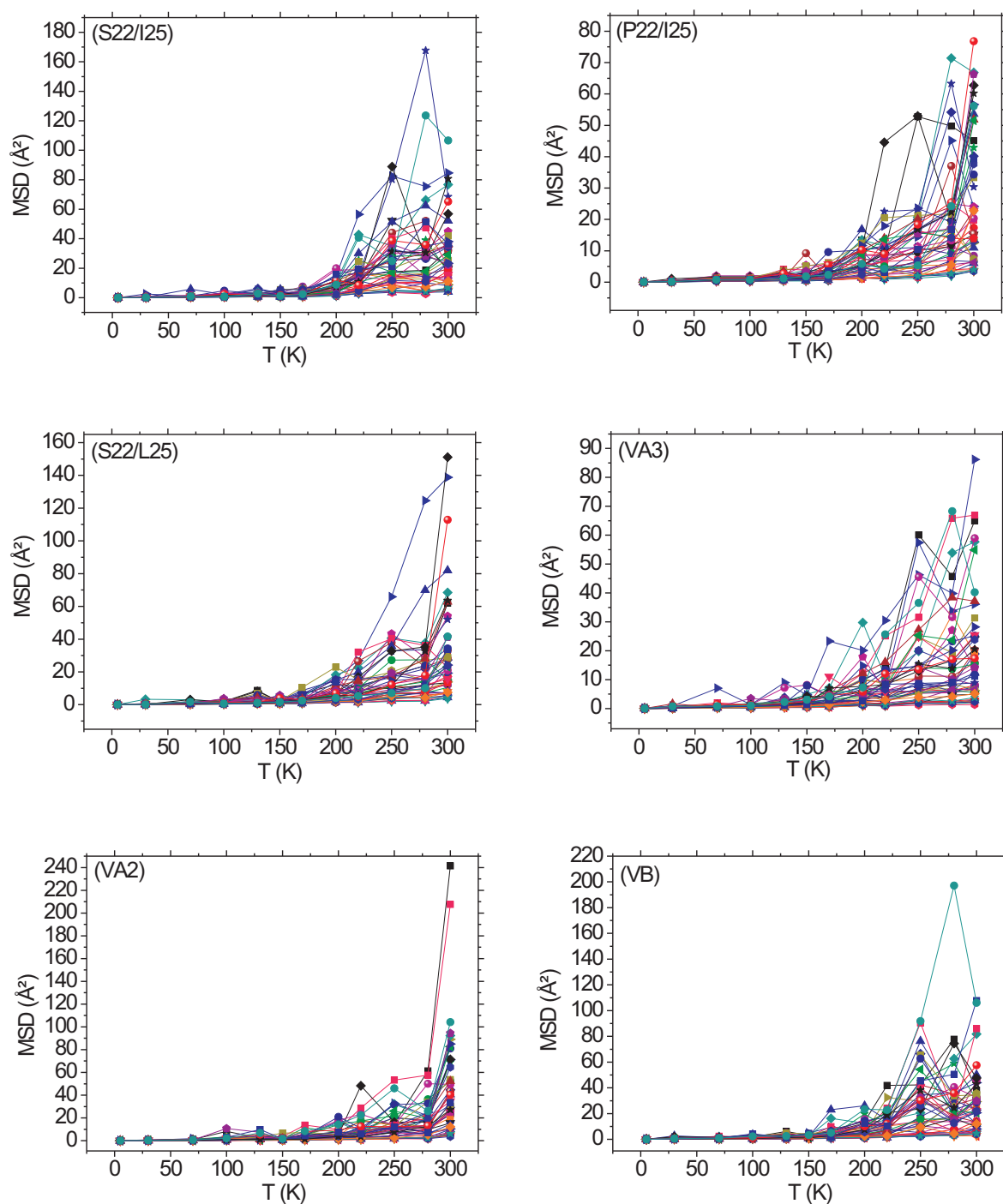


FIG. R.2 – Variation du MSD de chaque acide aminé avec la température pour les autres thionines hydratées, sur une trajectoire de 1ns.

Répartition de l'eau autour des protéines

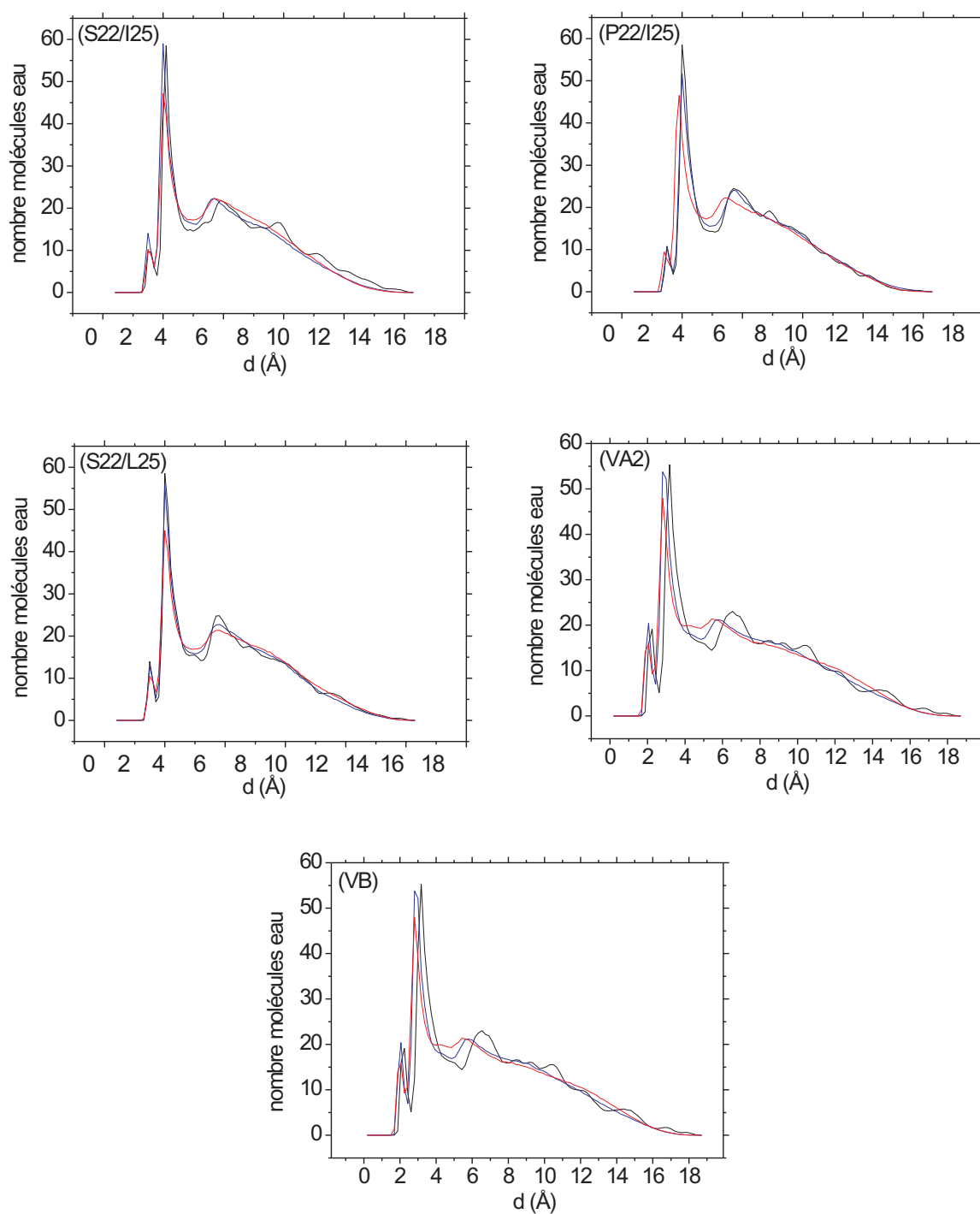


FIG. S.1 – Somme des distributions de probabilité de présence des molécules d'eau pour les protéines hydratées, à 150K (courbe noire), 220K (courbe bleue) et 300K (courbe rouge).

