

Université de Montréal

Étude théorique de la coopérativité dans la liaison des ions  $\text{Ca}^{2+}$  aux  
sites de la calbindine  $\text{D}_{9k}$

par

Sylvie Marchand

Département de physique

Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures  
en vue de l'obtention du grade de  
Philosophiæ Doctor (Ph.D.)  
en physique

octobre, 1997

©Sylvie Marchand, 1997



QC  
3  
U54  
1998  
V.007

Université de Montréal

Étude théorique de la régularité dans la liaison des jour C<sup>2</sup>, avec  
des de la calibration D<sup>2</sup>

par  
Sylvie Larivière  
Département de physique  
École des arts et des sciences

Thèse présentée à la Faculté des études supérieures  
en vue de l'obtention du grade de  
Philosophie Doctor (Ph.D.)  
en physique



reçu le 8 mai 1998  
Sylvie Larivière, 1998

Université de Montréal  
Faculté des études supérieures

Cette thèse intitulée:

Étude théorique de la coopérativité dans la liaison des ions  $\text{Ca}^{2+}$  aux  
sites de la calbindine  $\text{D}_{9k}$

présentée par:

Sylvie Marchand

a été évaluée par un jury composé des personnes suivantes:

Rémy Sauvé,	président-rapporteur
Benoît Roux,	directeur de recherche
John Gunn,	membre du jury
Alain St-Amant,	examineur externe

Thèse acceptée le: 24 novembre 1997

## SOMMAIRE

Le sujet principal de cette thèse consiste en l'étude de la coopérativité dans la liaison des ions  $\text{Ca}^{2+}$  aux sites de la calbindine  $\text{D}_{9k}$  par des méthodes de simulation de dynamique moléculaire (DM). La calbindine  $\text{D}_{9k}$  est une protéine de 75 acides aminés qui lie deux ions  $\text{Ca}^{2+}$  de façon coopérative, i.e. que la liaison d'un premier  $\text{Ca}^{2+}$  dans un site augmente l'affinité de l'autre site. Les simulations sont basées sur une fonction de potentiel classique (CHARMM) où les interactions sont calculées par paires d'atomes et où les effets à plusieurs corps ou non-additifs sont négligés. Les paramètres du champ de forces *param22* pour les acides aminés sont utilisés, ainsi que le modèle de molécule d'eau TIP3P.

Des paramètres de Lennard-Jones pour le  $\text{Ca}^{2+}$  reproduisant raisonnablement bien son énergie libre d'hydratation et les courbes d'interactions *ab initio*  $\text{Ca}^{2+}$ -ligands sont déterminés. La contribution non-additive dans les interactions  $\text{Ca}^{2+}$ - $(\text{H}_2\text{O})_n$  est évaluée par des méthodes *ab initio*. Cette contribution étant bien représentée par l'énergie de polarisation induite au second ordre, des polarisabilités pour les atomes lourds des modèles des ligands du  $\text{Ca}^{2+}$  dans les sites de la calbindine  $\text{D}_{9k}$  sont déterminées. Une tentative d'inclure un terme d'énergie de polarisation induite a révélé la grande sensibilité aux paramètres du champ de forces, ainsi que des difficultés inhérentes au développement d'une fonction de potentiel tenant compte des effets de polarisation induite. La fonction de potentiel à deux corps est donc utilisée pour le reste de cette thèse.

Dans ce contexte, des simulations de DM de 200 ps sont effectuées sur quatre états d'occupation de la calbindine  $\text{D}_{9k}$  incluse dans 2 285 molécules d'eau. Les structures simulées  $\text{S}_{\text{II}}$  et D sont demeurées proches de leur structure

expérimentale de départ, tandis que la structure simulée A s'en est éloignée plus considérablement. La structure simulée A( $S_{II}$ ), construite à partir de la structure expérimentale  $S_{II}$  demeure davantage proche de la structure expérimentale A. Structuralement, les états liés au  $Ca^{2+}$  se sont révélés similaires entre eux et différents de l'état A, signifiant que la majorité des changements structuraux se produit lors de la liaison du premier  $Ca^{2+}$ . Pour l'aspect dynamique, il s'avère que la liaison successive du  $Ca^{2+}$  rend la structure plus rigide, avec la majeure partie des effets concentrée dans la liaison d'un premier ion au site I ou au site II. Les résultats des simulations s'accordent qualitativement bien avec les données expérimentales. De plus, il est observé que les interactions hydrophobes pourraient jouer un rôle important dans la coopérativité.

Un modèle analytique est développé afin d'exprimer la partie librationnelle des paramètres d'ordre des vecteurs N-H ( $\langle S^2 \rangle_{N-H}$ ) en fonction de l'amplitude des librations effectuées par le vecteur N-H. Ces librations sont exprimées en termes des fluctuations anticorrélées des angles dièdres  $\phi$  et  $\psi$ . Les résultats suggèrent que des fluctuations trop grandes des  $\phi$  et  $\psi$  sont en partie la source des valeurs trop faibles calculées pour certains  $\langle S^2 \rangle_{N-H}$ , et une surestimation des librations par seulement quelques degrés suffisent à expliquer le désaccord avec les données expérimentales.

Les bases rigoureuses d'une méthode de calcul de la différence d'énergie libre associée à la mutation V61A dans la calbindine  $D_{9k}$  sont développées. Les calculs d'énergies libres faits par des méthodes perturbatives ne sont pas complètement convergés à ce point-ci, mais il semble que l'organisation des atomes autour de la chaîne latérale mutée (alanine) ne soit pas aussi compacte que dans la calbindine  $D_{9k}$  P43G originale (valine). Des temps plus longs de simulations devraient être accumulés à chaque fenêtre de perturbation afin de poursuivre le projet et vérifier cette hypothèse.

**Mots clés:** dynamique moléculaire, calbindine  $D_{9k}$ , coopérativité,  $Ca^{2+}$

## TABLE DES MATIÈRES

SOMMAIRE . . . . .	iii
TABLE DES MATIÈRES . . . . .	v
LISTE DES TABLEAUX . . . . .	xi
LISTE DES FIGURES . . . . .	xv
LISTE DES ABRÉVIATIONS . . . . .	xxiii
CHAPITRE 1: Introduction . . . . .	1
1.1 Le $\text{Ca}^{2+}$ et les protéines auxquelles il se lie . . . . .	1
1.2 La calbindine $\text{D}_{9\text{k}}$ . . . . .	2
1.3 Régulation allostérique et coopérativité . . . . .	7
1.4 La coopérativité chez la calbindine $\text{D}_{9\text{k}}$ . . . . .	8
1.4.1 Formulation . . . . .	8
1.4.2 Facteurs identifiés comme jouant un rôle dans la coopérativité	12
1.4.2.1 Facteurs structuraux . . . . .	13
1.4.2.2 Facteurs dynamiques . . . . .	14
1.4.2.3 Influence de l'eau . . . . .	15
1.4.3 Types d'interactions impliquées dans la coopérativité . . . . .	15

1.4.4	Modèle proposé . . . . .	16
1.5	But de la thèse . . . . .	18
1.6	Plan de la thèse . . . . .	18
CHAPITRE 2: Les méthodes . . . . .		20
2.1	Les calculs <i>ab initio</i> <sup>47, 48</sup> . . . . .	20
2.2	La dynamique moléculaire (DM) <sup>52, 53, 54, 55</sup> . . . . .	25
2.2.1	Le champ de forces et les contraintes <sup>56, 57, 58, 59</sup> . . . . .	26
2.2.2	Les conditions aux frontières <sup>61, 62, 63</sup> . . . . .	30
2.2.3	Étapes des simulations . . . . .	32
2.2.4	Le processus d'évolution dans le temps . . . . .	34
2.3	Les calculs d'énergies libres <sup>56, 62, 63, 67</sup> . . . . .	36
CHAPITRE 3: Étude des interactions entre le $\text{Ca}^{2+}$ et ses ligands dans la calbindine $\text{D}_{9k}$ . . . . .		41
3.1	Introduction . . . . .	41
3.2	Paramétrisation du $\text{Ca}^{2+}$ . . . . .	42
3.2.1	Calcul de l'énergie libre d'hydratation du $\text{Ca}^{2+}$ et propriétés statiques du $\text{Ca}^{2+}$ dans l'eau . . . . .	42
3.2.2	Calculs de courbes d'énergies d'interaction de systèmes $\text{Ca}^{2+}$ -ligands . . . . .	47
3.3	Évaluation de la partie non-additive des interactions $\text{Ca}^{2+}$ - $\text{H}_2\text{O}$ pour des agrégats de deux à six $\text{H}_2\text{O}$ . . . . .	60

3.4	Évaluation de polarisabilités approchées pour les atomes lourds liés au $\text{Ca}^{2+}$ dans les sites de la calbindine $\text{D}_{9k}$ . . . . .	64
3.5	De l'inclusion de l'énergie de polarisation induite par le $\text{Ca}^{2+}$ sur une molécule d'eau . . . . .	69
3.6	Conclusion . . . . .	73
CHAPITRE 4: Simulations de dynamique moléculaire de la calbindine $\text{D}_{9k}$ dans quatre de ses états d'occupation . . . . .		
4.1	Introduction . . . . .	75
4.2	Méthode . . . . .	81
4.2.1	Préparation des structures de départ . . . . .	81
4.2.2	Protocoles d'équilibration et de production des trajectoires . . . . .	84
4.3	Résultats . . . . .	85
4.3.1	Analyse comparée des propriétés structurales à l'équilibre . . . . .	85
4.3.1.1	Évolution temporelle de la différence structurale entre les structures expérimentales de départ et les structures simulées . . . . .	85
4.3.1.2	Différences structurales entre la structure expérimentale de départ et la structure simulée, et des structures simulées entre elles . . . . .	89
4.3.1.3	Différences de coordinence et de disposition des ligands entre les structures expérimentales de départ et les structures simulées . . . . .	95

4.3.1.4	Différences de disposition des ponts hydrogènes entre les structures expérimentales et les structures simulées, et des structures simulées entre elles . . .	100
4.3.1.5	Différences de compacité entre les structures simulées . . . . .	102
4.3.1.6	Différences d'hydratation des chaînes principales, des chaînes latérales et des sites entre les structures simulées . . . . .	108
4.3.2	Analyse comparée des propriétés dynamiques à l'équilibre .	112
4.3.2.1	Fluctuations moyennes autour de la position moyenne pour chaque acide aminé . . . . .	112
4.3.2.2	Fluctuations moyennes des angles dièdres $\phi$ et $\psi$	116
4.3.2.3	Paramètres d'ordre $\langle S^2 \rangle$ (RMN) des vecteurs parallèles aux liens NH et $C_\alpha H_\alpha$ ; comparaisons entre les valeurs expérimentales et calculées . . .	120
4.3.3	Contributions structurales et énergétiques à la coopérativité	128
4.3.3.1	Contribution électrostatique défavorable à la coopérativité associée à la liaison du $Ca^{2+}$ . . . .	128
4.3.3.2	Contribution des interactions hydrophobes: interactions de van der Waals entre la chaîne latérale de la Val61 et le reste de la protéine; distributions des énergies et valeurs moyennes . . . . .	130
4.3.3.3	Distributions des distances des deux ponts hydrogènes du feuillet- $\beta$ reliant les deux sites . . . . .	131

4.4 Synthèse et conclusion . . . . .	134
CHAPITRE 5: Développement d'une représentation analytique de la partie librationnelle des paramètres d'ordre des vecteurs N-H de la chaîne principale de la calbindine $D_{9k}$ . . . . .	139
5.1 Introduction . . . . .	139
5.2 Théorie . . . . .	143
5.3 Résultats . . . . .	147
5.4 Conclusion . . . . .	156
CHAPITRE 6: Calculs de la différence d'énergie libre associée à la mutation V61A et contribution à la coopérativité . . . . .	158
6.1 Introduction . . . . .	158
6.2 Méthode . . . . .	161
6.2.1 Formalisme . . . . .	161
6.3 Illustration et validation de la méthode . . . . .	166
6.4 Résultats et discussion de la mutation V61A dans la calbindine $D_{9k}$ P43G . . . . .	174
6.5 Conclusion . . . . .	186
CHAPITRE 7: Conclusion . . . . .	188
7.1 Rappel du but de la thèse . . . . .	188
7.2 Contributions principales . . . . .	189
7.3 Suggestions pour de futures simulations . . . . .	193

RÉFÉRENCES . . . . . 195

## LISTE DES TABLEAUX

I	Paramètres d'interaction de Lennard-Jones du $\text{Ca}^{2+}$ et de l'oxygène d'une molécule d'eau. . . . .	44
II	Contribution électrostatique à l'énergie libre d'hydratation du $\text{Ca}^{2+}$ calculée pour les cas où le $\text{Ca}^{2+}$ est simulé dans une sphère de $n$ molécules d'eau. . . . .	45
III	Distances optimales et énergies minimales de l'interaction $\text{Ca}^{2+}$ -ligand, calculées par la méthode HF (base 6-311G** <sup>82, 83, 84</sup> pour les atomes C, N, O et H, et base de Wachters <sup>87</sup> contractée par Gianolo et coll. <sup>88</sup> pour le $\text{Ca}^{2+}$ .) et par le logiciel CHARMM ( <i>param22</i> <sup>51</sup> ). . . . .	51
IV	Charge partielle présente sur chacun des atomes des molécules modèles à la distance $\text{Ca}^{2+}$ -O optimale (voir tableau III). Les charges <i>ab initio</i> proviennent d'analyses de charges de Mulliken (base 6-311G** <sup>82, 83, 84</sup> pour les atomes C, N, O et H et base de Wachters <sup>87</sup> contractée par Gianolo et coll. <sup>88</sup> pour le $\text{Ca}^{2+}$ ). . . . .	53
V	Évaluation de la contribution non-additive dans les calculs <i>ab initio</i> . La distance $r_{\text{min}}$ est en Å et les énergies en kcal/mol. . . . .	62
VI	Évaluation des énergies d'interactions à deux corps dans les calculs effectués avec CHARMM et <i>param22</i> . . . . .	63
VII	Charges sur les $\text{Ca}^{2+}$ et sur l'oxygène des molécules d'eau en fonction du nombre de ces dernières. . . . .	63

VIII	Polarisabilités en $\text{\AA}^3$ de molécules, groupements chimiques et atomes pour les molécules modèles. . . . .	68
IX	Différences structurales moyennes ( $\text{\AA}$ ) entre la structure expérimentale de départ et les structures simulées, calculées pour chaque élément de structure secondaire. L'abréviation c.p. signifie chaîne principale; tous les résultats excluent les atomes d'hydrogène et les $\text{Ca}^{2+}$ . . .	90
X	Différences structurales ( $\text{\AA}$ ) entre les quatre états d'occupation de la calbindine $\text{D}_{9k}$ , calculées à partir de structures simulées moyennes. L'abréviation c.p. signifie chaîne principale; tous les résultats excluent les atomes d'hydrogène et les $\text{Ca}^{2+}$ . . . . .	92
XI	Distances ( $\text{\AA}$ ) entre le $\text{Ca}^{2+}$ et ses ligands dans l'état D; exp. signifie dans la structure expérimentale de départ; simul. est la moyenne des distances évaluées dans les 20 000 configurations de l'ensemble. Les atomes de type $\text{OE}^*$ ou $\text{OD}^*$ appartiennent à une chaîne latérale et les atomes O appartiennent à la chaîne principale. . . . .	96
XII	Distances entre le $\text{Ca}^{2+}$ et ses ligands dans l'état $\text{S}_I(\text{D})$ . n.d. signifie non déterminé. Voir le tableau XI pour les détails. . . . .	98
XIII	Distances entre le $\text{Ca}^{2+}$ et ses ligands dans l'état $\text{S}_I(\text{D})$ . Voir le tableau XI pour les détails. . . . .	99

XIV	Ponts hydrogènes reliant les atomes HN à O de la chaîne principale qui sont présents dans les structures expérimentales (RMN) mais qui sont absents dans les structures simulées. La structure simulée A(S <sub>II</sub> ) est absente de ce tableau parce qu'elle a retenu tous les ponts hydrogènes de la structure expérimentale A. L'annotation p signifie que le pont hydrogène a été perdu; les autres ponts hydrogènes de la liste établissent des ponts hydrogènes avec un acide aminé adjacent. L'astérisque indique que ce pont hydrogène relie les deux sites. . . . .	101
XV	Compilation des rayons de giration moyens ( $r_G$ ) et des surfaces accessibles moyennes ( $\sigma_{acc.}$ ) pour les cinq structures simulées. . . . .	103
XVI	Comparaisons des valeurs moyennes (par élément de structure secondaire) des paramètres d'ordre des vecteurs N-H expérimentales <sup>24</sup> avec celles calculées à partir des trajectoires. . . . .	126
XVII	Dépendance du $\langle S^2 \rangle_{lib.}$ en fonction du $\sigma_{lib.}$ . . . . .	153
XVIII	Résultats des calculs de différences d'énergies libres (kcal/mol) à l'aller et au retour effectués par les méthodes des perturbations (P.T.), d'intégration thermodynamique (I.T.) et de l'algorithme WHAM. Les résultats aller/retour de WHAM signifient que toutes les données pour les 44 fenêtres ont été analysées simultanément.	169
XIX	Différences d'énergies libres d'hydratation entre l'alanine et la valine obtenues par les méthodes des perturbations (P.T.), d'intégration thermodynamique (I.T.) et WHAM. . . . .	172

XX	Différences structurales ( $\text{\AA}$ ) (équation 4.1) entre la structure simulée moyenne à la valeur de $\lambda = 0.9$ à l'aller et avec la structure moyenne à la même valeur de $\lambda$ au retour. Ces différences sont données pour chaque élément de structure secondaire. L'abréviation c.p. signifie chaîne principale; tous les résultats excluent les atomes d'hydrogène. . . . .	178
XXI	Résultats des calculs de différences d'énergies libres (kcal/mol) obtenus par l'algorithme WHAM. Les résultats totaux de WHAM signifient que toutes les données pour les 44 fenêtres ont été analysées simultanément. . . . .	179
XXII	Rapports des constantes macroscopiques entre les calbindines $D_{9k}$ P43G (Val) et P43G-V61A (Ala). . . . .	180
XXIII	Distances moyennes ( $\text{\AA}$ ) entre les atomes lourds de la chaîne latérale de la valine et les atomes lourds les plus proches appartenant à d'autres acides aminés non adjacents. ( $n = 20000$ configurations utilisées pour ces calculs.) . . . . .	181
XXIV	Distances moyennes ( $\text{\AA}$ ), dans les mutants V61A entre les atomes lourds de la chaîne latérale de l'alanine (Ala61) et les atomes lourds les plus proches appartenant à d'autres acides aminés non adjacents. ( $n = 20$ configurations utilisées pour ces calculs.) .	181