Programmes Fortran

T.1 Représentation des hydrogènes les plus mobiles par des sphères

Nous avons représenté les hydrogènes les plus mobiles par des sphères de rayon proportionnel au MSD par le programme fortran suivant :

```

c
c   Program with transforms values of a temperature field
c   in spheres of increasing radius for RASMOL representation
c   Linear interpolation between a minimum and maximum value
c   Mininum radius=0 Angstroem
c   Maximum radius=2.0 Angstroem (500 units)
c   A rasmol unit=0.004 angstroem
c   Output is a rasmol script : radius.spt
c
implicit none
integer*4 nmax,nlmax,nscale
parameter (nmax=4000,nlmax=30000,nscale=3)
integer*4 atpdb(nmax),tab(2,nmax),rad
integer*4 i,j,k,l,crd,model,nl,nmod,numero,isymb
integer*4 ioutpdb,ioutxyz,iout,nbegin,nstop,nid
integer*4 nsite,crdmodel,seqres(nmax),atomq(nmax)
integer*4 ter(100),iter
real*8 t(nmax,3),atomoccup(nmax),atomtemp(nmax),rms(2,nmax)
character*1 atomchain(nmax)
character*2 atomsymb(nmax)
character*3 res(nmax)
character*4 pdbcode,atomcode(nmax),segment(nmax),modelid
character*6 pdbrecord,pdbcrd(nmax)
character*10 label(nmax)
character*18 filepdb
character*30 nfile,nout,nfile2
character*80 title,lineb(nlmax),linef(nlmax),chartemp,charter(100)
c
c *** PURPOSE
c
write(6,*) '*** Write a rasmol script for CHARMM pdb file '
write(6,*) '*** Only a single chain protein treated'
write(6,*) '*** Multiplied by a factor scale :',scale
write(6,*)
write(6,*)
c
c *** PDB filename
c
write(6,*) '*** Type the filename of the PDB (CHARMM) file'
read(5,*) filepdb
c

```



```

c *** OPEN THE SCRIPT output file
c
  open(unit=21,file='radius.spt')
  write(21,*) "# Compute from charmm_temp_rasm for the data in"
  write(21,*) "#",filepdb
  write(21,*) "select all"
  write(21,*) "ssbonds off"
  write(21,*) "dots off"
  write(21,*) "background white"
  write(21,*) "select all"
  write(21,*) "color blue"
  write(21,*) "spacefill"
c
c *** PDB
c
  open(unit=1,file=filepdb)
  read(1,'(A80)') chartemp
  lineb(1)=chartemp
  read(1,'(A80)') chartemp
  lineb(2)=chartemp
  write(6,*) lineb(1)
  write(6,*) lineb(2)
  nl=2
c *** Read coordinates
  iter=0
  nsite=0
  nmod=0
  crd=1
  do while (crd > -1)
c *** Read the data for the file with one structure
  nsite=nsite+1
  nl=nl+1
  if (nsite.gt.nmax) stop
  read(1,'(A6,i5,1X,A4,1X,A3,1X,A1,i4,4X,3F8.3,2F6.2,6X,A4,A2,i2)')
    pdbcrd(nsite),atpdb(nsite),atomcode(nsite),
    res(nsite),atomchain(nsite),seqres(nsite),
    (t(nsite,k),k=1,3),atomoccup(nsite),atomtemp(nsite),
    segment(nsite),atomsymb(nsite),atomq(nsite)
  write(6,'(A6,i5,1X,A4,1X,A3,1X,A1,i4,4X,3F8.3,2F6.2,6X,A4,A2,i2)')
    pdbcrd(nsite),atpdb(nsite),atomcode(nsite),
    res(nsite),atomchain(nsite),seqres(nsite),
    (t(nsite,k),k=1,3),atomoccup(nsite),atomtemp(nsite),
    segment(nsite),atomsymb(nsite),atomq(nsite)
  if (pdbcrd(nsite).eq.'END ') then
  crd=-1
  nsite=nsite-1
  else

```

```

rad=scale*idnint(atomtemp(nsite))
write(21,'(1X,A15,i4)') "select atomno==" ,nsite
write(21,'(1X,A10,i4)') "spacefill ",rad
if (rad.gt.500) then
write(6,*) 'Maximum radius for atom ',nsite
pause
endif
endif
enddo
close(unit=1)
c
c *** End of the pdb file
c
close(unit=21)
stop
end

```

T.2 Calcul du rayon de giration le long d'une trajectoire

```

implicit none
integer*4 nmax,nfmax
parameter(nmax=1000,nfmax=10000)
character*4 typ(nmax)
character*3 residu(nmax)
integer*4 nresidu(nmax),natom,nframe,id,ncarbon,noxy,nhyd,nsulf,nnit
integer*4 idh,niso
real*8 r(nmax+2,3),wi(nmax),rprim2,rgir2,rprim(3)
real*8 rgiravg2,rgiravg,rprim2m
real*8 rgir2m,rgiravg2m,rgiravgm
real*8 rprimavg(nmax,3),rprimavgm(nmax,3)
real*8 rgirmean2,rgirmean2m,rgirmean,rgirmeanm,msd,msdmass
real*8 rgirexp2,rgirexp2m
real*8 rgirexpavg2,rgirexpavg2m,rgirexpavg,rgirexpavgm
character*1 nothing,isomer
character*2 nom,chain
character*64 filename
character*6 pdbid
character*4 flag,typtemp
character*12 mol
real*8 mass,occ,bfac(nmax),massprot,bfactemp
real*8 rtemp(3),pi
integer*4 i,j,k,l,num,nresidutemp
c
pi=4.0d0*datan(1.0d0)
c

```

```
c Output a "clean" pdb file (for isomers and experimental data)
c
open(unit=20,file='clean.pdb')
open(unit=10,file='resultsrgyr.dat')
write(6,*) '*** Type the filename (with .pdb) ***'
read(5,'(A64)') filename
c
c choose a isomer
c
write (6,*) '*** choose a isomer A, B or C or type Z (for all)'
read(5,*) isomer
open(unit=1,file=filename)
read(1,*)
read(1,*) natom,nframe
write(6,*) natom,nframe
if (natom.gt.nmax.or.nframe.gt.nfmax) then
write(6,*) '*** Nothing done, stop ***'
pause
stop
endif
rgiravg2=0.0d0
rgiravg2m=0.0d0
rgirexpavg2=0.0d0
rgirexpavg2m=0.0d0
c
c initialisation of rprimaverage
c
do i=1,natom
do k=1,3
rprimavg(i,k)=0.0d0
rprimavgm(i,k)=0.0d0
enddo
enddo
c
c loop on frames
c
do i=1,nframe
c initialisation of center of mass
r(natom+1,1)=0.0d0
r(natom+1,2)=0.0d0
r(natom+1,3)=0.0d0
r(natom+2,1)=0.0d0
r(natom+2,2)=0.0d0
r(natom+2,3)=0.0d0
c initialisation of mass protein
if (i.eq.1) then
massprot=0.0d0
```

```

ncarbon=0
noxy=0
nmit=0
nhyd=0
nsulf=0
endif
c
niso=0
do j=1,natom
read(1,'(A6,i5,1X,A4,A1,A3,1X,A2,i4,3x,3f8.3,2f6.2,A12)')
    & pdbid,num,typtemp,nothing,residu(j),
    & chain,nresidutemp,(rtemp(k),k=1,3),occ,bfactemp,mol
if (nothing.eq.isomer.or.nothing.eq.' '.or.isomer.eq.'Z') then
c weight wi computed for frame 1
niso=niso+1
if (nothing.ne.'Z') then
c output a clean pdb file = one isomer with atoms numbered correctly
write(20,'(A6,i5,1X,A4,A1,A3,1X,A2,i4,3x,3f8.3,2f6.2,A12)')
    & pdbid,niso,typtemp,nothing,residu(j),
    & chain,nresidutemp,(rtemp(k),k=1,3),occ,bfactemp,mol
endif
c
typ(niso)=typtemp
nresidu(niso)=nresidutemp
r(niso,1)=rtemp(1)
r(niso,2)=rtemp(2)
r(niso,3)=rtemp(3)
bfac(niso)=bfactemp
if (i.eq.1) then
nom=typ(niso)(1 :2)
    id=0
    if (nom(1 :1).eq.'C'.or.nom(2 :2).eq.'C') then
        mass=12.011d0
        id=1
        ncarbon=ncarbon+1
    endif
    if (nom(1 :1).eq.'O'.or.nom(2 :2).eq.'O') then
        mass=15.9994d0
        id=1
        noxy=noxy+1
    endif
    if (nom(1 :1).eq.'N'.or.nom(2 :2).eq.'N') then
        mass=14.0067d0
        id=1
        nmit=nmit+1
    endif
    if (nom(1 :1).eq.'T'.or.nom(2 :2).eq.'T') then

```

```
    mass=126.9045d0
    id=1
    nnit=nnit+1
    endif
    if (nom(1 :1).eq.'H'.or.nom(2 :2).eq.'H') then
    mass=1.00794d0
c Hydrogen mass
    if (1.eq.1) mass=0.0d0
    id=1
    nhyd=nhyd+1
    endif
    if (nom(1 :1).eq.'S'.or.nom(2 :2).eq.'S') then
    mass=32.06d0
    id=1
    nsulf=nsulf+1
    endif
    if (id.eq.0) then
c particular label for hydrogen atom (third position)
    nom=typ(niso)(2 :3)
    idh=0
    if(nom(2 :2).eq.'H') then
    idh=1
    mass=1.00794d0
c Hydrogen mass
    if (1.eq.1) mass=0.0d0
    nhyd=nhyd+1
    endif
    if (idh.eq.0) then
    write(6,*) '*** UNDEFINED ATOM TYPE, the program STOP'
    pause
    stop
    endif
    endif
    wi(niso)=mass
    endif
c    compute center of mass of frame i
do l=1,3
    r(natom+1,l)=r(natom+1,l)+r(niso,l)
    r(natom+2,l)=r(natom+2,l)+wi(niso)*r(niso,l)
enddo
c compute mass protein for frame 1
if (i.eq.1) then
massprot=massprot+mass
endif
c
endif
enddo
```

```

read(1,'(A3)') flag
if (flag.ne.'END') then
write(6,*) '*** Incorrect format, stop ***'
stop
endif
c
c end of the pdb output
c
close(unit=20)
c compute the square of these positions and sum up
rgir2=0.0d0
rgir2m=0.0d0
rgirexp2=0.0d0
rgirexp2m=0.0d0
do j=1,niso
rprim2=0.0d0
rprim2m=0.0d0
do k=1,3
rprim(k)=r(j,k)-r(natom+1,k)/niso
rprimavg(j,k)=rprimavg(j,k)+rprim(k)
rprim2=rprim2+rprim(k)*rprim(k)
rprim(k)=r(j,k)-r(natom+2,k)/massprot
rprimavgm(j,k)=rprimavgm(j,k)+rprim(k)
rprim2m=rprim2m+rprim(k)*rprim(k)
enddo
rgir2=rgir2+rprim2/niso
rgir2m=rgir2m+wi(j)*rprim2m/massprot
rgirexp2=rgirexp2+bfac(j)/8.0d0/pi**2/niso
rgirexp2m=rgirexp2m+wi(j)*(bfac(j)/8.0d0/pi**2)/massprot
enddo
c
c print the giration radius of frame i
c
rgiravg2=rgiravg2+rgir2
rgirexpavg2=rgirexpavg2+rgirexp2
rgiravg2m=rgiravg2m+rgir2m
rgirexpavg2m=rgirexpavg2m+rgirexp2m
enddo
c
c Radius computed from the mean rprim
c
rgirmean2=0.0d0
rgirmean2m=0.0d0
do j=1,niso
rgirmean2=rgirmean2+((rprimavg(j,1)**2+rprimavg(j,2)**2+rprimavg(j,3
&)**2)/nframe**2)/niso
rgirmean2m=rgirmean2m+wi(j)*((rprimavgm(j,1)**2+rprimavgm(j,2)**2+rprimavg

```

```

      &m(j,3)**2)/nframe**2)/massprot
    enddo
    rgirmean=dsqrt(rgirmean2)
    write(6,'(A24,f12.6)') 'RGYR (MEAN)= ',rgirmean
    rgirmeanm=dsqrt(rgirmean2m)
    write(6,'(A24,f12.6)') 'RGYR (MEAN, MASS)= ',rgirmeanm
    pause
    c
    c Radius of giration
    c
    rgiravg2=rgiravg2/nframe
    rgiravg=dsqrt(rgiravg2)
    rgirexpavg2=rgirexpavg2/nframe
    rgirexpavg2=rgirexpavg2+rgirmean2
    rgirexpavg=dsqrt(rgirexpavg2)
    if (isomer.ne.'Z') then
    write(6,*) 'Number of atoms in isomer',isomer,niso
    endif
    write(6,'(A6,f12.6)') 'RGYR= ',rgiravg
    rgiravg2m=rgiravg2m/nframe
    rgiravgm=dsqrt(rgiravg2m)
    rgirexpavg2m=rgirexpavg2m/nframe
    rgirexpavg2m=rgirexpavg2m+rgirmean2m
    rgirexpavgm=dsqrt(rgirexpavg2m)
    write(6,'(A6,f12.6)') 'RGYR(MASS)= ',rgiravgm
    write(6,*) '** PROTEIN MASS =',massprot
    write(6,*) 'test : carbon/oxy/nit/nhyd/nsulf', ncarbon,noxy,nnit,nhyd,nsulf
    c
    c exp
    c
    write(6,*) ' Computed from B factors '
    write(6,*) ' Rgyr (equal weights) ', rgirexpavg
    write(6,*) ' Rgyr (Mass) ', rgirexpavgm
    pause
    c
    c MSD
    c
    msd=rgiravg2-rgirmean2
    msdmass=rgiravg2m-rgirmean2m
    write(6,*) '*** MSD = ', msd
    write(6,*) '*** MSD (MASS)',msdmass
    write(10,'(A24,f12.6)') 'RGYR (MEAN)= ',rgirmean
    write(10,'(A24,f12.6)') 'RGYR (MEAN, MASS)= ',rgirmeanm
    write(10,'(A24,f12.6)') 'RGYR= ',rgiravg
    write(10,'(A24,f12.6)') 'RGYR(MASS)= ',rgiravgm
    write(10,*) '** PROTEIN MASS =',massprot
    write(10,*) 'test : carbon/oxy/nit/nhyd/nsulf', ncarbon,noxy,nnit,nhyd,nsulf

```

```
write(10,*) 'Number of atoms in isomer',isomer,niso
write(10,*) '*** MSD = ', msd
write(10,*) '*** MSD (MASS)',msdmass
write(10,*) ' Computed from B factors '
write(10,*) ' Rgyr (equal weights) ', rgirexpavg
write(10,*) ' Rgyr (Mass) ', rgirexpavgm
close(unit=1)
stop
end
```