## LISTE DES FIGURES

1	Structure de la main EF du domaine C-terminal de la parvalbumine et analogie avec celle d'une main droite qui est à l'origine de l'appellation. <sup>4</sup> . . . . .	3
2	Illustration schématique de la calbindine D <sub>9k</sub> où sont représentés les chaînes principales des hélices, des sites et du segment 36-45. . . . .	4
3	Configuration (en stéréo) des ligands autour du Ca <sup>2+</sup> lorsque dans un ou l'autre des sites de la calbindine D <sub>9k</sub> . Chaque point représente un atome d'oxygène (ligand) se trouvant sur une des deux pyramides à base pentagonale mitoyenne. Les chiffres de 1 à 7 indiquent le numéro du ligand dont la nature est donnée à la figure 4 pour les deux sites. . . . .	6
4	Séquences des acides aminés qui composent les sites I et II et identification des atomes liés au Ca <sup>2+</sup> (voir figure 3). . . . .	6
5	Courbe du nombre moyen de Ca <sup>2+</sup> lié aux deux sites de la calbindine D <sub>9k</sub> en fonction de la concentration libre dans le milieu intracellulaire en présence (tirets) et en absence (ligne continue) de coopérativité <sup>23</sup> . . . . .	9

6	Modèle à quatre états du processus de liaison de deux $\text{Ca}^{2+}$ à la calbindine $\text{D}_{9k}$ où les $k_I$ , $k_{II}$ , $k_{I,II}$ et $k_{II,I}$ sont les constantes microscopiques de liaison, chacune étant spécifique à un site <sup>23</sup> . Les $K_1$ et $K_2$ sont les constantes macroscopiques de liaison; P signifie la protéine. . . . .	10
7	Ponts hydrogènes reliant les acides aminés 23 à 61 qui appartiennent respectivement au site I et au site II. . . . .	14
8	Modèle de coopérativité chez la calbindine $\text{D}_{9k}$ proposé par M. Akke; <sup>25</sup> le dessin est tiré de Falke et coll. <sup>45</sup> . . . . .	17
9	Définition d'un angle dièdre. L'angle inclut quatre atomes et est déterminé par la rencontre des deux plans incluant les trois premiers et les trois derniers des quatre atomes. Par exemple, les atomes 1, 2, 3 et 4 pourraient être les atomes de carbone d'un groupe $\text{CH}_3\text{-CH}_2\text{-CH}_2\text{-CH}_3$ . . . . .	27
10	Définition d'un angle impropre. L'angle est formé par la rencontre des plans incluant les atomes 2, 3, 4 et 1, 3, 4. . . . .	28
11	Schéma de la méthode des perturbations. . . . .	39
12	Distribution radiale $g(r)$ ( $\text{\AA}^{-3}$ des atomes O (de l'eau) en fonction de la distance $r$ à partir du centre du $\text{Ca}^{2+}$ , calculée à partir des trajectoires de DM. . . . .	46

13	Structures et configurations des modèles utilisés pour représenter les interactions $\text{Ca}^{2+}$ -ligand dans les sites de la calbindine $\text{D}_{9k}$ : a) $\text{Ca}^{2+}$ - $\text{H}_2\text{O}$ ; b) $\text{Ca}^{2+}$ -acétamide et $\text{Ca}^{2+}$ -NMA; c) $\text{Ca}^{2+}$ -propionate (configuration <i>bidentée</i> ); d) $\text{Ca}^{2+}$ -propionate (configuration <i>monodentée</i> ). Les lignes pointillées représentent l'axe selon lequel est variée la distance $\text{Ca}^{2+}$ -O dans les calculs subséquents. . . . .	49
14	Courbes d'énergies d'interaction $\text{Ca}^{2+}$ - $\text{H}_2\text{O}$ en fonction de la distance telle que définie à la figure 13. . . . .	55
15	Courbes d'énergies d'interaction $\text{Ca}^{2+}$ -acétamide (haut) et $\text{Ca}^{2+}$ -NMA (bas) en fonction de la distance telle que définie à la figure 13. . . . .	57
16	Courbes d'énergies d'interaction $\text{Ca}^{2+}$ -propionate <i>monodenté</i> (haut) et <i>bidenté</i> (bas) en fonction de la distance telle que définie à la figure 13. . . . .	58
17	Courbes d'énergies d'interaction $\text{Ca}^{2+}$ -propionate <i>bidenté</i> . Dans la courbe de <i>param22</i> , le calcium a une charge de $+1.88e$ . La courbe HF est la même qu'à la figure 16 (bas). . . . .	59
18	Courbes de l'énergie d'interaction $\text{Ca}^{2+}$ - $\text{H}_2\text{O}$ en fonction de la distance telle que définie à la figure 13 et incluant les termes d'énergies de polarisation au premier et second ordre. . . . .	71
19	Structure de l'état D de la calbindine $\text{D}_{9k}$ entourée par 2 285 molécules d'eau. . . . .	83

20	Différences entre les structures expérimentales et les structures simulées ( $D_{\text{struc.}}$ en fonction du temps de simulation (production de la trajectoire) en considérant tous les atomes lourds de la calbindine $D_{9k}$ (haut) ou tous les atomes lourds de la chaîne principale, en excluant le $\text{Ca}^{2+}$ dans les deux cas. Les $D_{\text{struc.}}$ sont évaluées selon l'équation 4.1. . . . .	87
21	Matrices des contacts des structures simulées A et $A(S_{\text{II}})$ : compilation des atomes lourds de la chaîne principale se trouvant à moins de 5 Å d'un autre atome lourd de la chaîne principale. Les deux atomes doivent aussi être séparés par au moins 75 atomes dans la séquence primaire. . . . .	105
22	Matrices des contacts des structures simulées $S_{\text{I}}(D)$ et $S_{\text{II}}$ . Voir figure 21. . . . .	106
23	Matrice des contacts de la structure simulée D. Voir figure 21. . . . .	107
24	Hydratation de la chaîne principale des structures simulées. L'ordonnée donne le nombre de molécules d'eau dans un rayon de 4 Å du centre de masse de la chaîne principale d'un acide aminé. . . . .	110
25	Hydratation de la chaîne latérale de chaque acide aminé pour toutes les structures simulées. (Voir figure 24.) . . . . .	111
26	Fluctuations moyennes (Å) autour de la position moyenne de chaque acide aminé pondérées par la masse de chaque atome lourd inclus dans le calculs. . . . .	113

27	Comparaison entre les fluctuations moyennes ( $\text{\AA}$ ) de chaque atome lourd de la calbindine $D_{9k}$ calculées à partir de la trajectoire (simul.) et calculées à partir des facteurs $B$ déterminés par diffraction des rayons X (exp.). <sup>15</sup> . . . . .	114
28	Fluctuations moyennes autour de la valeur moyenne de chaque angle dièdre $\phi$ . . . . .	118
29	Fluctuations moyennes autour de la valeur moyenne de chaque angle dièdre $\psi$ . . . . .	119
30	Comparaisons des paramètres d'ordre ( $\langle S^2 \rangle$ ) des vecteurs N-H provenant de données expérimentales avec ceux calculés à partir des trajectoires. . . . .	124
31	Paramètres d'ordre ( $\langle S^2 \rangle$ ) des vecteurs $C_\alpha H_\alpha$ calculés à partir des trajectoires. . . . .	127
32	Distributions des énergies d'interaction de van der Waals entre la chaîne latérale de la Val61 et le reste de la protéine. . . . .	131
33	Distributions radiales des longueurs des ponts hydrogènes entre les acides aminés Leu23 et Val61 qui font partie respectivement du site I et II. . . . .	133
34	Modèle des librations d'un vecteur N-H: à la position d'équilibre du vecteur N-H ( $\Phi=0$ ), les atomes O, C, N et H se trouvent dans le plan $x-z$ ; les librations entraînent l'atome H exclusivement dans le plan $x-y$ , tandis que les atomes C et N demeurent dans le plan $x-z$ . . . . .	145

- 35 Comparaisons des paramètres d'ordre  $\langle S^2 \rangle$  des vecteurs N-H de l'état D de la calbindine D<sub>9k</sub> provenant de données expérimentales avec ceux calculés à partir des trajectoires de 200 et 400 ps. . . . . 148
- 36 Paramètres d'ordre des vecteurs C<sub>α</sub>-C<sub>α</sub> de l'état D de la calbindine D<sub>9k</sub> calculés à partir de la trajectoire de 400 ps. . . . . 150
- 37 Contribution des mouvements de librations au paramètre d'ordre total calculée par  $\langle S^2 \rangle_{\text{N-H}} / \langle S^2 \rangle_{\text{C}_\alpha\text{-C}_\alpha}$ . . . . . 151
- 38 Comparaison du modèle des librations (équation 5.7) avec les valeurs tirées de la trajectoire de 400 ps de l'état D . . . . . 152
- 39 Comparaison du modèle de Brühweiler et Wright (équation 5.17,  $\Theta = 62^\circ$ ) avec les valeurs des  $\langle S^2 \rangle_{\text{lib.}}^{\text{Brühsh.}}$  tirées de la trajectoire de 400 ps de l'état D et calculées selon l'équation 5.18, et les valeurs de  $\sigma_\phi$ . . . . . 156
- 40 Cycles thermodynamiques des changements d'énergies libres associés aux liaisons successives du Ca<sup>2+</sup> dans la calbindine D<sub>9k</sub> (coopérativité) dans les cas où l'acide aminé en position 61 est une valine (cycle interne) ou une alanine (cycle externe). Les étapes représentées par des flèches horizontales ou verticales correspondent aux changements d'énergies libres associées à la mutation V61A pour chaque état d'occupation de la calbindine D<sub>9k</sub>. 162
- 41 Représentation des acides aminés faisant partie de la mutation: les atomes faisant partie de la mutation sont les atomes qui ne sont pas communs aux deux acides aminés, c'est-à-dire que les deux groupements méthyles de la valine sont échangés pour deux HB dans l'alanine. . . . . 164

42	Variation du paramètre de couplage $\lambda$ en fonction de la fenêtre de simulation. Une fonction tanh a servi à générer les valeurs de $\lambda$ .	165
43	Cycle thermodynamique des changements d'énergies libres associés à la mutation Ala→Val. . . . .	167
44	Variation de l'énergie libre $\Delta G^{\text{vide}}$ lors des calculs de perturbations du tripeptide dans le vide en fonction de la valeur de $\lambda$ . La courbe du haut correspond aux résultats des méthodes P.T. et I.T., et celle du bas aux résultats obtenus de l'algorithme WHAM en utilisant soit les données pour l'aller seulement, le retour seulement, ou les données aller-retour combinées. . . . .	170
45	Voir figure 44. Lors de ces calculs, le tripeptide a été simulé dans 73 molécules d'eau. . . . .	171
46	Variations de l'énergie libre en fonction de la valeur de $\lambda$ pour la mutation effectuée dans la structure simulée A(S <sub>II</sub> ) (haut) et S <sub>I</sub> (D) (bas). . . . .	176
47	Variations de l'énergie libre en fonction de la valeur de $\lambda$ pour la mutation effectuée dans l'état S <sub>II</sub> (haut) et D (bas). . . . .	177
48	Représentation des acides aminés les plus proches des atomes soit de la valine ou de l'alanine et nomenclature à utiliser pour lire les tableaux XXIII et XXIV. . . . .	182
49	Évolution de la distance moyenne CG2(Val61)-CE1(Phe50) en fonction du temps de simulation (ps) (L'état D compte 400 ps). . . . .	184

50	Distributions radiales $\text{\AA}^{-3}$ des atomes O (de $\text{H}_2\text{O}$ ) en fonction de la distance à partir du centre de masse de la chaîne latérale de la Val61, calculées à partir des trajectoires de DM analysées au chapitre 4. . . . .	185
----	---	-----

## LISTE DES ABRÉVIATIONS

A15D *Aspartate en position 15 plutôt que alanine*

A *Etat apo ou non-lié de la calbindine D<sub>9k</sub>*

A(S<sub>II</sub>) *Etat non-lié de la calbindine D<sub>9k</sub> dont la simulation de DM a été effectuée à partir de la structure RMN S<sub>II</sub> dont le Ca<sup>2+</sup> a été supprimé*

ABNR *Adapted Basis Set Newton-Raphson (algorithme de minimisation d'énergie)*

Ac *Acétyle*

Ala *Alanine*

Asn *Asparagine*

Asp *Aspartate*

CHARMM *Chemistry at HARvard Macromolecular Mechanics (logiciel de DM)*

D *Etat où les deux sites de la calbindine D<sub>9k</sub> sont occupés par le Ca<sup>2+</sup>*

DM *Dynamique moléculaire*

D19N *Asparagine en position 17 plutôt que aspartate*

E17Q *Glutamine en position 17 plutôt que glutamate*

E26Q *Glutamine en position 26 plutôt que glutamate*

E60D *Aspartate en position 60 plutôt que glutamate*

E60Q *Glutamine en position 60 plutôt que glutamate*

Gln *Glutamine*

Glu *Glutamate*

Gly *Glycine*

HF *Hartree-Fock (méthode de calcul ab initio de structure électronique)*

I.T. *Intégration thermodynamique*

Leu *Leucine*

Lys *Lysine*

MP $x$  *Möller-Plesset perturbation theory du  $x$ ième ordre (méthode de calcul ab initio de structure électronique)*

N56A *Alanine en position 56 plutôt que asparagine*

NMA *N-méthyle acétamide*

NOE *Nuclear Overhauser Effet (transfert de magnétisation lors des interactions spin-spin en RMN)*

P20G *Glycine en position 20 plutôt que proline*

P43G *Glycine en position 43 plutôt que proline*

P43M *Méthionine en position 43 plutôt que proline*

Phe *Phénylalanine*

Pro *Proline*

P.T. *Perturbation thermodynamique*

RMN *Résonance magnétique nucléaire*

S<sub>I</sub>(D) *Etat où un Ca<sup>2+</sup> est lié au site I de la calbindine D<sub>9k</sub> dont la simulation de DM a été effectuée à partir de la structure RMN D dont le Ca<sup>2+</sup> dans le site II a été supprimé*

S<sub>II</sub> *Etat où un Ca<sup>2+</sup> est lié au site II de la calbindine D<sub>9k</sub>*

Ser *Sérine*

TIP3P *Transferable Interaction Potential for 3 Particles (modèle de molécule d'eau)*

Tyr *Tyrosine*

V61A *Alanine en position 61 plutôt que valine*

Val *Valine*

WHAM *Weighted Histogram Analysis Method*

Y13F *Phénylalanine en position 13 plutôt que tyrosine*

# CHAPITRE 1

## Introduction

### 1.1 Le $\text{Ca}^{2+}$ et les protéines auxquelles il se lie

Pour la plupart des gens, l'idée du calcium évoque surtout la formation et l'entretien du cartilage, des os et des dents. Pourtant, le calcium dans sa forme ionique ( $\text{Ca}^{2+}$ ) sert aussi à régler bon nombre d'activités cellulaires où il agit à titre de messenger secondaire, transmettant les signaux reçus à la surface cellulaire jusqu'à leur destination dans le milieu intracellulaire.<sup>1</sup> Concrètement, le  $\text{Ca}^{2+}$  contribue, entre autres, à la coagulation du sang, la contraction musculaire, le métabolisme du glycogène, la synthèse des protéines ainsi que la croissance et la différenciation des cellules.<sup>2</sup>

La concentration intracellulaire normale du  $\text{Ca}^{2+}$  libre est de l'ordre de  $0.1 \mu\text{M}$  et une élévation de cette concentration à  $1 \mu\text{M}$  initie les activités cellulaires énumérées au paragraphe précédent.<sup>3</sup> Le rôle du  $\text{Ca}^{2+}$ , qui se joue à l'intérieur d'un mince intervalle à de faibles concentrations, est orchestré par un mécanisme de contrôle sensible, précis et efficace qui régularise ou transmet l'information contenue dans les variations de la concentration de  $\text{Ca}^{2+}$  intracellulaire. Ce mécanisme de contrôle s'exerce principalement *via* des protéines qui lient le  $\text{Ca}^{2+}$ , et que l'on regroupe sous l'appellation de *superfamille des calmodulines*. Cette superfamille se décompose en familles dont chacune rassemble des membres ayant des structures et fonctions similaires. Parmi ces familles, mentionnons les calmodulines, les troponines C, les parvalbumines et les calbindines  $\text{D}_{9\text{k}}$ . Ces protéines

sont dites régulatrices si leur fonction consiste à transmettre à d'autres protéines le signal physiologique encodé dans la variation de la concentration de  $\text{Ca}^{2+}$  (calmodulines, troponines C); la liaison du  $\text{Ca}^{2+}$  à ces protéines s'accompagne généralement d'un changement majeur de leur conformation. Si leur fonction consiste à ramener la concentration de  $\text{Ca}^{2+}$  à la normale, elles font partie de la classe des tampons (parvalbumines, calbindines  $\text{D}_{9k}$ ) et présentent très peu de changements structuraux suite à la liaison de  $\text{Ca}^{2+}$ .

La sensibilité, la précision et l'efficacité avec lesquelles les protéines qui lient le  $\text{Ca}^{2+}$  accomplissent leurs fonctions sont tributaires d'une unité structurale constituée par un site de liaison au  $\text{Ca}^{2+}$  contigu à deux hélices- $\alpha$ . Cette structure caractéristique a été nommée *main EF* (*EF-hand*) par Kretsinger et Nockolds<sup>4</sup> à l'origine pour la parvalbumine. Ces derniers ont établi un parallèle entre cette séquence (hélice- $\alpha$  E-site-hélice- $\alpha$  F) et l'orientation de l'index (hélice- $\alpha$  E) et du pouce (hélice- $\alpha$  F) tel qu'illustré à la figure 1. Mais c'est l'assemblage de deux mains EF qui constitue l'unité fonctionnelle de ces protéines, puisqu'elles interagissent de façon coopérative et leur donnent les qualités requises de sensibilité, sélectivité et efficacité; une protéine donnée peut comporter une ou plusieurs de ces unités fonctionnelles. Ces concepts de main EF, de leur assemblage et leurs interactions coopératives sont détaillés dans les sections suivantes.