Bibliographie

- [Abraham 86] Abraham (F. F.), *Computational Statistical Mechanics. Methodology, Applications and Supercomputing*, Adv. Phys. **35** (1986), 1 :111.
- [Ahlström *et al.* 87] Ahlström (P.), Teleman (O.), Jönsson (B.), et Forsén (S.), *Molecular dynamics simulation of parvalbumin in aqueous solution*, J. Am. Chem. Soc. **109** (1987), 1541 :1551.
- [Alder et Wainwright 59] Alder (B. J.) et Wainwright (T. E.), *Studies in Molecular Dynamics. I. General Method*, J. Chem. Phys. **31** (1959), no. 2, 459.
- [Allinger *et al.* 89] Allinger (N. L.), Yuh (Y. H.), et Lii (J. H.), *Molecular mechanics. The MM3 force field for hydrocarbons*, J. Am. Chem. Soc. **111** (1989), no. 23, 8551 :8566.
- [Allinger *et al.* 96] Allinger (N. L.), Chen (K.), et Lii (J. H.), *An Improved Force Field (MM4) for Saturated Hydrocarbons*, J. Comp. Chem. **17** (1996), 642 :668.
- [Allinger et Sprague 73] Allinger (N. L.) et Sprague (J. T.), *Conformational analysis. XC. Calculation of the structures of hydrocarbons containing delocalized electronic systems by the molecular mechanics method*, J. Am. Chem. Soc. **95** (1973), 3893 :3907.
- [Allinger 77] Allinger (N. L.), *Conformational Analysis 130. MM2. A Hydrocarbon Force Field Utilizing V1 and V2 Torsional Terms*, Advan. Phys. Organ. Chem. **99** (1977), 8127 :8134.
- [Andersen 80] Andersen (H. C.), *Molecular-dynamics simulations at constant pressure and/or temperature*, J. Chem. Phys. **72** (1980), 2384 :2393.
- [Aparicio 05] Aparicio (F.), conférence privée, 2005.
- [Arrhenius 89] Arrhenius (S.), *On the Reaction Velocity of the Inversion of Cane Sugar by Acids*, Zeitschrift für Physikalische Chemie **4** (1889), 226 :232.
- [Austin *et al.* 75] Austin (R. H.), Beeson (K. W.), Eisenstein (L.), Frauenfelder (H.), et Gunsalus (I. C.), *Dynamics of ligand binding to myoglobin*, Biochemistry **14** (1975), no. 24, 5355 :5373.

- [Becke 93] Becke (A. D.), *Density-functional thermochemistry. III. The role of exact exchange*, J. Chem. Phys. **98** (1993), 5648 :5652.
- [Bekker *et al.* 93] Bekker (H.), Berendsen (H. J. C.), Dijkstra (E. J.), Achterop (S.), Drunen (R. Van), Spoel (D. Van Der), Sijbers (A.), Keegstra (H.), Reitsma (B.), et Renardus (M. K. R.), *Gromacs : A parallel computer for molecular dynamics simulations.*, Physics Computing 92 (1993), 252 :256.
- [Bellissent-Funel 04] Bellissent-Funel (M. C.), *Hydrophilic-hydrophobic interplay : from model systems to living systems*, C. R. Géoscience **337** (2004), 173 :179.
- [Bicout et Zaccai 01] Bicout (D. J.) et Zaccai (G.), *Protein flexibility from the dynamical transition : a force constant analysis*, Biophys J. **180** (2001), no. 3, 1115 :1123.
- [Bicout 05] Bicout (D. J.), *Diffusion incohérente des neutrons : modèles analytiques pour la dynamique interne des protéines*, J. phys., IV **130** (2005), 115 :132.
- [Bonvin *et al.* 93] Bonvin (A. M.), Rullmann (J. A.), Lamerichs (R. M.), Boelens (R.), et Kaptein (R.), *"Ensemble" iterative relaxation matrix approach : A new NMR refinement protocol applied to the solution structure of crambin*, Proteins **15** (1993), 385 :400.
- [Born et Oppenheimer 27] Born (M.) et Oppenheimer (R.), *Zur Quantentheorie der Molekeln*, Ann. Phys. **84** (1927), 457 :484.
- [Borwein et Borwein 87] Borwein (J. M.) et Borwein (P. B.), *PI and the AGM : A Study in Analytic Number Theory and Computational Complexity*, New York :Wiley, 1987.
- [Bouchaud et Georges 90] Bouchaud (J. P.) et Georges (A.), *Anomalous diffusion in disordered media : statistical mechanisms and physical applications*, Phys. Rep. **195** (1990), 127 :293.
- [Brooks *et al.* 83] Brooks (B. R.), Bruccoleri (R. E.), Olafson (B. D.), States (D. J.), Swaminathan (S.), et Karplus (M.), *CHARMM : A program for macromolecular energy, minimization, and dynamics calculations*, J. Comp. Chem. **4** (1983), 187 :217.
- [Brooks *et al.* 88] Brooks (C. L.), Karplus (M.), et Pettitt (B. M.), *Proteins : A theoretical perspective of dynamics, structure, and thermodynamics*, Wiley **71** (1988).
- [Caaveiro *et al.* 98] Caaveiro (J. M.), Molina (A.), Rodríguez-Palenzuela (P.), Goñi (F. M.), et González-Mañas (J. M.), *Interaction of wheat alpha-thionin with large unilamellar vesicles*, Protein Science **7** (1998), 2567 :2577.
- [Caliskan *et al.* 06] Caliskan (G.), Briber (R. M.), Thirumalai (D.), Garcia-Sakai (V.), Woodson (S. A.), et Sokolov (A. P.), *Dynamic transition in tRNA is solvent induced*, J. Am. Chem. Soc. **128** (2006), no. 1, 32 :33.

- [Callaway 74] Callaway (J.), *Quantum Theory of the Solid State*, Academic Press, New York, 1974.
- [Caves *et al.* 98] Caves (L. S.), Evanseck (J. D.), et Karplus (M.), *Locally accessible conformations of proteins : multiple molecular dynamics simulations of crambin*, *Protein Sci.* **7** (1998), no. 3, 649 :666.
- [Cheatham *et al.* 95] Cheatham (T. E.), Miller (J. H.), Fox (T.), Darden (P. A.), et Kollman (P. A.), *Molecular Dynamics Simulations on Solvated Biomolecular Systems : The Particle Mesh Ewald Method Leads to Stable Trajectories of DNA, RNA, and Proteins*, *J. Am. Chem. Soc.* **117** (1995), 4193 :4194.
- [Chong *et al.* 01] Chong (S. H.), Joti (Y.), Kidera (A.), Go (N.), Ostermann (A.), Gassmann (A.), et Parak (F. G.), *Dynamical transition of myoglobin in a crystal : comparative studies of x-ray crystallography and Mössbauer spectroscopy*, *Eur. Biophys. J.* **30** (2001), 319 :329.
- [Clark *et al.* 89] Clark (M.), Cramer (R. D.), et Opdenbosch (N. Van), *Validation of the General Purpose Tripos 5.2 Force Field*, *J. Comp. Chem.* **10** (1989), 982 :1012.
- [Coulon *et al.* 01] Coulon (A.), Berkane (E.), Sautereau (A. M.), Urech (K.), Roug a (P.), et Lopez (A.), *Modes of membrane interaction of a natural cysteine-rich peptide : viscotoxin A3*, *Biochimica et Biophysica Acta (BBA)* **1559** (2001), 145 :159.
- [Coulon *et al.* 02] Coulon (A.), Berkane (E.), Sautereau (A.), Urech (K.), Rouger (P.), et Lopez (A.), *Modes of membrane interaction of a natural cysteine-rich peptide : viscotoxin A3*, *Biochim. Biophys. Acta.* **1559** (2002), 145 :159.
- [Coulon *et al.* 03] Coulon (A.), Mosbah (A.), Lopez (A.), Sautereau (A. M.), Schaller (G.), Urech (K.), Rouge (P.), et Darbon (H.), *Comparative membrane interaction study of viscotoxins A3, A2 and B from mistletoe (*Viscum album*) and connections with their structures*, *Biochem. J.* **374** (2003), 71 :78.
- [Creighton 94] Creighton (T. E.), *Proteins, Structures and Molecular Properties*, 2  d., Freeman, New York, 1994.
- [Curtis *et al.* 04] Curtis (J. E.), Tarek (M.), et Tobias (D. J.), *Methyl group dynamics as a probe of the protein dynamical transition*, *J. Am. Chem. Soc.* **126** (2004), no. 49, 15928 :15929.
- [Dauber-Osguthorpe *et al.* 88] Dauber-Osguthorpe (P.), Roberts (V. A.), Osguthorpe (D. J.), Wolff (J.), Genest (M.), et Hagler (A. T.), *Structure and energetics of ligand binding to proteins : E. coli dihydrofolate reductase-trimethoprim, a drug-receptor system*, *Proteins* **4** (1988), 31 :47.
- [Debye 14] Debye (P.), *Interferenz von R ntgenstrahlen und W rmebewegung*, *Ann. Phys.* **348** (1914), 49 :92.

- [Diu *et al.* 89] Diu (B.), Guthmann (C.), Lederer (D.), et Roulet (B.), *Physique statistique*, Hermann, 1989.
- [Doster *et al.* 89] Doster (W.), Cusak (S.), et Petry (W.), *Dynamical transition of myoglobin revealed by inelastic neutron scattering*, *Nature* **337** (1989), 754 :756.
- [Doster et Settles 99] Doster (W.) et Settles (M.), *The Dynamical Transition In Proteins :the Role Of Hydrogen Bonds*, Workshop on Hydration Processes in Biology : Theoretical and Experimental Approaches, 1999.
- [Engler *et al.* 73] Engler (E. M.), Androse (J. D.), et Schleyer (P. R.), *J. Am. Chem. Soc.* **95** (1973), 8005 :8025.
- [Ermer et Lifson 73] Ermer (O.) et Lifson (S.), *Consistent force field calculations. III. Vibrations, conformations, and heats of hydrogenation of nonconjugated olefins*, *J. Am. Chem. Soc.* **95** (1973), 4121 :4132.
- [Evans et Holian 85] Evans (D. J.) et Holian (B. L.), *The Nose-Hoover thermostat*, *J. Chem. Phys.* **83** (1985), 4069 :4074.
- [Ewald 21] Ewald (P.), *Die Berechnung optischer und elektrostatischer Gitterpotentiale*, *Ann. Phys.* **64** (1921), 253 :287.
- [Eyal *et al.* 03] Eyal (E.), Najmanovich (R.), Edelman (M.), et Sobolev (V.), *Protein side chain rearrangement in regions of point mutations*, *Proteins* **50** (2003), 272 :282.
- [Feig *et al.* 04] Feig (M.), Karanicolas (J.), et Brooks (C. L.), *MMTSB Tool Set : enhanced sampling and multiscale modeling methods for applications in structural biology*, *Journal of Molecular Graphics and Modeling* **22** (2004), no. 5, 377 :395.
- [Feller *et al.* 95] Feller (S. E.), Zhang (Y.), Pastor (R. W.), et Brooks (B. R.), *Constant pressure molecular dynamics simulation : The Langevin piston method*, *J. Chem. Phys.* **103** (1995), no. 11, 4613 :4621.
- [Fenimore *et al.* 04] Fenimore (P. W.), Frauenfelder (H.), McMahon (B. H.), et Young (R. D.), *Bulk-solvent and hydration-shell fluctuations, similar to alpha- and beta-fluctuations in glasses, control protein motions and functions*, *Proc. Natl. Acad. Sci.* **101** (2004), 14408 :14413.
- [Fenimore *et al.* 05] Fenimore (P. W.), Frauenfelder (H.), McMahon (B. H.), et Young (R. D.), *Proteins are paradigms of stochastic complexity*, *Physica A* **351** (2005), 1 :13.
- [Ferrand *et al.* 93] Ferrand (M.), Dianoux (J.), Petry (W.), et Zaccai (G.), *Thermal motions and function of bacteriorhodopsin in purple membranes : effects of temperature and hydration studied by neutron scattering*, *Proc. Natl. Acad. Sci.* **90** (1993), 9668 :9672.

- [Fletcher et Reeves 64] Fletcher (R.) et Reeves (C. M.), *Function Minimization By Conjugate Gradient*, Computer Journal **7** (1964), 149 :154.
- [Flower 99] Flower (D. R.), *Rotational superposition : a review of methods*, J. Mol. Graph. Model **17** (1999), 238 :244.
- [Frauenfelder *et al.* 79] Frauenfelder (H.), Petsko (G. A.), et Tsernoglou (D.), *Temperature-Dependent X-ray Diffraction as a Probe of Protein Structural Dynamics*, Nature **280** (1979), 558 :563.
- [Frauenfelder *et al.* 91] Frauenfelder (H.), Sligar (S. G.), et Wolynes (P. G.), *The Energy Landscape and Motions of Proteins*, Science **254** (1991), 1598 :1603.
- [Frauenfelder et McMahon 98] Frauenfelder (H.) et McMahon (B.), *Dynamics and function of proteins : The search for general concepts*, Proc. Natl. Acad. Sci. **95** (1998), no. 9, 4795 :4797.
- [Frisch *et al.*] Frisch (M. J.), Trucks (G. W.), Schlegel (H. B.), Scuseria (G. E.), Robb (M. A.), Cheeseman (J. R.), et al. , *Gaussian 03, Revision C.02*, Gaussian, Inc., Wallingford, CT, 2004.
- [Fujita et Noda 78] Fujita (Y.) et Noda (Y.), *Effect of hydration on the thermal denaturation of lysozyme as measured by differential scanning calorimetry*, Bull. Chem. Soc. Jpn. **51** (1978), 1567 :1568.
- [Gabel et Bellissent-Funel 07] Gabel (F.) et Bellissent-Funel (M. C.), *C-Phycocyanin Hydration Water Dynamics in the Presence of Trehalose : An Incoherent Elastic Neutron Scattering Study at Different Energy Resolutions*, Biophysical Journal **92** (2007), 4054 :4063.
- [Gajewski *et al.* 90] Gajewski (J. J.), Gilbert (K. E.), et McKelvey (J.), *MMX an Enhanced Version of MM2*, Adv. in Molecular Modeling **2** (1990), 65 :92.
- [Giudici *et al.* 03] Giudici (M.), Pascual (R.), de la Canal (L.), Pf Fuller (K.), Pf Fuller (U.), et Villalain (J.), *Interaction of viscotoxins A3 and B with membrane model systems : implications to their mechanism of action*, Biophys J. **85** (2003), 971 :981.
- [Gregory et Chai 93] Gregory (R. B.) et Chai (K. J.), *A positron annihilation lifetime study of protein hydration : evidence for a glass transition*, vol. 3, International workshop on positron and positronium chemistry, no. 4, 1993, p. 305 :310.
- [Gregory 94] Gregory (R. B.), *Protein-solvent interactions*, Dekker, 1994.
- [Guex et Peitsch 97] Guex (N.) et Peitsch (M. C.), *SWISS-MODEL and the Swiss-PdbViewer : An environment for comparative protein modeling*, Electrophoresis **18** (1997), 2714 :2723.
- [Hagler *et al.* 74a] Hagler (A. T.), Huler (E.), et Lifson (S.), *Energy Functions for Peptides and Proteins : II. The Amide Hydrogen Bond and Calculation of Amide Crystal Properties*, J. Am. Chem. Soc. **96** (1974), 5327 :5335.
- [Hagler *et al.* 74b] Hagler (A.T.), Huler (E.), et Lifson (S.), *Energy Functions for Peptides and Proteins : I. Derivation of a Consistent Force*