## 1.2 La calbindine $\text{D}_{9k}$

La calbindine  $\text{D}_{9k}$  a été isolée pour la première fois en 1973 dans de l'intestin de boeuf par Fullmer et Wasserman,<sup>5</sup> d'où l'appellation ICaBP (Intracellular Calcium Binding Protein) donnée initialement à cette protéine. Mais puisqu'elle a depuis été isolée dans plusieurs autres tissus, son appellation a été remplacée par un terme plus générique, soit la calbindine  $\text{D}_{9k}$ . Des tissus où la calbindine  $\text{D}_{9k}$  est présente, on note l'intestin, le rein, le cartilage, les dents et les os où l'expression du gène de la calbindine  $\text{D}_{9k}$  peut être contrôlée par une forme active

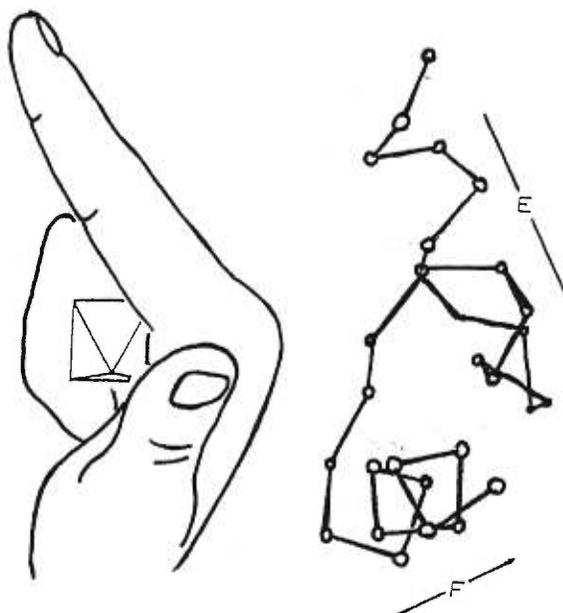


FIGURE 1. Structure de la main EF du domaine C-terminal de la parvalbumine et analogie avec celle d'une main droite qui est à l'origine de l'appellation. <sup>4</sup>

de la vitamine D,<sup>6</sup> l'utérus où l'expression de son gène peut être contrôlée par l'œstrogène,<sup>7, 8</sup> ainsi que le placenta et le poumon. Jusqu'en 1995, la calbindine D<sub>9k</sub> n'avait été isolée que dans des tissus mammifères, mais Zanello et coll.<sup>9</sup> ont démontré sa présence aussi dans les tissus de poulet. La plupart des fonctions spécifiques accomplies par la calbindine D<sub>9k</sub> restent à confirmer. En plus de son rôle de tampon appuyé par des preuves récentes,<sup>10</sup> notons qu'elle serait impliquée dans le transport intestinal du Ca<sup>2+</sup> de la membrane en bordure en brosse jusqu'à la membrane basolatérale<sup>11</sup> et dans la minéralisation des os et des dents.<sup>6</sup>

La calbindine D<sub>9k</sub> est le plus petit membre connu de la *superfamille des calmodulines* avec 78 acides aminés dans sa forme majeure et 75 acides aminés dans sa forme A mineure, i.e. sans le groupement N-terminal Ac-Ser-Ala-Lys;<sup>12</sup> sa masse moléculaire est de 8.7 kda d'où l'appellation 9k. La structure tridimensionnelle de la forme A mineure provenant de l'intestin de boeuf a été déterminée pour la première fois en 1981<sup>13</sup> par diffraction des rayons X à une résolution de 2.3 Å.

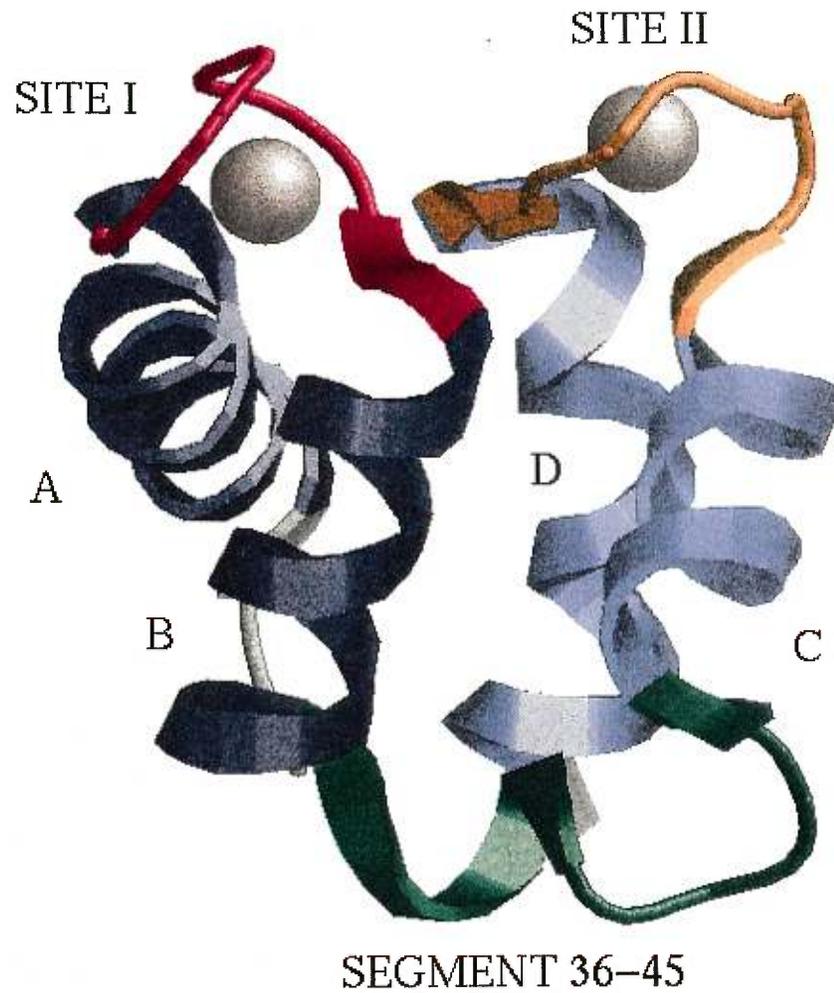


FIGURE 2. Illustration schématique de la calbindine D<sub>9k</sub> où sont représentés les chaînes principales des hélices, des sites et du segment 36-45.

La figure 2 donne une représentation tridimensionnelle de la protéine.

La calbindine D<sub>9k</sub> étudiée dans cette thèse est une forme mutante de la protéine originale, c'est-à-dire que les structures expérimentales disponibles dans les banques de données ont une glycine en position 43 plutôt qu'une proline (mutant P43G). La raison à l'origine de cette mutation est que la protéine sauvage hydratée présente un isomérisme *cis-trans* au lien entre la glycine 42 et la proline 43, ayant comme conséquence le dédoublement du signal spectroscopique dans les expériences de RMN<sup>14</sup> et de diffraction des rayons X.<sup>15</sup> Le mutant P43G ne présente pas d'isomérisme à ce lien et n'engendre qu'un seul signal. Enfin, il a été démontré que le mutant P43G n'affiche que quelques différences structurales et dynamiques négligeables par rapport à la forme originale de la protéine.<sup>14</sup>

La calbindine D<sub>9k</sub> possède deux sites de liaison pour le Ca<sup>2+</sup> dont chacun fait partie d'une main EF. Le site de l'extrémité C-terminale (appelé site II dans le reste du texte) est constitué de 12 acides aminés, et correspond au site standard, commun aux protéines de la *superfamille des calmodulines*. Les ligands du Ca<sup>2+</sup> sont invariablement des atomes d'oxygène dont cinq appartiennent aux chaînes latérales, un appartient à la chaîne principale et un dernier à une molécule d'eau. Le site de l'extrémité N-terminale, ou site I dans le reste du texte, compte pour sa part 14 acides aminés et constitue une exception à la règle que l'on retrouve aussi chez les protéines de la famille S-100. L'unité structurale associée à ce site a conséquemment été renommée pseudo main EF, pour laquelle les ligands du Ca<sup>2+</sup> incluent quatre atomes d'oxygène provenant de la chaîne principale, deux de chaînes latérales et un d'une molécule d'eau. Plusieurs propriétés du site I s'avèrent donc différentes de celles du site II, puisque les ligands qui les constituent proviennent majoritairement soit de la chaîne principale ou des chaînes latérales. Malgré cette différence, la géométrie des ligands autour des Ca<sup>2+</sup> demeure *grosso modo* la même: cinq des ligands forment approximativement un pentagone à l'intérieur duquel se trouve Ca<sup>2+</sup>; les deux autres ligands sont situés hors du plan,

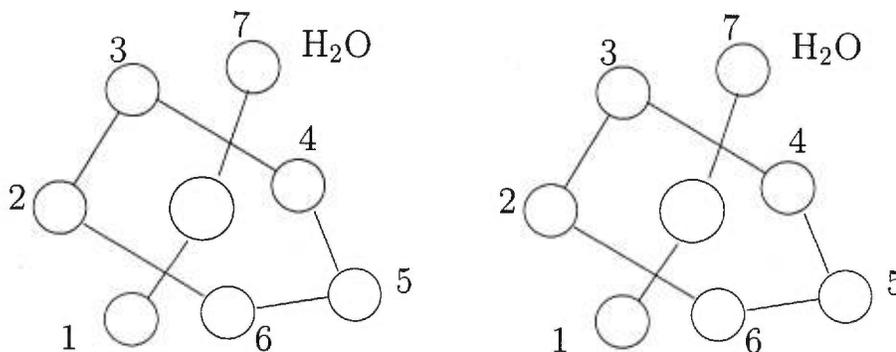


FIGURE 3. Configuration (en stéréo) des ligands autour du  $\text{Ca}^{2+}$  lorsque dans un ou l'autre des sites de la calbindine  $\text{D}_{9k}$ . Chaque point représente un atome d'oxygène (ligand) se trouvant sur une des deux pyramides à base pentagonale mitoyenne. Les chiffres de 1 à 7 indiquent le numéro du ligand dont la nature est donnée à la figure 4 pour les deux sites.

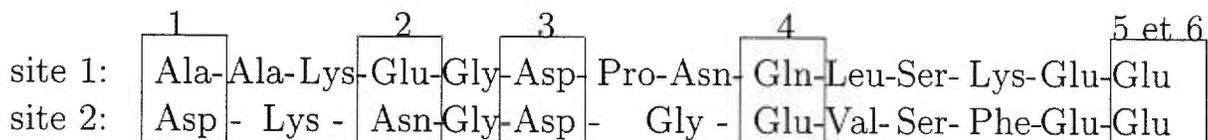


FIGURE 4. Séquences des acides aminés qui composent les sites I et II et identification des atomes liés au  $\text{Ca}^{2+}$  (voir figure 3).

un au-dessus et un au-dessous du  $\text{Ca}^{2+}$ , le tout formant ainsi deux pyramides à base pentagonale mitoyenne, tel qu'illustré à la figure 3. La séquence des deux sites de la calbindine  $\text{D}_{9k}$  ainsi que les différents atomes liés au  $\text{Ca}^{2+}$  sont illustrés à la figure 4. Les sites de la calbindine  $\text{D}_{9k}$  possèdent une grande affinité pour le  $\text{Ca}^{2+}$  (à 150 mM KCl,  $\log K_1 = 6.3$  et  $\log K_2 = 6.5$ ; à 0 mM KCl,  $\log K_1 = 8.2$ ,  $\log K_2 = 8.6$ )<sup>16, 17</sup> qu'elle lie de façon coopérative, c'est-à-dire que la présence d'un premier ion dans un des sites précipite la liaison d'un ion dans l'autre site. Notons enfin que les acides aminés chargés sont répartis presque exclusivement dans la région des sites et que la protéine possède une charge totale de  $-7e$  dans sa forme A mineure, c'est-à-dire sans le groupement N-terminal Ac-Ser-Ala-Lys

(sans ions).

### 1.3 Régulation allostérique et coopérativité

Les cellules vivantes utilisent des modes de communication et de régulation intra et extracellulaires dont les protéines sont les principaux vaisseaux. Ces protéines transmettent les messages cellulaires *via* des cascades d'interactions dont l'efficacité dépend de leur capacité à lier le ou les prochain(s) substrat(s) dans la cascade. La diversité de ces mécanismes est considérable, mais un point commun à beaucoup d'entre eux est d'effectuer des transitions allostériques, c'est-à-dire que l'attachement d'un premier substrat à un site induit des changements de conformations et conséquemment d'affinité des autres sites pour leur substrat. En particulier, dans le cas des protéines de la *superfamille des calmodulines*, la liaison d'un premier  $\text{Ca}^{2+}$  à un site augmente l'affinité d'un second site pour le  $\text{Ca}^{2+}$ .

Le modèle des transitions allostériques, tel que postulé à l'origine par Monod, Wyman et Changeux en 1963 et 1965<sup>18, 19</sup> ne concernait que des monomères identiques associés formant un oligomère, de sorte que le système possède au moins un axe de symétrie. La liaison d'un substrat au site d'un monomère entraînait une réaction concertée de changements conformationnels des autres monomères. Les propriétés catalytiques de la protéine s'en trouvaient activées ou inhibées *via* un changement d'affinité des sites des autres monomères. Dans leur modèle, l'élément de symétrie devait être conservé suite à ces changements structuraux, impliquant qu'il n'y avait que deux états conformationnels accessibles à l'oligomère et que les monomères devaient demeurer identiques entre eux. Un peu plus tard, ce modèle a été généralisé par Koshland et coll.<sup>20</sup> qui ont établi que plusieurs modèles simples étaient aptes à reproduire les courbes de saturation expérimentales des liaisons de substrats des systèmes allostériques. Même si leurs modèles ne s'en tenaient qu'aux oligomères, il n'était alors plus essentiel que les monomères qui les

composent soient identiques. Ce n'est qu'avec Weber en 1972<sup>21, 22</sup> qu'un modèle sera proposé pour des monomères pouvant lier des substrats de natures identiques ou différentes. Ce modèle exprime les transitions allostériques en termes de couplages d'énergies libres entre les ligands et est à la base des développements formels présentés à la section suivante.

## 1.4 La coopérativité chez la calbindine $D_{9k}$

### 1.4.1 Formulation

Le cas particulier de la calbindine  $D_{9k}$  est celui d'un monomère non-symétrique possédant deux sites non-identiques où s'attachent, de façon coopérative, deux ions de même nature ( $Ca^{2+}$ ) à deux sites séparés par environ 12 Å. Concrètement, la coopérativité dans la liaison du  $Ca^{2+}$  à la calbindine  $D_{9k}$  rend la fonction physiologique plus efficace en amincissant l'intervalle de concentration à l'intérieur duquel elle s'effectue, comme le montre la figure 5. L'ordonnée de cette courbe est le nombre moyen de  $Ca^{2+}$  lié à la protéine, et son expression est donnée par:

$$\langle N_{Ca^{2+} \text{ lié}} \rangle = 2P(D) + 1P(S_I) + 1P(S_{II}) + 0P(A) \quad (1.1)$$

où  $P(D)$ ,  $P(S_I)$ ,  $P(S_{II})$  et  $P(A)$  sont les probabilités de retrouver les états liés à deux  $Ca^{2+}$  (D), à un  $Ca^{2+}$  dans le site I ( $S_I$ ) ou II ( $S_{II}$ ) et à l'état A (nul). Ces probabilités individuelles sont données par:

$$P(D) = \frac{[D]}{[D] + [S_I] + [S_{II}] + [A]}, \quad (1.2)$$

$$P(S_I) = \frac{[S_I]}{[D] + [S_I] + [S_{II}] + [A]},$$

$$P(S_{II}) = \frac{[S_{II}]}{[D] + [S_I] + [S_{II}] + [A]}$$

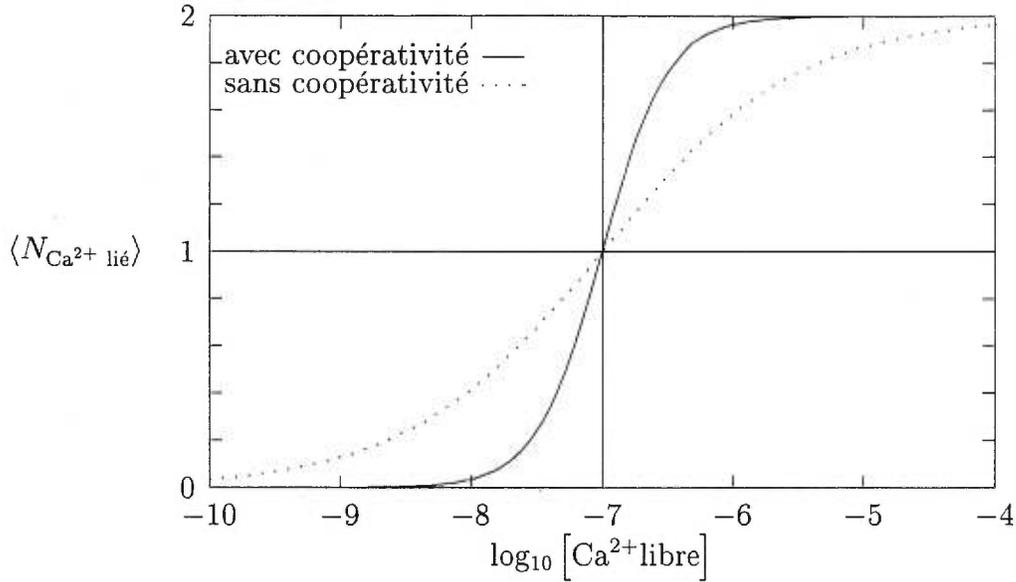


FIGURE 5. Courbe du nombre moyen de  $\text{Ca}^{2+}$  lié aux deux sites de la calbindine  $\text{D}_{9k}$  en fonction de la concentration libre dans le milieu intracellulaire en présence (tirets) et en absence (ligne continue) de coopérativité <sup>23</sup>.

où  $[A]$ ,  $[S_I]$ ,  $[S_{II}]$  et  $[D]$  représentent respectivement la concentration de calbindine  $\text{D}_{9k}$  dans l'état apo ( $A$ ; sans  $\text{Ca}^{2+}$ ), dans les états liés à un seul  $\text{Ca}^{2+}$  soit dans le site I ( $S_I$ ) ou dans le site II ( $S_{II}$ ), et dans l'état lié à deux  $\text{Ca}^{2+}$  ( $D$ ). L'équation 1.1 devient, après remplacement des équations 1.2:

$$\langle N_{\text{Ca}^{2+} \text{ lié}} \rangle = \frac{2[D] + [s]}{[D] + [s] + [A]}, \quad (1.3)$$

où  $[s] = [s_I] + [s_{II}]$ .

La figure 6 illustre un modèle à quatre états pour la calbindine  $\text{D}_{9k}$  qui convient aussi à n'importe quel autre système équivalent.<sup>16</sup> Chacun des passages d'un état à un autre est représenté par une constante microscopique de liaison ainsi qu'une différence d'énergie libre associée  $\Delta G_i = -kT \ln k_i$ . Notons que expérimentalement,<sup>16</sup>  $1 < k_{II}/k_I < 3$  selon les conditions de salinité, de pH, etc... Le processus global est défini en termes de constantes macroscopiques de

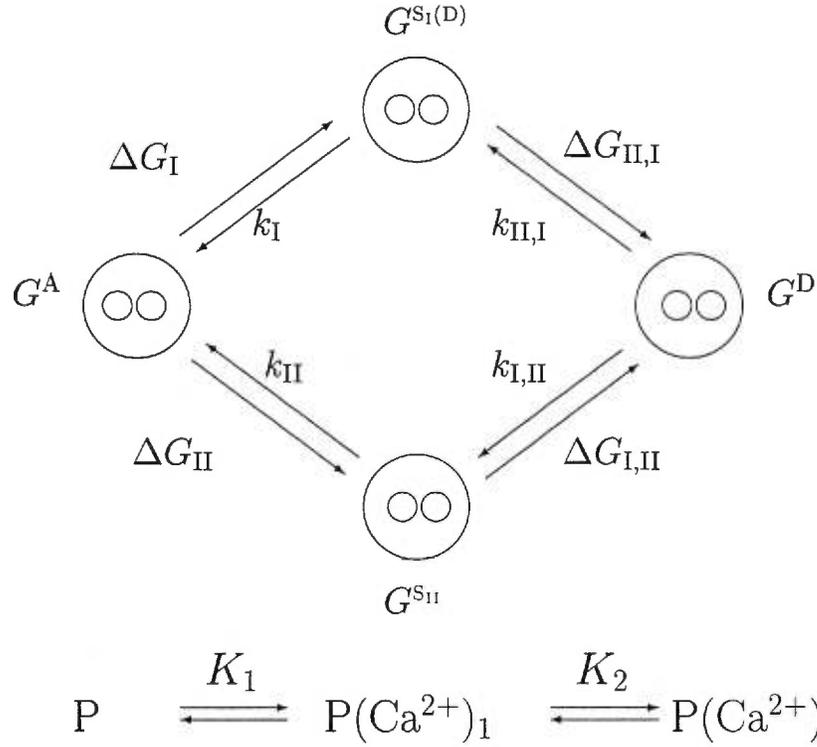


FIGURE 6. Modèle à quatre états du processus de liaison de deux  $\text{Ca}^{2+}$  à la calbindine  $\text{D}_{9k}$  où les  $k_I$ ,  $k_{II}$ ,  $k_{I,II}$  et  $k_{II,I}$  sont les constantes microscopiques de liaison, chacune étant spécifique à un site <sup>23</sup>. Les  $K_1$  et  $K_2$  sont les constantes macroscopiques de liaison; P signifie la protéine.

liaison  $K_1$  et  $K_2$ :

$$K_1 = \frac{[S_I + S_{II}]}{[A][\text{Ca}^{2+}]} \quad (1.4)$$

et

$$K_2 = \frac{[D]}{[S_I + S_{II}][\text{Ca}^{2+}]} \quad (1.5)$$

où  $[\text{Ca}^{2+}]$  est la concentration de  $\text{Ca}^{2+}$  non-lié. Les constantes macroscopiques sont reliées aux constantes microscopiques par:

$$K_1 = k_I + k_{II} \quad (1.6)$$

et

$$K_2 K_1 = k_I k_{II,I} = k_{II} k_{I,II} \quad (1.7)$$

considérant les deux voies de liaisons possibles. Dans ce contexte, la coopérativité positive chez la calbindine D<sub>9k</sub> se définit comme une différence d'énergies libres pouvant emprunter deux voies différentes:

$$\begin{aligned}\Delta\Delta G^{\text{coop.}} &= \Delta G_{\text{I,II}} - \Delta G_{\text{I}} < 0 \\ &= -kT \ln \left( \frac{k_{\text{I,II}}}{k_{\text{I}}} \right)\end{aligned}\quad (1.8)$$

ou encore

$$\begin{aligned}\Delta\Delta G^{\text{coop.}} &= \Delta G_{\text{II,I}} - \Delta G_{\text{II}} < 0 \\ &= -kT \ln \left( \frac{k_{\text{II,I}}}{k_{\text{II}}} \right).\end{aligned}\quad (1.9)$$

En substituant les équations (1.6) et (1.7) (pour la voie de liaison où le Ca<sup>2+</sup> se lie d'abord au site II) dans l'équation (1.9) on obtient une expression pour la coopérativité en termes des constantes macroscopiques de liaison:<sup>16</sup>

$$\begin{aligned}\Delta\Delta G^{\text{coop.}} &= -kT \ln \left( \frac{4K_2}{K_1} \right) - kT \ln \left[ \frac{(\zeta + 1)^2}{4\zeta} \right] \\ &= -kT \ln \left( \frac{k_{\text{II,I}}}{k_{\text{II}}} \right)\end{aligned}\quad (1.10)$$

où  $\zeta = k_{\text{II}}/k_{\text{I}}$  est le rapport des affinités du Ca<sup>2+</sup> pour chacun des sites en posant que l'affinité du site I est tout au plus égale à celle du site II ( $\zeta \geq 1$ ).

C'est à partir de la première égalité de cette équation et des valeurs des constantes macroscopiques de liaison, obtenues elles-mêmes par des expériences de titrations au Ca<sup>2+</sup> (voir figure 5), que sont évaluées expérimentalement la coopérativité et l'affinité relative des deux sites pour le Ca<sup>2+</sup> dans le cas où le Ca<sup>2+</sup> se lie d'abord au site II. A une concentration physiologique d'ions (150mM KCl), on a estimé la coopérativité à environ  $-1.10$  kcal/mol avec une affinité relative des sites  $k_{\text{II}}/k_{\text{I}} \approx 1$ , et  $K_2 > K_1$ .<sup>17</sup> Une évaluation de la coopérativité minimale du système peut être obtenue en posant que  $\zeta = 1$ , c'est à dire que les deux voies sont égales sur le plan énergétique:

$$\Delta\Delta G_{\zeta=1}^{\text{coop.}} = -kT \ln \left( \frac{4K_2}{K_1} \right)\quad (1.11)$$

d'où on voit que lorsque  $\zeta = 1$ , une coopérativité positive se résume à avoir  $K_2 > K_1$ . Soulignons enfin que la stricte définition de la coopérativité pour notre cas se résume à l'augmentation de l'affinité d'un site pour son ligand associée à la présence d'un premier ligand dans l'autre site.

Avec ces derniers développements, il est possible d'exprimer le nombre moyen de  $\text{Ca}^{2+}$  lié,  $\langle N_{\text{Ca}^{2+} \text{ lié}} \rangle$ , en fonction du  $\Delta\Delta G^{\text{coop.}}$ . En substituant d'abord les définitions des  $K_1$  et  $K_2$  en termes des concentrations des différents états (équations 1.4 et 1.5) avec  $[s] = [s_I] + [s_{II}]$  dans l'équation 1.3, on obtient l'expression suivante:

$$\langle N_{\text{Ca}^{2+} \text{ lié}} \rangle = \frac{2K_1K_2[\text{Ca}^{2+}]^2 + K_1[\text{Ca}^{2+}]}{K_1K_2[\text{Ca}^{2+}]^2 + K_1[\text{Ca}^{2+}] + 1} \quad (1.12)$$

Puis, si l'on remplace les  $K_1$  et  $K_2$  par leur équivalent en termes des constantes microscopiques (équations 1.6 et 1.7), en même temps que l'on substitue l'expression  $\exp(-\beta\Delta\Delta G^{\text{coop.}}) = k_{II,I}/k_{II}$  (équation 1.9), on obtient:

$$\langle N_{\text{Ca}^{2+} \text{ lié}} \rangle = \frac{2k_i[\text{Ca}^{2+}] + 2k_i^2[\text{Ca}^{2+}]^2 \exp(-\beta\Delta\Delta G^{\text{coop.}})}{1 + 2k_i[\text{Ca}^{2+}] + k_i^2[\text{Ca}^{2+}]^2 \exp(-\beta\Delta\Delta G^{\text{coop.}})} \quad (1.13)$$

où  $k_i$  représente la constante microscopique  $k_I$  ou  $k_{II}$  puisqu'on a posé que les deux voies de liaison étaient équivalentes ( $k_I \approx k_{II}$ )

#### 1.4.2 Facteurs identifiés comme jouant un rôle dans la coopérativité

Certains facteurs clefs impliqués dans le mécanisme de la coopérativité ont déjà été identifiés,<sup>16, 17, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37</sup> mais des données quantitatives quant aux différentes contributions ne sont pas encore disponibles. On peut diviser ces contributions selon que la coopérativité implique un changement de structure ou de dynamique de la calbindine  $D_{9k}$  et se demander si des interactions électrostatiques ou hydrophobes sont à l'origine de ces changements. Il faut mentionner qu'à cause de la coopérativité, les deux états où un seul  $\text{Ca}^{2+}$  est lié à la protéine ne peuvent être isolés en conditions naturelles. La majorité des connaissances sur la coopérativité proviennent d'expériences faites sur un

modèle de l'état où la calbindine  $D_{9k}$  n'est liée qu'à un seul  $Ca^{2+}$ , soit au site II ( $S_{II}$ ),<sup>25, 33</sup> ou au site I ( $S_I$ )<sup>35, 36</sup> qui n'est disponible que depuis 1995. Ces modèles moléculaires sont détaillés au chapitre 4.

#### 1.4.2.1 Facteurs structuraux

En ce qui concerne l'aspect structural, il a été vérifié que la calbindine  $D_{9k}$  ne subit pas de changement conformationnel majeur lors du processus de liaison des deux  $Ca^{2+}$ .<sup>25</sup> Plutôt, de petites réorganisations surviennent un peu partout dans la protéine: le site II est davantage modifié que le site I, et une réorganisation subtile du reste de la protéine s'effectue.<sup>32, 33</sup> Les changements induits par la liaison du premier  $Ca^{2+}$  sont plus importants que ceux induits par la liaison du second. De plus, les états  $S_I$ ,<sup>35, 36</sup>  $S_{II}$  et  $D$ <sup>25, 32, 33</sup> sont très similaires entre eux et tous trois différents de l'état A, suggérant que la liaison du premier ion est associée avec l'étape la moins favorable du point de vue d'un changement d'énergie libre, puisque la réorganisation suite à la liaison du second  $Ca^{2+}$  est de moindre envergure. Un élément structural possiblement impliqué dans la coopérativité consiste en deux ponts hydrogènes formant un mini feuillet- $\beta$  qui relie les deux sites (voir figure 7). Ces ponts hydrogènes pourraient être responsables de la transmission du signal indiquant au site resté vide qu'un ion vient de se lier à l'autre site, et initier les modifications d'affinité subséquentes. Bien que ces ponts hydrogènes soient pleinement formés dans les quatre états A,  $S_I$ ,  $S_{II}$  et D,<sup>28, 38, 39</sup> il apparaît qu'ils soient graduellement renforcés suite à la liaison du  $Ca^{2+}$ .<sup>25</sup> Il a donc été à ce jour difficile d'évaluer avec certitude la contribution de ces ponts hydrogènes.

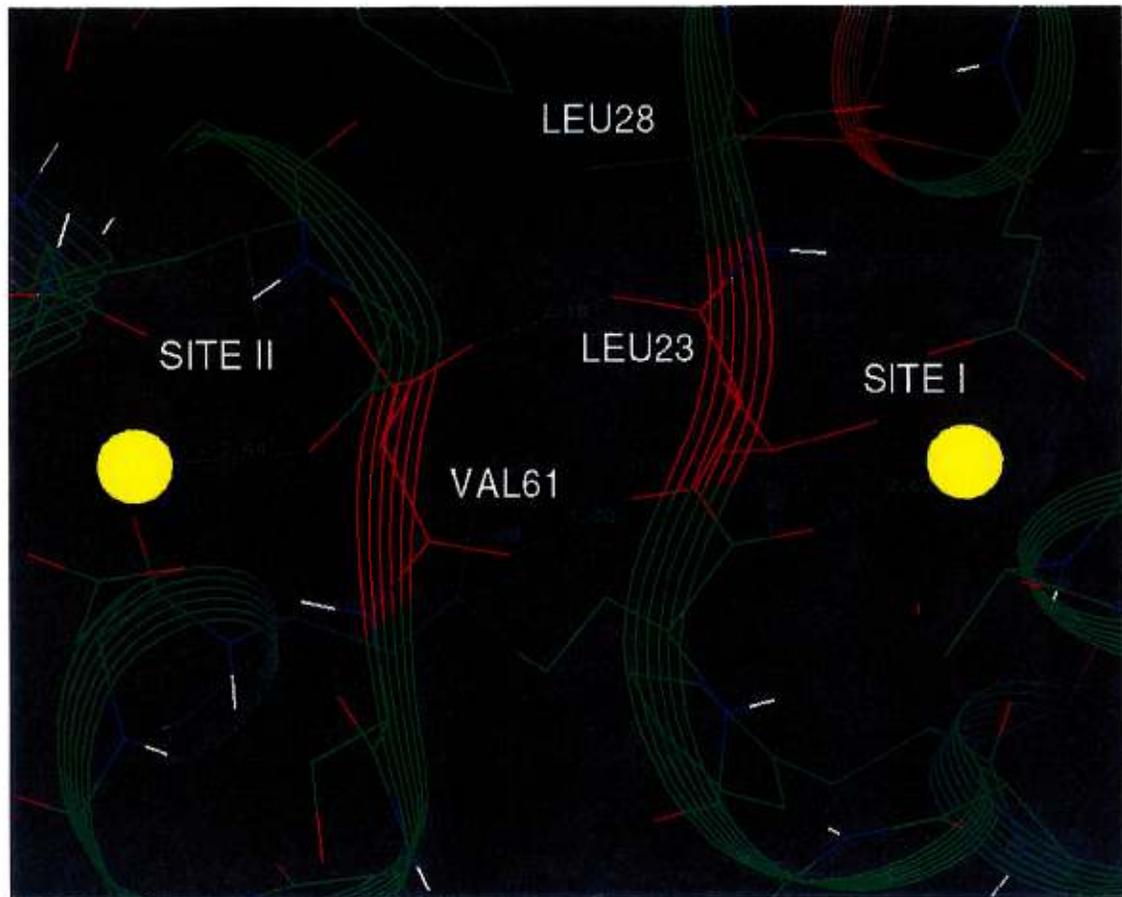


FIGURE 7. Ponts hydrogènes reliant les acides aminés 23 à 61 qui appartiennent respectivement au site I et au site II.

#### 1.4.2.2 Facteurs dynamiques

Des expériences de RMN<sup>24, 28, 33, 39</sup> ont permis de constater que les changements de dynamique entraînés par la liaison du premier  $\text{Ca}^{2+}$  sont plus grands que ceux attribués à la liaison du second. Les fluctuations dans la protéine (entropie) doivent donc être considérées comme jouant un rôle important dans la coopérativité. Dans le même ordre d'idées, Cooper et Dryden<sup>40</sup> avaient déjà proposé en 1984 un modèle thermostatique qui exprime la coopérativité positive en termes

de changements d'énergies libres imputables principalement à des changements de dynamique dans une protéine. Le détail des changements de structure et de dynamique relatifs à la liaison du  $\text{Ca}^{2+}$  et observés expérimentalement sont rappelés dans le chapitre 4.

#### 1.4.2.3 Influence de l'eau

Le rôle des molécules d'eau dans la déshydratation des sites et autres parties de la protéine suite à la liaison successive des deux  $\text{Ca}^{2+}$  est fort probablement un facteur jouant un rôle dans la coopérativité. Cet aspect a déjà été étudié pour l'hémoglobine<sup>41, 42</sup> où il s'est avéré que l'augmentation de la surface accessible au solvant suite à la liaison de l'oxygène était corrélée avec une augmentation du nombre de molécules d'eau liées à la protéine. Une telle étude n'existe pas pour la calbindine  $\text{D}_{9k}$  mais il est fort probable que les changements structuraux, surtout dans les sites, soient accompagnés d'un changement du profil d'hydratation de la protéine. Quant au rôle spécifique joué par les molécules d'eau qui occupent des positions stratégiques dans la calbindine  $\text{D}_{9k}$ , il a été proposé<sup>37</sup> que trois molécules d'eau dont deux liées aux  $\text{Ca}^{2+}$ , ainsi qu'une autre liée par de forts ponts hydrogènes à deux acides aminés reliant les deux mains EF, stabiliseraient la structure, ces trois molécules d'eau occupant la même position significativement plus longtemps que les autres.

#### 1.4.3 Types d'interactions impliquées dans la coopérativité

Plusieurs observations ont été faites quant aux types d'interactions impliquées dans les changements de structure et de dynamique: notamment, le rôle des interactions électrostatiques lors de la liaison du  $\text{Ca}^{2+}$  a été étudié,<sup>17, 31, 43, 44</sup> et il s'est avéré que les charges négatives à la surface de la calbindine  $\text{D}_{9k}$  (par des mutations d'acides aminés), ainsi que la concentration d'ions dans le milieu où

elle se trouve, peuvent modifier (réduire) les constantes macroscopiques  $K_1$  et  $K_2$  et par conséquent la coopérativité. De plus, ces acides aminés chargés n'ont pas besoin d'être directement liés au  $\text{Ca}^{2+}$  pour influencer son processus de liaison. Notons enfin que la coopérativité est plus grande à faible concentration d'ions, soit que  $\Delta\Delta G^{\text{coop.}} \approx -1.10$  kcal/mol à 150 mM KCl et  $\Delta\Delta G^{\text{coop.}} \approx -1.84$  kcal/mol à 50 mM KCl.<sup>17</sup>

En plus des charges négatives réparties asymétriquement à la surface et autour des sites, la protéine comporte un coeur d'acides aminés hydrophobes assez compact, constitué des quatre hélices- $\alpha$  reliées par un tricot de ponts hydrogènes plus ou moins serré. A partir d'expériences de mesures du taux d'échange des protons appartenant à des groupements amides avec les molécules d'eau du milieu, Skelton et coll.<sup>28</sup> ont proposé que les effets associés à la liaison d'un premier  $\text{Ca}^{2+}$  au site II pourraient être transmis *via* les ponts hydrogènes reliant les deux sites. Par exemple, la réduction de la flexibilité du site II se propagerait jusque dans la main EF du site I et y entraînerait aussi une réduction des fluctuations. Ceci permettrait d'obtenir des ponts hydrogènes plus forts et de modifier la configuration des chaînes latérales afin d'obtenir des interactions plus favorables.

#### 1.4.4 Modèle proposé

La figure 8, inspirée de Falke et coll.,<sup>45</sup> illustre le modèle proposé par M. Akke<sup>24, 25, 46</sup> auquel sont ajoutées quelques données supplémentaires provenant de travaux subséquents. Le modèle a été établi d'après des données expérimentales de RMN pour lesquelles le  $\text{Ca}^{2+}$  se lie d'abord au site II. Dans l'état A, la protéine est très flexible et explore un grand espace conformationnel dans lequel un nombre restreint de conformations possèdent une grande affinité pour le  $\text{Ca}^{2+}$ . Notons que, de façon générale, la chaîne principale du site II est plus flexible que celle du site I dont l'affinité pour le  $\text{Ca}^{2+}$  est légèrement inférieure, si elle est différente. Suite à la liaison du  $\text{Ca}^{2+}$  au site II, les deux mains EF voient le détail de leur structure

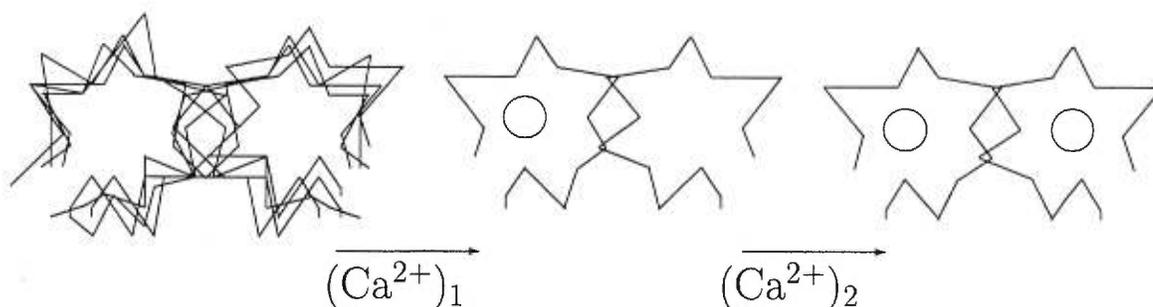


FIGURE 8. Modèle de coopérativité chez la calbindine  $D_{9k}$  proposé par M. Akke; <sup>25</sup> le dessin est tiré de Falke et coll. <sup>45</sup>

modifié et se retrouvent limitées à un nombre plus restreint de conformations dont la structure correspond à l'état occupé par le  $Ca^{2+}$ , et ce même pour le site I où il n'y a pas encore de  $Ca^{2+}$ . On peut donc dire que la liaison d'un premier  $Ca^{2+}$  au site II prépare le site I pour la venue du second  $Ca^{2+}$  et que cette préparation s'effectue plutôt au niveau de la réduction de l'espace conformationnel, puisque le site II subit une plus grande modification de sa structure que le site I qui est en grande partie préformé.