- Field Including the Hydrogen Bond from Amide Crystals*, J. Am. Chem. Soc. **96** (1974), 5319 :5327.
- [Halgren 92] Halgren (T. A.), *The representation of van der Waals (vdW) interactions in molecular mechanics force fields : potential form, combination rules, and vdW parameters*, J. Am. Chem. Soc. **114** (1992), 7827 :7843.
- [Han et al. 96] Han (K. H.), Park (K. H.), Yoo (H. J.), Cha (H.), Suh (S. W.), Thomas (F.), Moon (T. S.), et Kim (S. M.), *Determination of the three-dimensional structure of hordothionin by nuclear magnetic resonance*, Biochem. J. **313** (1996), 885 :892.
- [Hautman et al. 89] Hautman (J.), Halley (J. W.), et Rhee (Y. J.), *Molecular dynamics simulation of water between two ideal classical metal walls*, J. Chem. Phys. **91** (1989), 467 :472.
- [Hayward et al. 03] Hayward (J. A.), Finney (J. L.), Daniel (R. M.), et Smith (J. C.), *Molecular Dynamics Decomposition of Temperature-Dependent Elastic Neutron Scattering by a Protein Solution*, Biophys J. **85** (2003), 679 :685.
- [Henzler-Wildman et al. 07] Henzler-Wildman (K. A.), Lei (M.), Thai (V.), Kerns (S. Jordan), Karplus (M.), et Kern (D.), *A hierarchy of timescales in protein dynamics is linked to enzyme catalysis*, Nature **450** (2007), 913 :916.
- [Hermans et al. 84] Hermans (J.), Berendsen (H. J. C.), Gunsteren (W. F. Van), et Postma (J. P. M.), *A Consistent Empirical Potential for Water-Protein Interactions*, Biopolymers **23** (1984), 1513 :1518.
- [Hu et al. 01] Hu (K.), Ivanov (P. Ch.), Chen (Z.), Carpena (P.), et Stanley (H. E.), *Effect of trends on detrended fluctuation analysis*, Phys. Rev. E **64** (2001), 011114.
- [Hughes et al. 00] Hughes (P.), Dennis (E.), Whitecross (M.), Llewellyn (D.), et Gage (P.), *The cytotoxic plant protein, b-purothionin, forms ion channels in lipid membranes*, J. Biol. Chem. **275** (2000), 823 :827.
- [Humphrey et al. 96] Humphrey (W.), Dalke (A.), et Schulten (K.), *VMD - Visual Molecular Dynamics*, J. Molec. Graphics **14** (1996), 33 :38.
- [Hwang et al. 94] Hwang (M. J.), Stockfish (T. P.), et Hagler (A. T.), *Derivation of Class II Force Fields. II. Derivation and Characterization of a Class II Force Field, CFF93, for the Alkyl Functional Group and Alkane Molecules*, J. Am. Chem. Soc. **116** (1994), 2515 :2525.
- [Ishikawa et al. 08] Ishikawa (H.), Kwak (K.), Chung (J. K.), Kim (S.), et Fayer (M. D.), *Direct observation of fast protein conformational switching*, Proc. Natl. Acad. Sci. **105** (2008), 8619 :8624.
- [Jelsch et al. 00] Jelsch (C.), Teeter (M.), Lamzin (V.), Pichon-Pesme (V.), Blessing (R. H.), et Lecomte (C.), *Accurate protein crystallo-*

- graphy at ultra-high resolution : valence electron distribution in crambin*, Proc. Natl. Acad. Sci. **97** (2000), 3171 :3176.
- [Jorgensen *et al.* 83] Jorgensen (W. L.), Chandrasekhar (J.), Madura (J.), Impey (R. W.), et Klein (M. L.), *Comparison of simple potential functions for simulating liquid water*, J. Chem. Phys. **79** (1983), 926 :935.
- [Jorgensen et Tirado-Rives 88] Jorgensen (W. L.) et Tirado-Rives (J.), *The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin*, J. Am. Chem. Soc. **110** (1988), no. 6, 1657 :1666.
- [Kabsch 76] Kabsch (W.), *A solution for the best rotation to relate two sets of vectors*, Acta Cryst. **A32** (1976), 922 :923.
- [Kendrew *et al.* 60] Kendrew (J. C.), Dickerson (R. E.), Strandberg (B. E.), Hart (R. G.), Davies (D. R.), Phillipps (D. C.), et Shore (V. C.), *Structure of Myoglobin : A Three-Dimensional Fourier Synthesis at 2 Å. Resolution*, Nature **185** (1960), 422 :427.
- [Kidera et Go 92a] Kidera (A.) et Go (N.), *Normal mode refinement : crystallographic refinement of protein dynamic structure. I. Theory and test by simulated diffraction data*, J. Mol. Biol. **225** (1992), no. 2, 457 :475.
- [Kidera et Go 92b] Kidera (A.) et Go (N.), *Normal mode refinement : crystallographic refinement of protein dynamic structure. II. Application to human lysozyme*, J. Mol. Biol. **225** (1992), no. 2, 477 :486.
- [Kneller et Hinsen 04] Kneller (G. R.) et Hinsen (K.), *Fractional Brownian dynamics in proteins*, J. Chem. Phys. **121** (2004), 10278 :10283.
- [Kneller 04] Kneller (G. R.), *Comment on Using quaternions to calculate RMSD*, J. Comp. Chem. **26** (2004), no. 15, 1660 :1662.
- [Krishtal *et al.* 06] Krishtal (A.), Senet (P.), Yang (M.), et Alsenoy (C. Van), *A Hirshfeld partitioning of polarizabilities of water clusters*, J. Chem. Phys. **125** (2006), 034312.
- [Kumar *et al.* 06] Kumar (P.), Yan (Z.), Xu (L.), Mazza (M. G.), Buldyrev (S. V.), Chen (S. H.), Sastry (S.), et Stanley (H. E.), *Glass Transition in Biomolecules and the Liquid-Liquid Critical Point of Water*, Phys. Rev. Lett. **97** (2006), 177802–1 :177802–4.
- [Landau et Lifshitz 84] Landau (L. D.) et Lifshitz (E. M.), *Statistical Physics (Course of Theoretical Physics, Volume 5)*, 3 éd., Butterworth-Heinemann, 1984.
- [Lee *et al.* 00] Lee (A. L.), Kinnear (S. A.), et Wand (A. J.), *Redistribution and loss of side chain entropy upon formation of a calmodulin peptide complex*, Nature Structural Biology **7** (2000), 72 :77.

- [Lee et Wand 01] Lee (A. L.) et Wand (A. J.), *Microscopic origins of entropy, heat capacity and the glass transition in proteins*, Nature **411** (2001), 501 :504.
- [LeGoff et al. 02] LeGoff (L.), Hallatschek (O.), Frey (E.), et Amblard (F.), *Tracer studies on F-Actin Fluctuations*, Phys. Rev. Lett. **89** (2002), 258101 :258104.
- [Lii et Allinger 89a] Lii (J. H.) et Allinger (N. L.), *Molecular mechanics. The MM3 force field for hydrocarbons. 2. Vibrational frequencies and thermodynamics*, J. Am. Chem. Soc. **111** (1989), no. 23, 8566 :8575.
- [Lii et Allinger 89b] Lii (J. H.) et Allinger (N. L.), *Molecular mechanics. The MM3 force field for hydrocarbons. 3. The van der Waals' potentials and crystal data for aliphatic and aromatic hydrocarbons*, J. Am. Chem. Soc. **111** (1989), no. 23, 8576 :8582.
- [Luo et al. 06] Luo (G.), Andricioaei (I.), Xie (X. S.), et Karplus (M.), *Dynamic Distance Disorder in Proteins Is Caused by Trapping*, J. Phys. Chem. B **110** (2006), 9363 :9367.
- [MacKerell] MacKerell (A.), *Personal communication*.
- [Mandelbrot et Ness 68] Mandelbrot (B. B.) et Ness (J. W. Van), *Fractional Brownian Motion, Fractional Noises and Applications*, SIAM Review **10** (1968), 422 :437.
- [Marx et Hutter 00] Marx (D.) et Hutter (J.), *Ab initio molecular dynamics : Theory and Implementation*, vol. 1, Modern Methods and Algorithms of Quantum Chemistry, 2000.
- [Mayo et al. 90] Mayo (S. L.), Olafson (B. D.), et Goddard (W. A.), *DREIDING : a generic force field for molecular simulations*, J. Phys. Chem. **94** (1990), no. 26, 8897 :8909.
- [McCammon et al. 77] McCammon (J. A.), Gelin (B. R.), et Karplus (M.), *Dynamics of Folded Proteins*, Nature **267** (1977), 585 :590.
- [Millet et al. 03] Millet (O.), Mittermaier (A.), Baker (D.), et Kay (L. E.), *The effects of mutations on motions of side-chains in protein L studied by ^2H NMR dynamics and scalar couplings*, J. Mol. Biol. **329** (2003), 551 :563.
- [Min et al. 05] Min (W.), Luo (G.), Cherayil (B. J.), Kou (S. C.), et Xie (X. S.), *Observation of a Power-Law Memory Kernel for Fluctuations within a Single Protein Molecule*, Phys. Rev. Lett. **94** (2005), 198302(1) :198302(4).
- [Miyazaki et al. 00] Miyazaki (Y.), Matsuo (T.), et Suga (H.), *Low-temperature heat capacity and glassy behavior of lysozyme crystal*, J. Phys. Chem. B. **104** (2000), no. 33, 8044 :8052.
- [Momany et al. 75] Momany (F. A.), McGuire (R. F.), Burgess (A. W.), et Scheraga (H. A.), *Energy parameters in polypeptides. VII. Geometric parameters, partial atomic charges, nonbonded interactions, hydrogen bond interactions, and intrinsic torsional*