L'étude expérimentale du modèle de l'état  $S_I$  <sup>35, 36</sup> a permis de corroborer plusieurs aspects de ce modèle proposé pour la coopérativité. La venue du  $Ca^{2+}$  dans le site I induit des changements de structure et de dynamique dans toute la protéine. En fait, la structure de cet état est très similaire à celle de l'état D. Généralement, les conséquences de la liaison du  $Ca^{2+}$  au site I sont similaires à celles induites par la liaison du  $Ca^{2+}$  au site II. Même si ces résultats n'éclaircissent pas le mécanisme des deux voies de liaison du  $Ca^{2+}$ , ils indiquent que la liaison d'un seul  $Ca^{2+}$  à un site ou à l'autre transforme déjà la calbindine  $D_{9k}$  vers une structure très similaire à celle de l'état D, ce qui s'accorde avec la définition même de la coopérativité.

## 1.5 But de la thèse

La coopérativité chez la calbindine  $D_{9k}$ , et en général, ne se présente certes pas comme un phénomène facile à expliquer. L'absence de changement majeur de conformation de la calbindine  $D_{9k}$  rend la tâche d'autant plus difficile qu'elle nous place devant un phénomène résultant de changements subtils, répartis asymétriquement dans toute la protéine. L'étude de la nature de ces changements et de la façon dont ils sont reliés, bref les bases moléculaires et atomiques, nécessite sans contredit un modèle atomique détaillé du système.

Cette thèse vise donc l'étude théorique de divers aspects des conséquences de la liaison des ions  $Ca^{2+}$  et de la coopérativité chez la calbindine  $D_{9k}$  par des méthodes de modélisation moléculaire. Ces méthodes permettront de suivre l'évolution temporelle du système, d'obtenir une analyse au niveau atomique et d'évaluer les différentes contributions qui en déterminent l'allure. En comparant certains de nos résultats avec leurs homologues expérimentaux il est possible d'évaluer la méthode de modélisation utilisée lors des simulations et éventuellement d'aller au-delà de ce qui est déjà connu.

## 1.6 Plan de la thèse

Les méthodes utilisées sont exposées de façon détaillée au chapitre 2, suivi du chapitre 3 où les paramètres des interactions  $Ca^{2+}$ -ligand sont déterminés et où la qualité de ces interactions est évaluée. Le chapitre 4 propose, à l'aide de la méthode présentée au chapitre 2 et des paramètres développés au chapitre 3, une étude comparative de la structure et de la dynamique des quatre états du modèle de la calbindine  $D_{9k}$ . La qualité du modèle est aussi évaluée. Au chapitre 5, un modèle analytique propose une interprétation mécanique quant à l'origine des fluctuations (flexibilité) des vecteurs dipolaires NH pour l'état D de la calbindine

$D_{9k}$ . Puis, au chapitre 6, à partir d'une méthode de perturbations permettant les calculs de différences d'énergies libres associées à une mutation, l'effet de la substitution d'une valine en alanine sur la coopérativité est calculée et examinée en détail. Dans tous les cas où des données expérimentales sont disponibles, des comparaisons avec les valeurs calculées sont effectuées.

## CHAPITRE 2

### Les méthodes

Plusieurs méthodes de calculs sont utilisées lors de cette thèse. Elles sont ici présentées dans l'ordre de leur apparition dans les chapitres subséquents et auxquels il faut se référer pour obtenir de l'information concernant les protocoles de calculs spécifiques aux systèmes étudiés.

#### 2.1 Les calculs *ab initio* <sup>47, 48</sup>

Les méthodes *ab initio* traitent les systèmes moléculaires à partir des postulats de la mécanique quantique et consistent à trouver une solution à l'équation de Schrödinger multiélectronique. Bien que s'inspirant des principes premiers, ces méthodes reposent sur l'approximation de Born-Oppenheimer dont l'interprétation phénoménologique se résume à dire que la vitesse des noyaux est négligeable devant celle des électrons, ou que les électrons peuvent être traités dans un champ des noyaux fixes, c'est-à-dire que leurs positions sont traitées comme des paramètres dans l'équation aux valeurs propres qui est donnée par:

$$H^{\text{BO}}\Phi^{\text{BO}}(\mathbf{r}_1, \dots, \mathbf{r}_n; \mathbf{R}_1, \dots, \mathbf{R}_N) = E^{\text{BO}}(\mathbf{R}_1, \dots, \mathbf{R}_N)\Phi^{\text{BO}}(\mathbf{r}_1, \dots, \mathbf{r}_n; \mathbf{R}_1, \dots, \mathbf{R}_N) \quad (2.1)$$

où BO signifie Born-Oppenheimer et  $\mathbf{r}_i$  et  $\mathbf{R}_i$  sont respectivement les coordonnées des  $n$  électrons et des  $N$  noyaux;  $\Phi^{\text{BO}}$  est la fonction d'onde pour les électrons en fonction des coordonnées des noyaux et  $E^{\text{BO}}$  est l'énergie propre (potentielle) des électrons en fonction des coordonnées des noyaux. Le coeur de ces calculs consiste

à trouver les valeurs et fonctions propres de l'hamiltonien dit multiélectronique  $H^{\text{BO}}$  qui est l'opérateur d'énergie interne totale pour la molécule (en omettant l'énergie cinétique des noyaux).<sup>49</sup> L'hamiltonien est défini par  $H^{\text{BO}} = T^{\text{BO}} + V^{\text{BO}}$  où  $T^{\text{BO}}$  est l'opérateur d'énergie cinétique électronique et  $V^{\text{BO}}$  est l'énergie potentielle électrostatique pour toutes les particules. Etant donné que ce sont les systèmes moléculaires à  $N$  noyaux et  $n$  électrons qui nous intéressent, les opérateurs d'énergie cinétique et d'énergie potentielle sont dans ce contexte définis par

$$T^{\text{BO}} = - \sum_{i=1}^n \left( \frac{\hbar^2}{2m_e} \right) \nabla_i^2 \quad (2.2)$$

où  $m_e$  est la masse d'un électron,  $\hbar = h/(2\pi)$ ,  $h$  étant la constante de Planck, et

$$V^{\text{BO}} = - \sum_{\alpha=1}^N \sum_{i=1}^n \frac{Z_{\alpha} e^2}{4\pi\epsilon_0 r_{i\alpha}} + \sum_{i<j}^n \frac{e^2}{4\pi\epsilon_0 r_{ij}} + \sum_{\alpha<\beta}^N \frac{Z_{\alpha} Z_{\beta} e^2}{4\pi\epsilon_0 \mathbf{R}_{\alpha\beta}} \quad (2.3)$$

où  $Z_{\alpha}$  et  $Z_{\beta}$  sont les nombres de protons pour les noyaux  $\alpha$  et  $\beta$ ,  $e$  est la charge électronique,  $r_{i\alpha}$  est la distance noyau-électron,  $r_{ij}$  est la distance entre les électrons  $i$  et  $j$ ,  $\mathbf{R}_{\alpha\beta}$  est la distance entre les noyaux  $\alpha$  et  $\beta$ , et  $\epsilon_0$  est la permittivité du vide. Il n'existe pas de solution analytique aux calculs des systèmes moléculaires multiélectroniques. Comme il est expliqué plus bas, une solution à ce problème passe par la méthode du champ auto-cohérent appliquée à l'équation de Hartree-Fock, elle-même une approximation à l'équation (de Schrödinger) multiélectronique dans l'approximation de Born-Oppenheimer (voir équation 2.1).

La méthode Hartree-Fock consiste à assigner une forme fonctionnelle particulière à la fonction d'onde qui représente les  $n$  électrons d'un atome, soit un produit antisymétrisé de fonctions monoélectroniques:

$$|\Phi_{\text{élec}}^{\text{HF}}\rangle = A |\chi_1 \chi_2 \cdots \chi_{n-1} \chi_n\rangle \quad (2.4)$$

où  $\chi_i$  est une spin-orbitale, c'est-à-dire une fonction d'onde monoélectronique qui décrit à la fois la distribution spatiale et le spin de l'électron. L'opérateur  $A$  assure que la fonction  $\Phi_{\text{élec}}^{\text{HF}}$  a la propriété d'antisymétrie par rapport à toute permutation de deux électrons afin de respecter le principe d'exclusion de Pauli. Le principe

variationnel stipule que la "meilleure" fonction d'onde  $\Phi_{\text{élec}}^{\text{HF}}$  est celle pour laquelle on trouve un minimum à l'énergie  $E_0^{\text{HF}} = \langle \Phi_{\text{élec}}^{\text{HF}} | H | \Phi_{\text{élec}}^{\text{HF}} \rangle$ . La valeur minimale  $E_0^{\text{HF}}$  est alors la meilleure approximation (et supérieure) à la première valeur propre de  $H^{\text{BO}}$ , l'énergie fondamentale du système  $E_0^{\text{BO}}$  ( $E_0^{\text{HF}} > E_0^{\text{BO}}$ ). Comme les systèmes biologiques possèdent presque toujours des couches électroniques fermées, les spin-orbitales sont couplées:

$$\chi_i = \psi_i \alpha \quad (2.5)$$

ou

$$\chi_i = \psi_i \beta \quad (2.6)$$

où  $\psi_i$  est une orbitale spatiale commune à deux électrons et  $\alpha$  et  $\beta$  sont respectivement les fonctions de spin *up* et *down*. Toute la souplesse d'ajustement de la fonction  $\Phi_{\text{élec}}^{\text{HF}}$  afin d'obtenir une  $E_0^{\text{HF}}$  minimale se trouve dans les fonctions spatiales  $\psi_i$  comme nous le verrons plus bas.

L'équation de Hartree-Fock découle de la minimisation de  $E_0^{\text{HF}}$ , procédé qui devient algébriquement simple à cause de la forme simple de  $\Phi_{\text{élec}}^{\text{HF}}$ . Au bout de la dérivation, on arrive à une équation de Schrödinger monoélectronique où chaque électron se trouve dans le champ moyen de tous les autres électrons. L'équation de Hartree-Fock dans le cas de couches fermées est donc une équation aux valeurs propres exprimée par:

$$f(i)\psi_i = \epsilon_i \psi_i \quad (2.7)$$

où  $\epsilon_i$  est la valeur propre électronique obtenue par la méthode et  $\psi_i$  est la fonction propre associée. Dans ce contexte, l'hamiltonien de Hartree-Fock pour une molécule devient

$$\mathcal{H}^{\text{HF}} = \sum_{i=1}^n f(i) \quad (2.8)$$

avec son équation aux valeurs propres associée

$$\mathcal{H}^{\text{HF}} | \Phi_{\text{élec}}^{\text{HF}} \rangle = \sum_{i=1}^n \epsilon_i | \Phi_{\text{élec}}^{\text{HF}} \rangle. \quad (2.9)$$

Notons que  $\mathcal{H}^{\text{HF}} \neq H$  et que  $\sum_{i=1}^n \epsilon_i \neq E_0^{\text{HF}}$ . L'opérateur monoélectronique de Fock  $f(i)$  est défini par:

$$f(i) = h(i) + v^{\text{HF}}(i) \quad (2.10)$$

où

$$h(i) = - \left( \frac{\hbar^2}{2m_e} \right) \nabla_i^2 - \sum_{\alpha=1}^n \frac{Z_\alpha e^2}{4\pi\epsilon_0 r_i} \quad (2.11)$$

et

$$v^{\text{HF}}(i) = \sum_{a=1}^{n/2} 2J_a(i) - K_a(i). \quad (2.12)$$

$J_a(i)$  est appelé opérateur coulombien, tandis que  $K_a(i)$  est l'opérateur d'échange qui émerge de la nature antisymétrique de  $\Phi_{\text{elec}}^{\text{HF}}$ . Ces deux opérateurs sont des opérateurs intégraux. Le potentiel  $v^{\text{HF}}(i)$  est le potentiel moyen senti par l'électron  $i$  dû à la présence des autres électrons et dépend donc des spin-orbitales des autres électrons.

L'équation non-linéaire de Hartree-Fock doit être résolue par une méthode itérative. La méthode du champ auto-cohérent employée commence par l'introduction d'orbitales que l'on pose comme hypothèse de départ afin de calculer le champ moyen  $v^{\text{HF}}$  senti par chaque électron. Puis, on résoud l'équation aux valeurs propres (2.7) afin de calculer de nouvelles orbitales (fonctions propres) qui seront réutilisées pour recalculer le terme du champ moyen, et le cycle recommence jusqu'à ce que les orbitales introduites dans le champ moyen soient identiques, à un critère de convergence près, aux fonctions propres recalculées. La solution à l'équation de Hartree-Fock fournit un ensemble d'orbitales électroniques orthonormales  $\{\psi_i\}$  ainsi que les énergies orbitales  $\{\epsilon_i\}$  correspondantes.

De façon plus détaillée, la démarche consiste à développer chacune des  $n/2$  orbitales moléculaires  $\psi_i$  sur une base de  $k$  fonctions  $\{\phi_\mu\}$  connues tel que

$$\psi_i = \sum_{\mu=1}^k C_{\mu i} \phi_\mu, \quad i = 1, 2, 3, \dots, k > n/2 \quad (2.13)$$

où les  $C_{\mu i}$  sont des coefficients à déterminer et les  $\phi_\mu$  sont les fonctions de la base, typiquement des approximations à des orbitales atomiques. Les fonctions

gaussiennes sont parmi les plus pratiques. Plus l'ensemble des fonctions de base est vaste, plus on s'approche de la limite inférieure en énergie que peut fournir la méthode de Hartree-Fock ( $E_0^{\text{HF}} > E_0^{\text{BO}}$ ). La limite de Hartree-Fock correspond à utiliser une base complète, donc infinie.

L'équation de Hartree-Fock (2.7) développée dans sa forme intégrale et remaniée devient alors une équation matricielle (les équations de Roothaan) qui se résume à

$$\mathbf{FC} = \mathbf{SC}\epsilon \quad (2.14)$$

où  $\mathbf{F}$  est la matrice  $k \times k$  de l'opérateur de Fock ( $f(i)$ ) dans la base  $\{\phi_\mu\}$ ,  $\mathbf{C}$  est la matrice  $k \times k$  des coefficients  $C_{\mu i}$ ,  $\mathbf{S}$  est la matrice  $k \times k$  de recouvrement des fonctions de la base  $\{\phi_\mu\}$  qui émerge lorsque ces fonctions ne forment pas une base orthonormée, et enfin,  $\epsilon$  est la matrice diagonale des énergies orbitales.

Le principal défaut à l'approximation de l'équation de Hartree-Fock dont l'origine est la forme imposée à  $\Phi_{\text{elec}}^{\text{HF}}$  (équation 2.4), est de ne pas retenir les effets de corrélations entre les électrons. La différence entre l'énergie  $E_0^{\text{HF}}$  à la limite de Hartree-Fock et  $E_0^{\text{BO}}$  est, justement, l'énergie de corrélation. Des méthodes plus raffinées comme l'interaction de configurations ou la théorie des perturbations à plusieurs corps peuvent être utilisées pour solutionner des problèmes pour lesquels la corrélation est déterminante comme par exemple la dissociation d'atomes liés. Notons brièvement que la méthode dite *Möller-Plesset perturbation theory* du second ordre (MP2) apporte une correction peu coûteuse à l'énergie calculée par la méthode Hartree-Fock. La perturbation à laquelle fait référence cette méthode est  $H^{\text{BO}} - \mathcal{H}^{\text{HF}}$ .

Etant donné que les équations de Hartree-Fock sont résolues dans le cadre de l'approximation de Born-Oppenheimer, on obtient comme résultat des surfaces d'énergies potentielles ( $E_0^{\text{HF}}$ ) en fonction des coordonnées des noyaux. Ce sont ces surfaces d'énergies potentielles, calculées par le programme *Gaussian 90*,<sup>50</sup> qui sont utilisées comme courbes de références afin d'ajuster certains paramètres

faisant partie du champ de force *param22*<sup>51</sup> utilisé pour nos simulations dynamique moléculaire.

## 2.2 La dynamique moléculaire (DM) <sup>52, 53, 54, 55</sup>

La méthode principale utilisée pour cette thèse consiste à simuler par ordinateur la dynamique de systèmes moléculaires à l'aide de méthodes classiques de modélisation moléculaire et d'un modèle atomique détaillé de la molécule à simuler. Plus précisément, on simule les mouvements des atomes d'un système moléculaire en utilisant les équations du mouvement de Newton, c'est-à-dire:

$$m_i \frac{d^2 \mathbf{R}_i(t)}{dt^2} = -\nabla_i U(\mathbf{R}_1, \dots, \mathbf{R}_N). \quad (2.15)$$

où  $m_i$  est la masse de chacun des  $N$  atomes dont les coordonnées sont les  $\mathbf{R}_i$ .  $U(\mathbf{R}_1, \dots, \mathbf{R}_N)$  est la fonction d'énergie potentielle du système. Idéalement, les forces seraient calculées à partir d'un potentiel *ab initio*, mais la calbindine D<sub>9k</sub> comprend à elle seule 1195 atomes, et demeure donc pour l'instant impossible à simuler en utilisant les principes premiers de la mécanique quantique. La fonction d'énergie potentielle classique  $U(\mathbf{R}_1, \dots, \mathbf{R}_N)$  est donc utilisée en remplacement de  $E_0^{\text{HF}}(\mathbf{R}_1, \dots, \mathbf{R}_N)$ . En général, pour des raisons de temps, de mémoire et de précision, les calculs *ab initio* demeurent à exclure pour l'étude de protéines complètes et se limitent donc à des fragments moléculaires suffisamment petits. Cependant, il est possible d'élaborer un modèle classique consistant en une fonction de potentiel qui capture l'essentiel des interactions physiques les plus déterminantes du système. C'est à partir de cette fonction de potentiel qu'il nous est possible de générer une suite discrète de configurations du système, un peu comme un film où chaque image sur la pellicule est une configuration générée par la simulation et dont l'animation se fait par la lecture successive de toutes les configurations échantillonnées. Cette suite de configurations est communément appelée la trajectoire, en référence au trajet suivi par les atomes du système. Le calcul

effectué par la simulation fournit une configuration à chaque pas successif dans le temps et la trajectoire totale consiste aussi en un ensemble statistique de configurations représentatives des conditions désirées. C'est à partir de cet ensemble des configurations que les valeurs moyennes des observables physiques seront calculées. Tous les calculs présentés ici ont été effectués avec le logiciel de simulations de dynamique moléculaire CHARMM.<sup>56</sup> Toutes les descriptions des différentes étapes des calculs, quoique assez générales, ne font référence qu'à ce logiciel spécifique.

### 2.2.1 Le champ de forces et les contraintes <sup>56, 57, 58, 59</sup>

La fonction d'énergie potentielle totale  $U_{\text{tot.}}(\mathbf{R}_1, \dots, \mathbf{R}_N)$  telle qu'incluse dans le logiciel CHARMM<sup>56</sup> comporte les termes suivants:

$$U_{\text{tot.}} = U_{\text{liens}} + U_{\text{angles}} + U_{\text{dièdres}} + U_{\text{impropres}} + U_{\text{charges}} + U_{\text{Lennard-Jones}}. \quad (2.16)$$

Chacun de ces termes est défini comme suit:

1.

$$U_{\text{liens}} = \sum_{\text{lien } i} \frac{1}{2} K_i^{[r]} (r_i - r_i^e)^2 \quad (2.17)$$

Les atomes d'une liaison covalente sont modélisés par des sphères reliées par des ressorts. L'énergie associée à la longueur  $r_i$  d'un lien chimique entre deux atomes est représentée par un potentiel harmonique où  $K_i^{[r]}$  est la constante de force de rappel. Cette approximation harmonique est valable dans les cas où les liens entre deux atomes oscillent autour d'une valeur d'équilibre (ici  $r_i^e$ ) à des températures acceptables pour des systèmes biologiques.

2.

$$U_{\text{angles}} = \sum_{\text{angle } i} \frac{1}{2} K_i^{[\theta]} (\theta_i - \theta_i^e)^2 \quad (2.18)$$

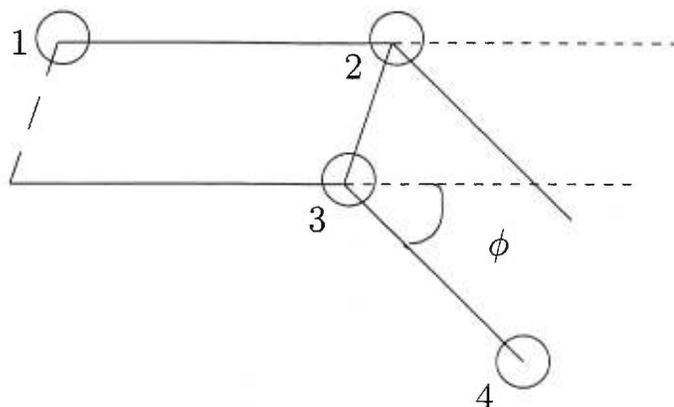


FIGURE 9. Définition d'un angle dièdre. L'angle inclut quatre atomes et est déterminé par la rencontre des deux plans incluant les trois premiers et les trois derniers des quatre atomes. Par exemple, les atomes 1, 2, 3 et 4 pourraient être les atomes de carbone d'un groupe  $\text{CH}_3\text{-CH}_2\text{-CH}_2\text{-CH}_3$ .

Considérant toujours une géométrie à l'équilibre, l'énergie associée à la déformation de l'angle polaire  $\theta_i$  formé par trois atomes liés chimiquement est aussi représentée par un potentiel harmonique où  $K_i^{[\theta]}$  et  $\theta_i^e$  sont respectivement la constante de force de rappel et l'angle à l'équilibre.

3.

$$U_{\text{dièdres}} = \sum_{\text{dièdre } i} \sum_n K_{in}^{[\phi]} (\cos(n\phi_i) + 1) \quad (2.19)$$

Ce potentiel maintient les angles dièdres formés par quatre atomes liés chimiquement (voir figure 9). Puisque les angles dièdres admettent des valeurs de  $0^\circ$  et  $360^\circ$ , un potentiel périodique est utilisé où  $K_{in}^{[\phi]}$  est la constante d'énergie imposée au dièdre (amplitude du mouvement),  $\phi_i$  est l'angle du dièdre et  $n$  est la périodicité de l'angle ( $n = 1, 2, 3, 4$  et  $6$ ). Il peut y avoir plusieurs contributions provenant de différentes valeurs de  $n$  et donc de  $K_{in}^{[\phi]}$ .

4.

$$U_{\text{impropres}} = \sum_{\text{impropre } i} \frac{1}{2} K_i^{[\omega]} (\omega_i - \omega_i^e)^2 \quad (2.20)$$

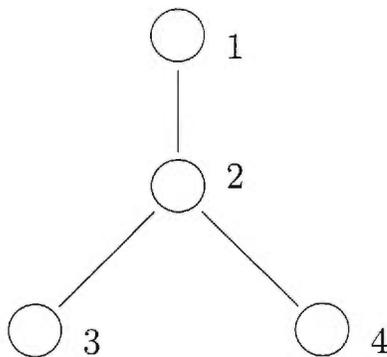


FIGURE 10. Définition d'un angle impropre. L'angle est formé par la rencontre des plans incluant les atomes 2, 3, 4 et 1, 3, 4.

Les angles impropres sont définis par la rencontre de deux plans (voir figure 10) et le potentiel harmonique qui leur est imposé sert à maintenir la chiralité ou la planéité respectivement d'éléments structuraux (quatre atomes) tétraédriques ou planaires.

5.

$$U_{\text{charges}} = \sum_{\text{paire } ij} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad (2.21)$$

et

$$U_{\text{Lennard-Jones}} = \sum_{\text{paire } ij} 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \quad (2.22)$$

Ces équations rendent compte des interactions entre atomes non-liés chimiquement. D'une part, l'équation 2.21 rend compte des interactions électrostatiques entre charges et/ou charges partielles;  $q_i$  et  $q_j$  sont les charges d'une paire d'atomes,  $\epsilon_0$  est la permittivité du vide et  $r_{ij}$  est la distance entre la paire d'atomes.

D'autre part, l'équation 2.22 émerge du principe d'exclusion de Pauli qui donne lieu à une répulsion à courte portée lors du chevauchement des nuages électroniques de deux atomes non-liés chimiquement. Ce potentiel est modélisé par un terme en  $1/r^{12}$ . L'énergie d'interaction (attraction) entre les dipôles électriques momentanés de deux atomes neutres est modélisée par un terme en  $-1/r^6$  qui, joint au terme de répulsion, forme le potentiel de

Lennard-Jones. A l'équation 2.22,  $r_{ij}$  est la distance entre les atomes,  $\epsilon_{ij}$  est la profondeur du puits de potentiel et  $\sigma_{ij}$  est le diamètre de Lennard-Jones, tous deux pour les deux espèces atomiques  $i$  et  $j$  combinées, c'est-à-dire

$$\sigma_{ij} = \frac{(\sigma_i + \sigma_j)}{2} \quad (2.23)$$

et

$$\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}. \quad (2.24)$$

Le calcul des interactions électrostatiques à longue portée requiert une partie considérable du temps de calcul. Aussi, plutôt que de les couper à une certaine distance souvent trop courte, il est utile de déterminer une distance au-delà de laquelle ces interactions ne sont plus calculées par paires d'atomes. L'équation 2.25 représente la façon dont les interactions électrostatiques sont traitées lorsqu'on impose une telle distance de coupure:

$$U_{\text{charges}} = \sum_{r_{ij} < r_{\text{coup.}}} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} + \sum_{r_{ij} > r_{\text{coup.}}} q_i \left[ \frac{1}{2} V_i + (\mathbf{r}_i - \mathbf{r}_{i0}) \cdot \mathbf{E}_i \right] \quad (2.25)$$

Le  $r_{\text{coup.}}$  est la distance au-delà de laquelle les interactions ne sont plus calculées par paires d'atomes et le deuxième terme traite les interactions au-delà de cette distance de coupure.  $V_i$  et  $\mathbf{E}_i$  sont respectivement le potentiel et champ électriques approximatifs calculés à la position de la charge atomique  $q_i$  et auxquels ne contribuent que les atomes  $j$  présents au-delà de la distance de coupure;  $\mathbf{r}_{i0}$  sont les coordonnées de l'atome  $i$  au pas d'intégration précédent. Ces interactions à longue portée sont calculées en effectuant un développement de multipôles (coupée aux quadrupôles) pour des groupes de moins de six atomes dont la somme des charges est habituellement zéro, à moins que l'acide aminé dont ils font partie soit chargé.

Parmi les approximations incluses dans ce modèle, l'approximation harmonique qui suppose que tous les liens et angles oscillent autour et près d'un minimum a déjà été mentionnée. Notons aussi, pour les atomes non-liés chimiquement, l'absence d'un terme explicite représentant la polarisation induite sur les atomes.

Malgré l'absence de ce terme explicite d'interactions, certains de leurs effets physiques ont été incorporés dans les différents paramètres du champ de force de façon à reproduire le mieux possible les données expérimentales pour les acides aminés. Tous les paramètres, sauf le  $\sigma$  du calcium, proviennent du champ de force *param22*<sup>51</sup> développé pour CHARMM. Notons que le modèle de molécule d'eau TIP3P (*Transferable Interaction Potential for 3 Particles*)<sup>60</sup> a été utilisé et le développement du paramètre  $\sigma$  de Lennard-Jones pour le  $\text{Ca}^{2+}$  est présenté au chapitre suivant. En ce qui concerne les autres paramètres du champ de force *param22*, les longueurs de liens ou angles à l'équilibre ont été évaluées par des méthodes *ab initio* sur des molécules modèles simples en phase gazeuse et aussi à partir de données de structures expérimentales (rayons X, RMN, etc. . .). Les constantes de forces pour les termes harmoniques peuvent être déterminées à partir de spectres vibrationnels expérimentaux, toujours dans le vide. Une autre méthode consiste à minimiser l'énergie de la structure pour ensuite trouver les modes normaux et les comparer avec les spectres connus. Les constantes de force et les autres paramètres sont ensuite lissés par une méthode des moindres carrés jusqu'à ce que l'on reproduise les données expérimentales. Les paramètres de Lennard-Jones ont été ajustés à des données expérimentales de viscosité, de diffusion, de structures cristallographiques ou de calculs *ab initio* de haut niveau. Le développement du paramètre  $\sigma$  de Lennard-Jones du  $\text{Ca}^{2+}$ , présenté au chapitre suivant, donnera une idée de ces méthodes.

### 2.2.2 Les conditions aux frontières <sup>61, 62, 63</sup>

La méthode utilisée dans cette thèse pour représenter adéquatement l'état hydraté de la calbindine  $\text{D}_{9k}$  permet de n'inclure qu'un nombre minimum de molécules d'eau explicites dans les simulations et s'avère donc avantageuse comparativement aux autres méthodes disponibles. Dans toutes les simulations la calbindine  $\text{D}_{9k}$  est entourée par 2285 molécules d'eau, c'est-à-dire le nombre minimum

requis pour que tous les atomes de la protéine demeurent entourés de molécules d'eau pendant toutes les simulations et ce à pression constante. Les conditions appliquées aux frontières du système permettent de reproduire les propriétés statistiques correspondant au cas d'une protéine immergée dans un bain infini de molécules d'eau. Un premier fondement théorique de cette méthode consiste à séparer la fonction de partition du système en deux contributions:

$$Z = \int d(\mathbf{X}) \frac{1}{n!} \int d(\mathbf{1}) \dots d(\mathbf{n}) \frac{1}{(N-n)!} \int' d(\mathbf{n}+1) \dots d(\mathbf{N}) \exp[-\beta U] \quad (2.26)$$

où  $\mathbf{X}$  et  $(\mathbf{1} \dots \mathbf{N})$  sont respectivement les degrés de liberté de la protéine et des  $N$  molécules de solvant,  $U$  est l'énergie potentielle totale du système,  $\beta = 1/(k_B T)$  où  $k_B$  est la constante de Boltzmann et  $T$  la température. Une première contribution provient des  $n$  molécules d'eau dans le voisinage de la protéine, et une deuxième attribuable aux  $(N-n)$  autres molécules se trouvant au-delà des  $n$  plus proches. Le prime sur l'intégrale indique que ces  $\{n+1 \dots N\}$  molécules d'eau de la couche externe sont restreintes à une région délimitée de l'espace afin d'éviter un double comptage de configurations. En pratique, toutes les configurations où une des  $\{n+1 \dots N\}$  molécules d'eau externes se retrouve plus près du soluté qu'une des  $n$  molécules d'eau internes sont rejetées.

Une autre étape clef consiste à exprimer de façon rigoureuse les propriétés statistiques à l'équilibre d'un système infini en termes des  $n$  molécules explicites. La définition suivante est posée en condition:

$$\exp[-\beta W(\mathbf{X}; \mathbf{1} \dots \mathbf{n})] \equiv \frac{1}{C} \int' d(\mathbf{n}+1) \dots d(\mathbf{N}) \exp[-\beta U(\mathbf{X}; \mathbf{1} \dots \mathbf{N})] \quad (2.27)$$

où  $W(\mathbf{X}; \mathbf{1} \dots \mathbf{n})$  est le potentiel effectif (potentiel de force moyenne<sup>64</sup>) exercé par les  $(N-n)$  molécules en dehors de la frontière sur les  $n$  molécules explicites internes et  $C$  une constante arbitraire. Les propriétés statistiques à l'équilibre peuvent être réévaluées par

$$\langle A \rangle = \frac{\int d(\mathbf{X}) \frac{1}{n!} \int d(\mathbf{1}) \dots d(\mathbf{n}) A(\mathbf{X}) \exp[-\beta W]}{\int d(\mathbf{X}) \frac{1}{n!} \int d(\mathbf{1}) \dots d(\mathbf{n}) \exp[-\beta W]} \quad (2.28)$$

sans que l'influence des  $(N-n)$  molécules n'apparaisse explicitement. Ce potentiel de force moyenne exercé par les  $(N-n)$  molécules externes sur les  $n$  molécules explicites internes est la base du potentiel imposé à la frontière sphérique du système. De plus, les  $n$  molécules explicites internes sont contenues dans une sphère dont le rayon varie selon la configuration instantanée de ces  $n$  molécules internes.

Pour notre cas particulier où une protéine est immergée dans 2285 molécules d'eau explicites, le potentiel de force moyenne imposé à la frontière sphérique du système correspond en fait à l'énergie libre d'hydratation  $\Delta W$  du système (protéine + 2285 molécules d'eau). On a approximé que cette énergie libre d'hydratation pouvait être décomposée en trois contributions:

$$\Delta W = \Delta W_{\text{cavité}} + \Delta W_{\text{vdW}} + \Delta W_{\text{élec.}} \quad (2.29)$$

où  $\Delta W_{\text{cavité}}$  est l'énergie libre requise pour former la cavité,  $\Delta W_{\text{vdW}}$  est l'énergie libre requise pour assembler le système moléculaire neutre à l'intérieur de la cavité et  $\Delta W_{\text{élec.}}$  qui est l'énergie libre requise pour charger les atomes du système contenus dans la cavité. Ces contributions s'ajoutent aux autres termes du potentiel énoncés à la section précédente.

### 2.2.3 Étapes des simulations

- Coordonnées initiales

Les coordonnées initiales des atomes des trois états de la calbindine  $D_{9k}$  proviennent de diverses sources. La structure A du mutant P43G a été déterminé par des expériences de RMN ( $^1\text{H}$ ) en solution. Les coordonnées ont ensuite été raffinées à l'aide de la DM dans le vide (Skelton, 1995;<sup>32</sup> PDB1clb dans la banque Brookhaven de données des protéines). La structure  $S_{II}$  du mutant P43G a été déterminée par les mêmes techniques (Akke, 1995<sup>33</sup>), sauf que lors des expériences c'est un ion  $\text{Cd}^{2+}$  qui était présent

dans le site II, parce que l'état  $S_{II}$  lié au  $Ca^{2+}$  ne peut être isolé expérimentalement. Cependant, cet ion  $Cd^{2+}$  a été remplacé par un ion  $Ca^{2+}$  dans les simulations afin de rendre cet état comparable aux autres. La structure de la protéine sauvage dans l'état D a été déterminée par diffraction des rayons X à une résolution de 1.6 Å (Svensson, 1992;<sup>15</sup> PDB4icb dans la banque de données Brookhaven). En ce qui concerne l'état  $S_I$ , comme aucune structure expérimentale n'était disponible, il a été simulé à partir de la structure expérimentale de l'état D duquel le  $Ca^{2+}$  dans le site II a été effacé ( $S_I(D)$  dans le reste du texte) pour ensuite être soumis à un protocole d'équilibration qui sera détaillé au chapitre 4. Chacune des structures expérimentales est ensuite placée dans une sphère de 2285 molécules d'eau générée par CHARMM.<sup>56</sup>

- Minimisation de l'énergie potentielle

Le système total comprenant les coordonnées de la calbindine  $D_{9k}$ , les ions et les molécules d'eau est amélioré du point de vue énergétique en utilisant un algorithme de minimisation. Dans notre cas, de tels algorithmes ont été utilisés afin d'éliminer le plus possible les tensions causées par des contacts défavorables entre atomes rapprochés dans l'espace. Les deux algorithmes de minimisation de l'énergie *steepest-descent* et *Adapted Basis Set Newton-Raphson* (ABNR) ont été utilisés à cette fin.

- Équilibration du système

Cette étape vise à amener le système à un état stable où il reproduit les caractéristiques de l'ensemble statistique désiré, soit dans notre cas l'ensemble microcanonique, c'est-à-dire un nombre  $N$  constant de particules contenues dans un volume  $V$  déterminé, le tout correspondant à une énergie qui varie entre  $E$  et  $E \pm \delta E$  à une température  $T$  donnée. Notons que le système a une probabilité constante de visiter tous les états situés dans cet intervalle d'énergie. Il existe plusieurs protocoles afin d'en arriver à ce point, et celui utilisé ici consiste à assigner d'emblée à chacun des atomes une

vitesse initiale de sorte que l'ensemble des vitesses constitue une distribution gaussienne dont le centre correspond à la température voulue. Une conséquence du théorème d'équipartition est que la température en tout temps est reliée aux vitesses par

$$TN_{\text{libres}}k_B = \sum_{i=1}^{N_{\text{atomes}}} m_i v_i^2 \quad (2.30)$$

où  $T$  est la température,  $N_{\text{libres}}$  est le nombre de degrés de liberté,  $k_B$  est la constante de Boltzmann,  $N_{\text{atomes}}$  est le nombre d'atomes et  $m_i$  et  $v_i$  sont respectivement la masse et la vitesse de l'atome  $i$ . La distribution des vitesses est réajustée jusqu'à ce que la température du système soit à l'équilibre, moment où commence l'enregistrement des configurations qui feront partie de l'ensemble statistique.