- potentials for the naturally occurring amino acids*, J. Phys. Chem. **79** (1975), 2361 :2381.
- [Moritsugu et Smith 06] Moritsugu (K.) et Smith (J. C.), *Temperature-Dependent Protein Dynamics : A Simulation-Based Probabilistic Diffusion-Vibration Langevin Description*, J. Phys. Chem. B **110** (2006), no. 11, 5807 :5816.
- [Mosbah et al. 01] Mosbah (A.), Coulon (A.), Bernard (C.), Urech (K.), Rouge (P.), et Darbon (H.), *Comparative membrane interaction study of viscotoxins A3, A2 and B from mistletoe (Viscum album) and connections with their structures*, to be published (2001).
- [Nakano et al. 93] Nakano (A.), Bi (L.), Kalia (R. K.), et Vashista (P.), *Structural correlations in porous silica : Molecular dynamics simulation on a parallel computer*, Phys. Rev. Lett. **71** (1993), 85 :88.
- [Neusius et al. 08] Neusius (T.), Daidone (I.), Sokolov (I. M.), et Smith (J. C.), *Subdiffusion in Peptides Originates from the Fractal-Like Structure of Configuration Space*, Phys. Rev. Lett. **100** (2008), 188103.
- [Nilsson et Karplus 86] Nilsson (R.) et Karplus (M.), *Empirical energy functions for energy minimisation and dynamics of nucleic acids*, J. Comp. Chem. **7** (1986), no. 5, 591 :616.
- [Nishikawa et al. 74] Nishikawa (K.), Momany (F. A.), et Scheraga (H. A.), *Low-Energy Structures of Two Dipeptides and Their Relationship to Bend Conformations*, Macromolecules **7** (1974), 797 :806.
- [Oldziej et al. 04] Oldziej (S.), Lagiewka (J.), Liwo (A.), Czaplewski (C.), Chinchio (M.), Nancias (M.), et Scheraga (H. A.), *Optimization of the UNRES force field by hierarchical design of the potential-energy landscape. 3. Use of many proteins in optimization*, J. Phys. Chem. B **108** (2004), 16950 :16959.
- [Orecchini et Paciaroni 04] Orecchini (A.) et Paciaroni (A.), *High-temperature dynamical transition in beta-lactoglobulin*, Physica B. **350** (2004), 595 :598.
- [Paciaroni et al. 02] Paciaroni (A.), Cinelli (S.), et Onori (G.), *Effect of the Environment on the Protein Dynamical Transition : A Neutron Scattering Study*, Biophys J. **1983** (2002), 1157 :1164.
- [Parak et al. 80] Parak (F.), Frolov (E. N.), Kononenko (A. A.), Mössbauer (R. L.), Goldansi (V. I.), et Rubin (A. B.), *Evidence for a correlation between the photoinduced electron transfer and dynamic properties of the chromatophore membranes from Rhodospirillum rubrum*, Elsevier **117** (1980), 368 :372.
- [Parak et al. 81] Parak (F.), Frolov (E. N.), Mössbauer (R. L.), et Goldanskii (V. I.), *Dynamics of metmyoglobin crystals investigated by nuclear gamma resonance absorption*, J. Mol. Biol. **145** (1981), 825 :833.

- [Parak 86] Parak (F.), *Correlation of protein dynamics with water mobility : Mossbauer spectroscopy and microwave absorption methods*, Methods Enzymol. **127** (1986), 196 :206.
- [Parak 03] Parak (F.), *Physical aspects of protein dynamics*, Rep. Prog. Phys. **66** (2003), 103 :129.
- [Parr et Weitao 89] Parr (R. G.) et Weitao (Y.), *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, New York, 1989.
- [Paulaitis et al. 92] Paulaitis (M. E.), Sowa (M. J.), et McMinn (J. H.), *Effect of enzyme hydration on the catalytic activity of chymotrypsin in nearly anhydrous organic suspensions*, Ann. NY Acad. Sci. **672** (1992), 278 :282.
- [Pelegriini et Franco 05] Pelegriini (P. B.) et Franco (O. L.), *Plant beta-thionins : Novel insights on the mechanism of action of a multi-functional class of defense proteins*, Int. J. Biochem. Cell Biol. **37** (2005), no. 11, 2239 :2253.
- [Peng et al. 94] Peng (C. K.), Buldyrev (S. V.), Havlin (S.), Simons (M.), Stanley (H. E.), et Goldberger (A. L.), *Mosaic organization of DNA nucleotides*, Phys. Rev. E **49** (1994), 1685 :1689.
- [Perrin 28] Perrin (F.), *Etude mathématique du mouvement Brownien de rotation*, Annales scientifiques de IE.N.S. **45** (1928), 1 :51.
- [Pissis 89] Pissis (P.), *Dielectric studies of protein hydration*, J. Mol. Liquids **41** (1989), 271 :289.
- [Prestrelski et al. 93] Prestrelski (S. J.), Tedeschi (N.), Arakawa (T.), et Carpenter (J. F.), *Dehydration-induced conformational transitions in proteins and their inhibition by stabilizers*, Biophys J. **65** (1993), 661 :671.
- [Rahman 64] Rahman (A.), *Correlations in the Motion of Atoms in Liquid Argon*, Phys. Rev. **136** (1964), 405 :411.
- [Rappe et al. 92] Rappe (A. K.), Casewit (C. J.), Colwell (K. S.), III (W. A. Goddard), et Skiff (W. M.), *UFF, a Rule-Based Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations*, J. Am. Chem. Soc. **114** (1992), 10035.
- [Rasmussen et al. 92] Rasmussen (B. F.), Stock (A. M.), Ringe (D.), et Petsko (G. A.), *Crystalline ribonuclease A loses function below the dynamical transition at 220 K*, Nature **357** (1992), 423 :424.
- [Reat et al. 00] Reat (V.), Dunn (R.), Ferrand (M.), Finney (J. L.), Daniel (R. M.), et Smith (J. C.), *Solvent dependence of dynamic transitions in protein solutions*, Proc. Natl. Acad. Sci. **97** (2000), no. 18, 9961 :9966.
- [Rigdena et al. 06] Rigdena (D. J.), Littlejohna (J. E.), Joshic (H. V.), de Grootc (B. L.), et Jedrzejasa (Mark J.), *Alternate Structural Conformations of Streptococcus pneumoniae Hyaluronan Lyase : In-*