- Production de la trajectoire

C'est lors de la période de production que l'on enregistre la trajectoire. L'hypothèse d'ergodicité est à la base de la dynamique moléculaire: si on considère un système à l'équilibre, tous les états ont une probabilité égale d'être occupés, mais comme il n'est possible de générer qu'une seule configuration à un instant donné, on suppose qu'en un temps raisonnable, on passera par un nombre suffisant de configurations pour former un ensemble statistique valable. Alors, les moyennes d'ensembles peuvent être remplacées par des moyennes temporelles.

## 2.2.4 Le processus d'évolution dans le temps

L'étape de production consiste à calculer la trajectoire proprement dite des atomes du système, c'est-à-dire à résoudre les équations du mouvement pour des temps successifs et ainsi obtenir les coordonnées et les vitesses en fonction du temps. Les atomes de la protéine se déplacent selon les équations de Newton, tandis que les molécules d'eau (atome d'oxygène) sont traitées comme des particules

browniennes et leurs équations du mouvement se résument à l'équation de Langevin:

$$m_i \ddot{\mathbf{R}}_i(t) = -\nabla_i U(\mathbf{R}_i) + \mathbf{f}_i(t) - \gamma_i \dot{\mathbf{R}}_i(t), \quad i = 1, 2, \dots, N \quad (2.31)$$

où  $-\nabla_i U(\mathbf{R}_i)$  est la force appliquée à un atome de masse  $m_i$  due au potentiel  $U(\mathbf{R}_i)$ . La force aléatoire  $\mathbf{f}_i(t)$  est attribuée selon une distribution gaussienne dont l'amplitude des fluctuations est proportionnelle au coefficient de friction  $\gamma_i$  et à la température  $T$ , c'est-à-dire

$$\langle \mathbf{f}_i(t) \cdot \mathbf{f}_i(0) \rangle = 2k_B T \gamma_i \delta(t) \quad (2.32)$$

où le delta de Dirac indique que la corrélation est de 0 entre deux pas successifs dans le temps d'où le caractère aléatoire. La force de friction est donnée par  $-\gamma_i \dot{\mathbf{R}}_i(t)$ . Si on pose  $\gamma_i = 0$  comme tel est le cas pour les atomes de la protéine, on retrouve simplement les équations de Newton. La démarche afin de produire une trajectoire consiste à calculer, dans un premier temps, le potentiel  $U(\mathbf{R}_i(0))$  d'une configuration définie à partir des coordonnées initiales  $\{\mathbf{R}_i(0)\}$  de la structure moléculaire. Puis, dans un deuxième temps, la force  $\mathbf{F}_i$  exercée sur chacun des atomes est calculée en dérivant numériquement le potentiel par rapport aux coordonnées de chaque atome. Des forces atomiques on obtient directement les accélérations atomiques qui sont intégrées numériquement pour finalement recueillir leurs vitesses  $\{\mathbf{v}_i(t)\}$  et positions  $\{\mathbf{R}_i(t)\}$  au temps  $t$ . L'algorithme suivant décrit les étapes principales d'une simulation de DM:

- Configuration initiale du système de  $N$  atomes à  $t = 0$ ,  $\{\mathbf{R}_i(0), i = 1, \dots, N\}$ .
- Calcul des  $\{U(\mathbf{R}_i(t))\}$ .
- Calcul des  $\{\mathbf{F}_i = -\nabla_i U(\mathbf{R}_i) + \mathbf{f}_i(t) - \gamma_i \dot{\mathbf{R}}_i(t)\}$ .
- Calcul des  $\{\ddot{\mathbf{R}}_i(t) = \mathbf{F}_i(t)/m_i\}$ .
- Calcul des  $\{\mathbf{R}_i(t + \Delta t), \dot{\mathbf{R}}_i(t + \Delta t)\}$ .

Parmi les divers algorithmes permettant ce dernier calcul, les algorithmes du second ordre de Verlet et *leap frog* ont été utilisés. Il sont définis comme suit:

$$\begin{aligned}\mathbf{R}(t + \Delta t) &= 2\mathbf{R}(t) - \mathbf{R}(t - \Delta t) + \Delta t^2 \frac{F}{m} \\ \dot{\mathbf{R}}(t) &= \frac{[\mathbf{R}(t + \Delta t) - \mathbf{R}(t - \Delta t)]}{2\Delta t}\end{aligned}\quad (2.33)$$

(Verlet) et

$$\begin{aligned}\dot{\mathbf{R}}\left(t + \frac{\Delta t}{2}\right) &= \dot{\mathbf{R}}\left(t - \frac{\Delta t}{2}\right) + \Delta t \frac{F}{m} \\ \mathbf{R}(t + \Delta t) &= \mathbf{R}(t) + \dot{\mathbf{R}}\left(t + \frac{\Delta t}{2}\right) \Delta t\end{aligned}\quad (2.34)$$

(*leap frog*). Il est important que le pas d'intégration demeure inférieur au temps caractéristique du mouvement le plus rapide du système. Habituellement ces mouvements rapides sont dus aux liens chimiques impliquant un atome hydrogène. Le temps total de la simulation dépend donc de la longueur du pas d'intégration puisque nous sommes limités en partie par l'utilisation du temps de calcul sur ordinateur. Une solution est de contraindre ces liens à une longueur fixe, et c'est ce que fait l'algorithme SHAKE<sup>65, 66</sup> utilisé dans tous nos calculs.

### 2.3 Les calculs d'énergies libres <sup>56, 62, 63, 67</sup>

La notion d'énergie libre revient plus d'une fois dans cette thèse. Il en est question dès le chapitre 3 et l'énergie libre est le sujet principal des calculs effectués au chapitre 6. Plus spécifiquement, ce sont des différences d'énergies libres qui y sont calculées. Les prochains paragraphes exposent brièvement les notions de base ainsi que la façon dont ces calculs sont effectués quand on dispose d'une méthode de production de trajectoire de DM.

La mécanique statistique permet le passage d'une distribution microscopique d'un système à des observables macroscopiques décrivant des propriétés moyennes. La notion de fonction de partition est donc centrale à la mécanique statistique,

puisque'elle représente la somme de tous les états accessibles au système. Sa définition est

$$Z = \int \dots \int \exp[-\beta U(\mathbf{X})] d\mathbf{X} \quad (2.35)$$

où  $U(\mathbf{X})$  est l'énergie d'une configuration dont les  $\mathbf{X}$  sont les coordonnées qui la définissent (ici le volume  $V$  est considéré comme une variable dynamique et la pression est constante) et  $\beta = 1/(k_B T)$  où  $k_B$  est la constante de Boltzmann et  $T$  la température. La fonction de probabilité de Boltzmann étant donnée par

$$P(\mathbf{X}) = \frac{\exp[-\beta U(\mathbf{X})]}{Z} = \frac{\exp[-\beta U(\mathbf{X})]}{\int \dots \int \exp[-\beta U(\mathbf{X})] d\mathbf{X}}, \quad (2.36)$$

l'énergie moyenne du système

$$\langle U(\mathbf{X}) \rangle = \int U(\mathbf{X}) P(\mathbf{X}) d\mathbf{X} \quad (2.37)$$

correspond à une moyenne sur l'ensemble des configurations. L'énergie libre de Gibbs  $G$  du système est reliée à la fonction de partition par

$$G = -k_B T \ln Z \quad (2.38)$$

de sorte qu'une différence d'énergies libres  $\Delta G$  entre deux états bien définis d'un système, soit les états 0 et 1, est donnée par

$$\Delta G = G_1 - G_0 = -k_B T \ln \frac{Z_1}{Z_0}. \quad (2.39)$$

Il est maintenant possible d'établir le lien qui nous permet en pratique d'effectuer ce calcul, c'est-à-dire de transformer l'équation précédente et y retrouver une moyenne d'ensemble. Si, dans le rapport  $Z_1/Z_0$

$$\frac{Z_1}{Z_0} = \frac{\int \dots \int \exp[-\beta U_1(\mathbf{X})] d\mathbf{X}}{\int \dots \int \exp[-\beta U_0(\mathbf{X})] d\mathbf{X}} \quad (2.40)$$

on insère l'identité suivante

$$1 \equiv \exp[-\beta U_0(\mathbf{X})] \exp[\beta U_0(\mathbf{X})], \quad (2.41)$$

on obtient

$$\begin{aligned}
\frac{Z_1}{Z_0} &= \frac{\int \dots \int \exp[-\beta U_1(\mathbf{X})] \exp[\beta U_0(\mathbf{X})] \exp[-\beta U_0(\mathbf{X})] d\mathbf{X}}{\int \dots \int \exp[-\beta U_0(\mathbf{X})] d\mathbf{X}} \quad (2.42) \\
&= \int \dots \int \exp[-\beta(\Delta U(\mathbf{X}))] P_0(\mathbf{X}) d\mathbf{X} \\
&= \langle \exp[-\beta \Delta U(\mathbf{X})] \rangle_{(0)}
\end{aligned}$$

où  $\Delta U(\mathbf{X}) = U_1(\mathbf{X}) - U_0(\mathbf{X})$  est le changement d'énergie configurationnelle associé au passage de l'état 0 à l'état 1 et où l'indice 0 indique que la moyenne d'ensemble est effectuée à partir des configurations de l'état initial du système.

Dans ce cas,

$$\begin{aligned}
\Delta G(0 \rightarrow 1) &= -k_B T \ln \langle \exp[-\beta \Delta U(\mathbf{X})] \rangle_{(0)} \quad (2.43) \\
\Delta G(1 \rightarrow 0) &= -k_B T \ln \langle \exp[+\beta \Delta U(\mathbf{X})] \rangle_{(1)}
\end{aligned}$$

puisque le processus est réversible et que la définition de ce qui est initial ou final est arbitraire. Cette méthode n'est cependant efficace que pour les cas où les états initial et final sont très proches en énergie, c'est-à-dire les situations pour lesquelles un état peut être considéré comme une perturbation par rapport à l'autre. Tout ceci parce que les changements doivent être adiabatiques afin que le système demeure autant que possible à l'équilibre.

Dans notre cas, où l'on varie la charge ou le paramètre de Lennard-Jones  $\sigma_i$  d'un atome, la différence d'énergie libre entre ces deux états est trop grande pour n'être considérée que comme une perturbation. La méthode des perturbations thermodynamiques (P.T.)<sup>68</sup> permet de calculer la différence d'énergie libre entre ces deux états que l'on relie par un paramètre de couplage  $\lambda$  et dont dépend l'énergie configurationnelle  $U(\mathbf{X}; \lambda)$  entre une valeur initiale et une valeur finale de sorte que l'on puisse passer de  $U(\mathbf{X}; \lambda_{\text{initial}})$  à  $U(\mathbf{X}; \lambda_{\text{final}})$ . L'énergie libre devient donc une fonction continue entre ces deux états tel que

$$G(\lambda_{\text{final}}) = G(\lambda_{\text{initial}}) - k_B T \ln \left[ \frac{Z(\lambda_{\text{final}})}{Z(\lambda_{\text{initial}})} \right] \quad (2.44)$$

où

$$Z(\lambda) = \int \dots \int \exp[-\beta U(\mathbf{X}; \lambda)] d\mathbf{X}. \quad (2.45)$$

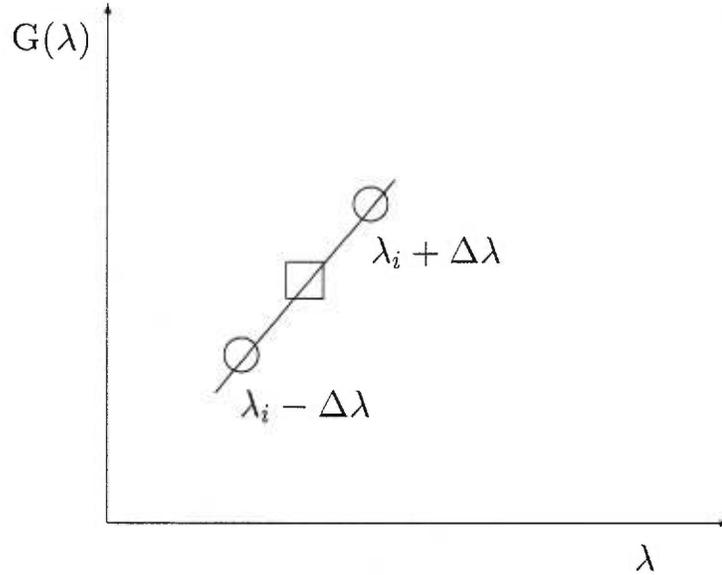


FIGURE 11. Schéma de la méthode des perturbations.

Le changement total d'énergie est divisé en un nombre fini d'intervalles  $\Delta G$  dont chacun peut être considéré comme une perturbation:

$$\Delta G(\lambda_i \rightarrow \lambda_i \pm \Delta\lambda) = -k_B T \ln \langle \exp[-\beta(U(\mathbf{X}; \lambda \pm \Delta\lambda) - U(\mathbf{X}; \lambda))] \rangle_{(\lambda_i)} \quad (2.46)$$

où  $(U(\mathbf{X}; \lambda \pm \Delta\lambda) - U(\mathbf{X}; \lambda))$  est cette fois le changement d'énergie configurationnelle associé à la perturbation de  $\lambda_i$ . La figure 11 illustre le processus global et la méthode utilisée. Le carré représente une valeur arbitraire de  $\lambda$ , soit par exemple  $\lambda_i$ . Le système est tout d'abord équilibré à cette valeur, puis un ensemble de configurations est produit. Les énergies configurationnelles sont ensuite recalculées à partir du même ensemble de configurations mais à  $\lambda_i \pm \Delta\lambda$  de sorte à obtenir la perturbation de l'énergie configurationnelle qui permet le calcul de la perturbation sur l'énergie libre *via* l'équation (2.46). La différence totale d'énergie libre est obtenue en additionnant tous les  $\Delta G$ :

$$G(\lambda_{\text{final}}) = G(\lambda_{\text{initial}}) + \sum_{i=1}^n -\Delta G(\lambda_i \rightarrow \lambda_i - \Delta\lambda) + \Delta G(\lambda_i \rightarrow \lambda_i + \Delta\lambda) \quad (2.47)$$

De plus, ces calculs doivent être effectués dans les deux directions, comme montré à la figure 11, et la similitude des résultats de l'aller et du retour est une indication de la qualité du calcul. Entre autres, cela est un indicateur à savoir si l'ensemble

des configurations produites est statistiquement suffisant. Ces critères de convergence et la qualité des calculs sont discutés dans les chapitres concernés.

La méthode d'intégration thermodynamique (I.T.)<sup>64</sup> adopte une approche légèrement différente pour le calcul du  $\Delta G(\lambda)$ . Reformulons la différence d'énergie libre  $\Delta G$  comme suit:

$$\Delta G = G(1) - G(0) = \int_0^1 \frac{\partial G(\lambda)}{\partial \lambda} d\lambda \quad (2.48)$$

où

$$\frac{\partial G(\lambda)}{\partial \lambda} = -k_B T \frac{\partial \ln Z(\lambda)}{\partial \lambda} = -\frac{k_B T}{Z(\lambda)} \frac{\partial Z(\lambda)}{\partial \lambda}. \quad (2.49)$$

Considérant que

$$Z(\lambda) = \int \dots \int \exp[-\beta U(\mathbf{X}; \lambda)] d\mathbf{X} \quad (2.50)$$

on calcule que

$$\frac{\partial Z(\lambda)}{\partial \lambda} = -\beta \int \dots \int \frac{\partial U(\mathbf{X}; \lambda)}{\partial \lambda} \exp[-\beta U(\mathbf{X}; \lambda)] d\mathbf{X}. \quad (2.51)$$

En substituant ce résultat dans l'équation 2.49 on obtient

$$\frac{\partial G(\lambda)}{\partial \lambda} = \frac{1}{Z(\lambda)} \int \dots \int \frac{\partial U(\mathbf{X}; \lambda)}{\partial \lambda} \exp[-\beta U(\mathbf{X}; \lambda)] d\mathbf{X} \quad (2.52)$$

que l'on simplifie en utilisant les équations 2.36 et 2.37, ce qui donne:

$$\int \dots \int \frac{\partial U(\mathbf{X}; \lambda)}{\partial \lambda} P(\mathbf{X}) d\mathbf{X} = \left\langle \frac{\partial U(\mathbf{X}; \lambda)}{\partial \lambda} \right\rangle \quad (2.53)$$

où  $\langle \dots \rangle$  indique une moyenne d'ensemble. On obtient alors l'expression suivante en remplacement de l'équation 2.49:

$$\frac{\partial G(\lambda)}{\partial \lambda} = \left\langle \frac{\partial U(\mathbf{X})}{\partial \lambda} \right\rangle_{(\lambda)}. \quad (2.54)$$

Finalement, le résultat équivalent à l'équation 2.47 pour la méthode d'intégration thermodynamique est le suivant:

$$G(\lambda_{\text{final}}) = G(\lambda_{\text{initial}}) + \int_{\lambda_{\text{initial}}}^{\lambda_{\text{final}}} \left\langle \frac{\partial U(\mathbf{X}; \lambda)}{\partial \lambda} \right\rangle_{(\lambda)} d\lambda \quad (2.55)$$

Pratiquement, il suffit de faire autant de simulations que de sauts discrets de  $\lambda_i$  et d'intégrer numériquement l'équation (2.55) à partir des valeurs de  $\langle U(\mathbf{X}) \rangle_{\lambda_i}$  calculées pour des  $\lambda$  successifs allant de l'état initial à l'état final, de façon similaire à la méthode des perturbations thermodynamiques.