- sights into Enzyme Flexibility and Underlying Molecular Mechanism of Action*, J. Mol. Biol. **358** (2006), 1165 :1178.
- [Ringe et Petsko 03] Ringe (D.) et Petsko (G. A.), *The 'glass transition' in protein dynamics : what it is, why it occurs, and how to exploit it*, Biophys. Chem. **105** (2003), 667 :680.
- [Roh et al. 05] Roh (J. H.), Novikov (V. N.), Gregory (R. B.), Curtis (J. E.), Chowdhuri (Z.), et Sokolv (A. P.), *Onsets of anharmonicity in protein dynamics*, Phys. Rev. Lett. **95** (2005), 038101.
- [Romagnoli et al. 00] Romagnoli (S.), Ugolini (R.), Fogolari (F.), Schaller (G.), Urech (K.), Giannattasio (M.), Ragona (L.), et Molinari (H.), *NMR structural determination of viscotoxin A from *Viscum album L**, Biochem. J. **350** (2000), 569 :577.
- [Rousseau et al. 00] Rousseau (B.), Peeters (A.), et Alsenoy (C. Van), *Systematic study of the parameters determining stockholder charges*, Chem. Phys. Letter **324** (2000), no. 1, 189 :194.
- [Rueda et al. 07] Rueda (M.), Ferrer (C.), Meyer (T.), Pérez (A.), Camps (J.), Hospital (A.), Gelpí (J. L.), et M.Orozco , *A consensus view of protein dynamics*, Proc. Natl. Acad. Sci. **104** (2007), no. 3, 796 :801.
- [Rupley et al. 83] Rupley (J. A.), Gratten (E.), et Careri (G.), *Water and globular proteins*, Trends Biochem. Sci. **8** (1983), 18 :22.
- [Ryckaert et al. 77] Ryckaert (J. P.), Ciccotti (G.), et Berendsen (H. J.), *Numerical integration of the cartesian equations of motion of a system with constraints : molecular dynamics of n-alkanes*, J. Comput. Phys. **23** (1977), 327 :341.
- [Sacquin-Mora et Lavery 06] Sacquin-Mora (S.) et Lavery (R.), *Investigating the local flexibility of functional residues in hemoproteins*, Biophys J. **90** (2006), no. 8, 2706 :2717.
- [Sayle et Milner-White 95] Sayle (R. A.) et Milner-White (E. J.), *RasMol : Biomolecular graphics for all*, Trends in Biochemical Sciences **20** (1995), 374 :376.
- [Senet et Aparicio 07] Senet (P.) et Aparicio (F.), *DFT Fragment descriptors to quantify the reactivity of a molecular family. Application to amino acids*, J. Chem. Phys. **126** (2007), 145105.
- [Settles et Doster 96] Settles (M.) et Doster (W.), *Anomalous diffusion of adsorbed water : A neutron scattering study of hydrated myoglobin*, Faraday Discuss. **103** (1996), 269 :279.
- [Shai 99] Shai (Y.), *Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by alpha-helical antimicrobial and cell non-selective membrane-lytic peptides*, Biochim Biophys Acta. **1462** (1999), 55 :70.
- [Sprague et al. 87] Sprague (J. T.), Taij (J. C.), Young (Y.), et Allinger (N. L.), *The MMP2 calculational method*, J. Comp. Chem. **8** (1987), 581 :603.

- [Stec *et al.* 04] Stec (B.), Markman (O.), Rao (U.), Heffron (G.), Henderson (S.), Vernon (L. P.), Brumfeld (V.), et Teeter (M.), *Proposal for molecular mechanism of thionins deduced from physico-chemical studies of plant toxins*, The J. Peptide Res. **64** (2004), 210.
- [Steinbach et Brooks 93] Steinbach (P. J.) et Brooks (B. R.), *Protein hydration elucidated by molecular dynamics simulation*, Proc. Natl. Acad. Sci. **90** (1993), no. 19, 9135 :9139.
- [Stil *et al.* 90] Stil (W. C.), Tempczyk (A.), Hawley (R. C.), et Hendrickson (T.), *A General Treatment of Solvation for Molecular Mechanics*, J. Am. Chem. Soc. **112** (1990), 6127 :6129.
- [Stillinger et Rahman 74] Stillinger (F.H.) et Rahman (A.), *Improved simulation of liquid water by molecular dynamics*, J. Chem. Phys. **60** (1974), 1545 :1557.
- [Swope *et al.* 82] Swope (W. C.), Andersen (H. C.), Berens (P. H.), et Wilson (K. R.), *A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules : Application to small water clusters*, J. Chem. Phys. **76** (1982), 637 :649.
- [Teeter *et al.* 93] Teeter (M.), Roe (S. M.), et Heo (N. H.), *Atomic resolution (0.83 Å) crystal structure of the hydrophobic protein crambin at 130 K*, J. Mol. Biol. **230** (1993), 292 :311.
- [Teeter *et al.* 01] Teeter (M.), Yamano (A.), Stec (B.), et Mohanty (U.), *On the nature of a glassy state of matter in a hydrated protein : Relation to protein function*, Proc. Natl. Acad. Sci. **98** (2001), 11242 :11247.
- [Teeter et Whitlow 87] Teeter (M. M.) et Whitlow (M. D.), *Hydrogen Bonding in the High Resolution Structure of the Protein Crambin*, Transactions ACA **2** (1987), 75 :88.
- [Tilton *et al.* 92] Tilton (R. F. Jr.), Dewan (J. C.), et Petsko (G. A.), *Effects of temperature on protein structure and dynamics : X-ray crystallographic studies of the protein ribonuclease-A at nine different temperatures from 98 to 320K*, Biochemistry **31** (1992), 2469 :2481.
- [Tournier *et al.* 03] Tournier (A. L.), Jiancong (X.), et Smith (J. C.), *Translational hydration water dynamics drives the protein glass transition*, Biophysical Journal **85** (2003), 1871 :1875.
- [Verlet 67] Verlet (L.), *Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules*, Phys. Rev. **159** (1967), 98 :103.
- [Waller 23] Waller (I.), *Zur Frage der Einwirkung der Wärmebewegung auf die Interferenz von Röntgenstrahlen*, Zeitschrift für Physik **17** (1923), 398 :408.

- [Warshel 73] Warshel (A.), *Quantum Mechanical Consistent Force Field (QCF/PI) Method : Calculations of Energies, Conformations and Vibronic Interactions of Ground and Excited States of Conjugated Molecules*, Israel J. Chem. **11** (1973), no. 5, 709.
- [Watanabe 91] Watanabe (H.), *Thermal Dilatation of Water between 0 °C and 44 °C*, Metrologia **28** (1991), 33 :43.
- [Weik et al. 01] Weik (M.), Ravelli (R. B. G.), Silman (I.), Sussman (J. L.), Gros (P.), et Kroon (J.), *Specific protein dynamics near the solvent glass transition assayed by radiation-induced structural changes*, Protein Science **10** (2001), 1953 :1961.
- [Weiner et al. 84] Weiner (S. J.), Kollman (P. A.), Case (D. A.), Singh (U. C.), Ghio (C.), Alagona (G.), Profeta (S. JR), et Weiner (P.), *A new force field for molecular mechanical simulation of nucleic acids and proteins*, J. Am. Chem. Soc. **106** (1984), no. 3, 765 :784.
- [Wuthrich 90] Wuthrich (K.), *Protein structure determination in solution by NMR spectroscopy*, J. Biol. Chem. **265** (1990), no. 36, 22059 :22062.
- [Xu et al. 99] Xu (Y.), Wu (J.), Gorenstein (D.), et Braun (W.), *Automated 2D NOESY assignment and structure calculation of Crambin(S22/I25) with the self-correcting distance geometry based NOAH/DIAMOD programs*, J. Magn Reson. **136** (1999), 76 :85.
- [Yoshida et al. 02] Yoshida (T.), Hiroyasu (T.), Miki (M.), Ogura (M.), et Okamoto (Y.), *Energy Minimization of Protein Tertiary Structure by Parallel Simulated Annealing using Genetic Crossover*, Proceedings of 2002 Genetic and Evolutionary Computation Conference, 2002, p. 49 :51.
- [Zaccai 00] Zaccai (G.), *How Soft Is a Protein? A Protein Dynamics Force Constant Measured by Neutron Scattering*, Science **288** (2000), no. 5471, 1604 :1607.
- [Zanotti et al. 99] Zanotti (J. M.), Bellissent-Funel (M. C.), et Parello (J.), *Hydration-Coupled Dynamics in Proteins Studied by Neutron Scattering and NMR : The Case of the Typical EF-Hand Calcium-Binding Parvalbumin*, Biophys. J. **76** (1999), no. 5, 2390 :2411.
- [Zanotti et al. 05] Zanotti (J. M.), Bellissent-Funel (M. C.), et Chen (S. H.), *Experimental evidence of a liquid-liquid transition in interfacial water*, Europhysics letters **71** (2005), no. 1, 91 :97.
- [Zanotti et al. 07] Zanotti (J. M.), Bellissent-Funel (M. C.), et Kolesnikov (A. I.), *Phase transitions of interfacial water at 165 and 240K. Connections to bulk water physics and protein dynamics*, The European Physical Journal. Special Topics **141** (2007), 227 :233.