## CHAPITRE 3

# Étude des interactions entre le $\text{Ca}^{2+}$ et ses ligands dans la calbindine $\text{D}_{9k}$

### 3.1 Introduction

La première étape de ce travail consiste à paramétrer un ion calcium dans un contexte où l'énergie totale des interactions non-liées est la somme des interactions entre paires d'atomes, ce qui implique que les effets reliés aux interactions à plusieurs corps sont négligés. De plus, cette détermination de paramètres se situe dans un contexte où les charges sur les atomes sont fixes. Dans ce chapitre, seuls les paramètres pour le  $\text{Ca}^{2+}$  sont déterminés et cela en fonction des paramètres déjà présents dans le champ de force *param22*.<sup>51</sup> La charge du  $\text{Ca}^{2+}$  est fixée à  $+2e$ , où  $e$  est l'unité de charge élémentaire, puisque les paramètres des acides aminés dans le champ de force *param22*<sup>51</sup> ont été déterminés dans un contexte où ils interagissent avec des molécules neutres. Mais en fait, le  $\text{Ca}^{2+}$  en présence de ligands chargés négativement subit un transfert de charge assez important, pouvant amener sa charge jusqu'aux environs de  $+1.6e$  lorsque sa coordinence est de sept ou huit (voir section 3.2). En premier lieu, les paramètres  $\sigma$  et  $\epsilon$  du potentiel de Lennard-Jones sont ajustés afin d'être combinés à d'autres espèces atomiques et représenter le mieux possible les réalités physiques des interactions  $\text{Ca}^{2+}$ -ligands. La principale méthode d'ajustement utilisée à cette fin est de tenter de reproduire le mieux possible l'énergie libre d'hydratation du  $\text{Ca}^{2+}$  mesurée expérimentalement par des simulations de dynamique moléculaire (DM). Ce choix est basé sur l'argument<sup>69</sup> que l'ajustement de paramètres à des courbes d'interactions en phase gazeuse ne

garantit pas que ces mêmes paramètres soient aptes à reproduire des propriétés du système en phase condensée. Puis, le paramètre ajusté est testé en comparant les courbes d'énergies d'interactions  $\text{Ca}^{2+}$ -ligand calculées avec CHARMM<sup>56</sup> à celles calculées par des méthodes *ab initio* à l'aide du logiciel Gaussian90.<sup>50</sup>

La section 3.3 évalue la contribution des interactions non-additives ou effets à plusieurs corps qui sont implicites dans les calculs *ab initio* mais qui n'apparaissent pas de façon explicite dans la fonction de potentiel de CHARMM.<sup>56</sup> Cette contribution est cependant d'une certaine façon incluse dans le fait que les paramètres du potentiel à deux corps sont déterminés pour reproduire l'énergie libre d'hydratation. En prévision d'inclure dans CHARMM un terme d'énergie qui rendrait compte de façon explicite et approximative de la contribution non-additive, la section 3.4 décrit la méthode utilisée pour déterminer la polarisabilité atomique de chacun des atomes liés au  $\text{Ca}^{2+}$  dans les sites. Enfin, la section 3.5 présente une tentative d'inclusion d'un terme d'énergie de polarisation induite pour le cas test de l'interaction entre le  $\text{Ca}^{2+}$  et les molécules d'eau.

## 3.2 Paramétrisation du $\text{Ca}^{2+}$

### 3.2.1 Calcul de l'énergie libre d'hydratation du $\text{Ca}^{2+}$ et propriétés statiques du $\text{Ca}^{2+}$ dans l'eau

L'énergie libre d'hydratation expérimentale du  $\text{Ca}^{2+}$  est la principale mesure en fonction de laquelle les paramètres  $\sigma_{\text{Ca}^{2+}}$  et  $\epsilon_{\text{Ca}^{2+}}$  de Lennard-Jones sont ajustés. Les calculs de l'énergie libre d'hydratation du  $\text{Ca}^{2+}$  sont réalisés par DM de Langevin selon la méthode exposée à la section 2.3. La première étape consiste à simuler un système à pression constante où le  $\text{Ca}^{2+}$  est déjà inclus dans une sphère de molécules d'eau (modèle TIP3P<sup>60</sup>) de sorte à ne pas avoir à calculer l'énergie libre associée à la formation d'une cavité pour loger le  $\text{Ca}^{2+}$ . Ce dernier est fixé au centre de la sphère et les conditions aux frontières sont celles exposées

à la section 2.2.2. Ces calculs ont été effectués avec des sphères comptant 6, 25, 50 et 100 molécules d'eau afin d'observer la convergence de la méthode du potentiel appliqué à la frontière sphérique du solvant<sup>61</sup> en fonction du nombre de molécules d'eau incluses. Le même protocole a été utilisé pour chacun des quatre systèmes de 6, 25, 50 et 100 molécules d'eau.

Tout d'abord, l'énergie du système subit 100 pas de minimisation ABNR (voir section 2.2.3). S'ensuit une dynamique de Langevin simulée à 300 K avec un coefficient de friction  $\gamma_i/m_i = 0.2 \text{ ps}$  appliqué seulement aux oxygènes ce qui correspond à un temps caractéristique de relaxation de  $5 \text{ ps}^{-1}$  et qui est déterminé par la viscosité de l'eau liquide et du coefficient de friction d'une molécule d'eau dans l'eau.

Un pas d'intégration de 0.001 ps est utilisé dans tous les cas. Afin d'économiser du temps de calcul, seule la charge du  $\text{Ca}^{2+}$  est perturbée lors du calcul de l'énergie libre, ce qui signifie que seule la contribution électrostatique à l'énergie libre d'hydratation du  $\text{Ca}^{2+}$  est évaluée. L'énergie libre associée à la formation de la cavité où loge le  $\text{Ca}^{2+}$  est approximée à partir des données homologues<sup>61</sup> pour le  $\text{Na}^+$  ( $+2.6 \pm 0.5 \text{ kcal/mol}$ ) qui possède un rayon ionique ( $r_{\text{Na}^{2+}} = 0.95 \text{ \AA}$ )<sup>57</sup> proche de celui du  $\text{Ca}^{2+}$  ( $r_{\text{Ca}^{2+}} = 0.99 \text{ \AA}$ ).<sup>57</sup> Cette contribution pour le sodium a été calculée avec les mêmes conditions aux frontières et la même méthode de perturbations. L'énergie libre d'hydratation expérimentale du  $\text{Ca}^{2+}$  utilisée comme référence étant de  $-380.8 \text{ kcal/mol}$ ;<sup>70</sup> on vise donc à calculer une valeur autour de  $-383.4 \text{ kcal/mol}$  pour la contribution électrostatique.

La charge du  $\text{Ca}^{2+}$  est perturbée de  $+2e$  à  $0e$  en simulant le système à des valeurs successives de  $\lambda$ , partant de 1.95 jusqu'à 0.05 par sauts de 0.10, pour un total de 20 fenêtres. Le calcul est aussi effectué en sens inverse ( $0e$  à  $+2e$ ) et c'est la moyenne des deux directions (aller et retour) qui est retenue comme résultat final. A chaque fenêtre de  $\lambda$ , on effectue 10 ps d'équilibration du système, suivi par une période de 10 ps de production de trajectoire. Un total de

1000 configurations forme l'ensemble statistique servant au calcul de chacun des 20  $\Delta G_i$  dont la somme (voir équation 2.47) est la contribution électrostatique totale à l'énergie libre d'hydratation du  $\text{Ca}^{2+}$ . La méthode d'essais et erreurs a été utilisée afin de s'approcher du  $\sigma_{\text{Ca}^{2+}}$  optimal, en ayant préalablement fixé le paramètre  $\epsilon_{\text{Ca}^{2+}}$  à  $-0.1200$  kcal/mol qui est la valeur attribuée au  $\text{Ca}^{2+}$  par Prod'hom dans *param22*<sup>51</sup> lors d'une première tentative de paramétrisation du  $\text{Ca}^{2+}$ . Le fait de fixer le paramètre  $\epsilon$  à une valeur déjà attribuée et de varier  $\sigma$  vient du fait que l'énergie du terme de van der Waals a une plus grande sensibilité aux fluctuations de ce dernier (voir équation 2.22).

Les systèmes comptant 6 et 25 molécules d'eau ont d'abord été utilisés puisque ces calculs sont relativement rapides, et ensuite les systèmes comptant 50 et 100 molécules d'eau pour un raffinement final. Le paramètre  $\sigma_{\text{Ca}^{2+}}$  a été fixé au moment où on reproduisait raisonnablement la valeur expérimentale de référence. La valeur finale pour le  $\text{Ca}^{2+}$  ainsi que les autres paramètres sont donnés au tableau I.

	$\epsilon$	$\sigma$
	(kcal/mol)	Å
$\text{Ca}^{2+}$	-0.1200	2.4357
O (de $\text{H}_2\text{O}$ TIP3P)	-0.1521	3.1575

TABLEAU I. Paramètres d'interaction de Lennard-Jones du  $\text{Ca}^{2+}$  et de l'oxygène d'une molécule d'eau.

Les résultats des calculs d'énergies libres faits à partir du paramètre final sont présentés au tableau II. Le résultat du calcul effectué avec la sphère de 100 molécules d'eau est de  $-384.6$  kcal/mol pour la partie électrostatique. L'énergie libre d'hydratation obtenue est donc de  $-382.0$  kcal/mol après recalcul incluant la contribution de la formation de la cavité.

D'autres ensembles de paramètres de Lennard-Jones du  $\text{Ca}^{2+}$  ont déjà été

$n$	$\Delta G$ (kcal/mol)
6	-388.7
25	-381.7
50	-382.6
100	-384.6

TABLEAU II. Contribution électrostatique à l'énergie libre d'hydratation du  $\text{Ca}^{2+}$  calculée pour les cas où le  $\text{Ca}^{2+}$  est simulé dans une sphère de  $n$  molécules d'eau.

utilisés dans des simulations effectuées par Ahlström et coll.,<sup>71</sup> Sussman et Weinstein,<sup>72</sup> Åqvist,<sup>69</sup> Prod'homme et Karplus,<sup>73</sup> et Obst et Bradaczek.<sup>74</sup> Lors d'une étude comparative de différents ensembles de paramètres, Åqvist<sup>69</sup> a utilisé les paramètres de Ahlström et coll.<sup>71</sup> et Sussman et Weinstein<sup>72</sup> ainsi que ceux développés dans son travail afin de reproduire l'énergie libre d'hydratation du  $\text{Ca}^{2+}$ . Il a trouvé que les paramètres des deux autres groupes d'auteurs<sup>71, 72</sup> ne reproduisaient pas la référence expérimentale de l'énergie libre d'hydratation et que seuls ses propres paramètres donnaient une réponse juste.<sup>69</sup> Cette étude laisse entendre que si les paramètres n'ont pas été développés en tenant compte de cette référence expérimentale, ils ont alors peu de chances de bien la reproduire.

Des trajectoires de 100 ps ont été générées pour des systèmes incluant le  $\text{Ca}^{2+}$  et 25, 50 ou 100 molécules d'eau. Ceci afin de comparer certaines propriétés statiques calculées avec le nouveau paramètre à leurs homologues provenant de données expérimentales. La figure 12 illustre la fonction de distribution radiale des atomes d'oxygène à partir du centre du  $\text{Ca}^{2+}$ . Le premier maximum de la fonction de distribution radiale correspond dans ce cas-ci à la distance  $\text{Ca}^{2+}$ -O (de l'eau) la plus probable mais aussi au rayon auquel on retrouve la première coquille d'hydratation. Dans tous les cas on trouve une valeur de 2.3 Å qui semble indépendante du nombre de molécules d'eau incluses dans la simulation, d'où

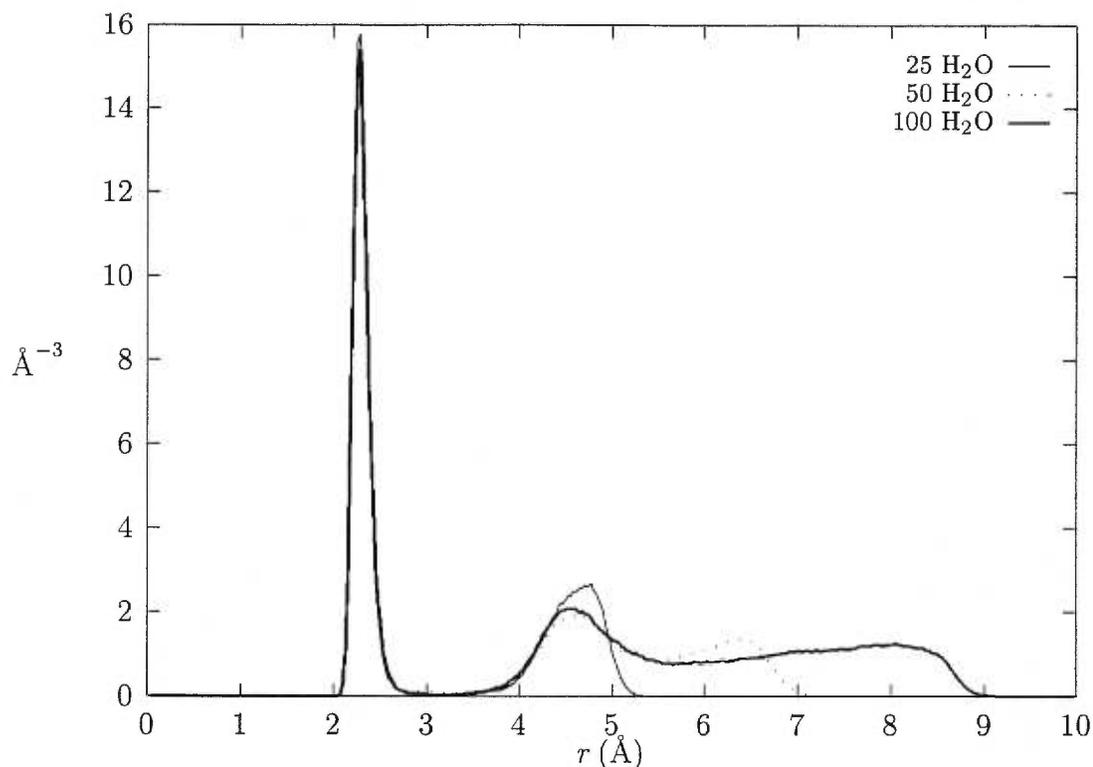


FIGURE 12. Distribution radiale  $g(r)$  ( $\text{\AA}^{-3}$  des atomes O (de l'eau) en fonction de la distance  $r$  à partir du centre du  $\text{Ca}^{2+}$ , calculée à partir des trajectoires de DM.

on peut conclure que des propriétés de dilution infinie ont bien été simulées. Des données expérimentales obtenues par diffraction de neutrons ou de rayons X stipulent que la majorité des distances  $\text{Ca}^{2+}$ -O (de l'eau) se retrouvent entre 2.3 et 2.5  $\text{\AA}$ ,<sup>75, 76, 77, 78</sup> avec une moyenne d'environ 2.42  $\text{\AA}$ , et varie proportionnellement avec la coordinence du  $\text{Ca}^{2+}$ . Bien qu'il ne s'agisse que d'une approximation statique à la définition dynamique de la coordinence, qui tient compte du temps de résidence des molécules d'eau autour de l'ion, une coordinence de 7.3 molécules d'eau a été obtenue en moyenne pour les trois cas en totalisant l'aire sous la courbe du premier maximum d'hydratation. Selon les données expérimentales,<sup>76, 78</sup> la

coordinnence la plus probable du  $\text{Ca}^{2+}$  est de sept ou huit molécules d'eau faisant partie d'une première coquille d'hydratation, mais quelques rares cas de six ou neuf liens de coordinnence ont été observés, et ce nombre dépend lui-même de la concentration de  $\text{Ca}^{2+}$  en solution à laquelle il est inversement proportionnel. On note aussi la présence de la seconde et la troisième coquille d'hydratation qui apparaissent avec l'augmentation de 25, à 50 et à 100 molécules d'eau.

Des paramètres de Lennard-Jones ont été attribués au  $\text{Ca}^{2+}$  de sorte à reproduire son énergie libre d'hydratation. De plus, les propriétés statiques du  $\text{Ca}^{2+}$  hydraté avec ces paramètres (fonctions de distribution radiale et coordinnences) sont très similaires aux données expérimentales. La prochaine section examine en détail la qualité des paramètres de Lennard-Jones en vérifiant leur comportement dans l'interaction avec d'autres molécules qui modélisent les ligands du  $\text{Ca}^{2+}$  *in situ*.

### 3.2.2 Calculs de courbes d'énergies d'interaction de systèmes $\text{Ca}^{2+}$ -ligands

Afin d'évaluer la qualité de la fonction de potentiel au niveau microscopique, les distances  $\text{Ca}^{2+}$ -ligand sont optimisées (minimisation de l'énergie) et les courbes d'énergie d'interaction  $\text{Ca}^{2+}$ -ligands sont toutes deux calculées et comparées avec leurs homologues *ab initio* ainsi qu'avec des données expérimentales lorsque disponibles.

Le  $\text{Ca}^{2+}$  interagit en premier lieu avec des atomes d'oxygène appartenant à un acide aminé ou à une molécule d'eau, et avec lesquels il établit des liaisons ioniques avec très peu de caractère covalent.<sup>45</sup> Afin d'alléger la taille des systèmes  $\text{Ca}^{2+}$ -acide aminé et de permettre des calculs *ab initio* en des temps raisonnables, chacun des acides aminés concernés est pris comme interagissant seul avec le  $\text{Ca}^{2+}$  et est représenté par une molécule modèle se prêtant plus facilement aux calculs.

Les acides aminés dont l'atome d'oxygène (lié au  $\text{Ca}^{2+}$ ) fait partie de la chaîne principale (site I: Ala15, Glu17, Asp19 et Gln22; site II: Glu60) sont modélisés par le N-méthyle-acétamide (NMA), ou par l'acétamide dans le cas de l'Asn56. Les autres acides aminés, dont l'atome d'oxygène fait partie d'une chaîne latérale (site I: Glu27 (*bidenté*); site II: Asp54, Asp58 et Glu65 (*bidenté*)), sont modélisés par une molécule de propionate. Le terme *bidenté* signifie que les deux atomes d'oxygène du groupement carboxylique du propionate sont liés symétriquement au  $\text{Ca}^{2+}$ . Les charges attribuées aux atomes d'une molécule modèle sont les mêmes que celles présentes sur l'acide aminé représenté dans le champ de force *param22*.<sup>51</sup> Ces molécules modèles d'acides aminés sont couramment utilisées aux mêmes fins et reconnues comme étant des modèles valables pouvant représenter des groupements peptidiques d'une protéine.<sup>79</sup> Dans les calculs classiques, les molécules d'eau liées au  $\text{Ca}^{2+}$  sont représentées par le modèle TIP3P<sup>60</sup> comme dans tous les calculs de DM de cette thèse. Et c'est aussi la géométrie du modèle TIP3P,<sup>60</sup> très proche de la géométrie d'une molécule d'eau en phase gazeuse, qui a été utilisée pour les calculs *ab initio*, mais dont la distribution des charges varie, bien sûr, en fonction des calculs *ab initio*. Les coordonnées de l'acétamide et du NMA proviennent de structures déterminées expérimentalement par diffraction d'un gaz d'électrons.<sup>80, 81</sup> Dans le cas du propionate, un ensemble complet de coordonnées déterminées expérimentalement pour tous les atomes n'était pas disponible. La structure du propionate a donc été optimisée avec la base 6-311G\*\*<sup>82, 83, 84</sup> au niveau Hartree-Fock (HF), à partir de coordonnées initiales elles-mêmes provenant d'une optimisation avec une base moins complète.<sup>85, 86</sup>

La géométrie intra-moléculaire a été fixée lors de tous les calculs. La figure 13 présente le modèle utilisé pour chaque acide aminé ayant un atome d'oxygène lié au  $\text{Ca}^{2+}$ . L'orientation relative entre le  $\text{Ca}^{2+}$  et la molécule est fixée dans tous les calculs, tel qu'illustré à la figure 13. Seule la distance entre le  $\text{Ca}^{2+}$  et l'atome d'oxygène de la molécule est variée et c'est en fonction de cette distance qu'est minimisée l'énergie (distance optimale) ou calculées les courbes d'énergies

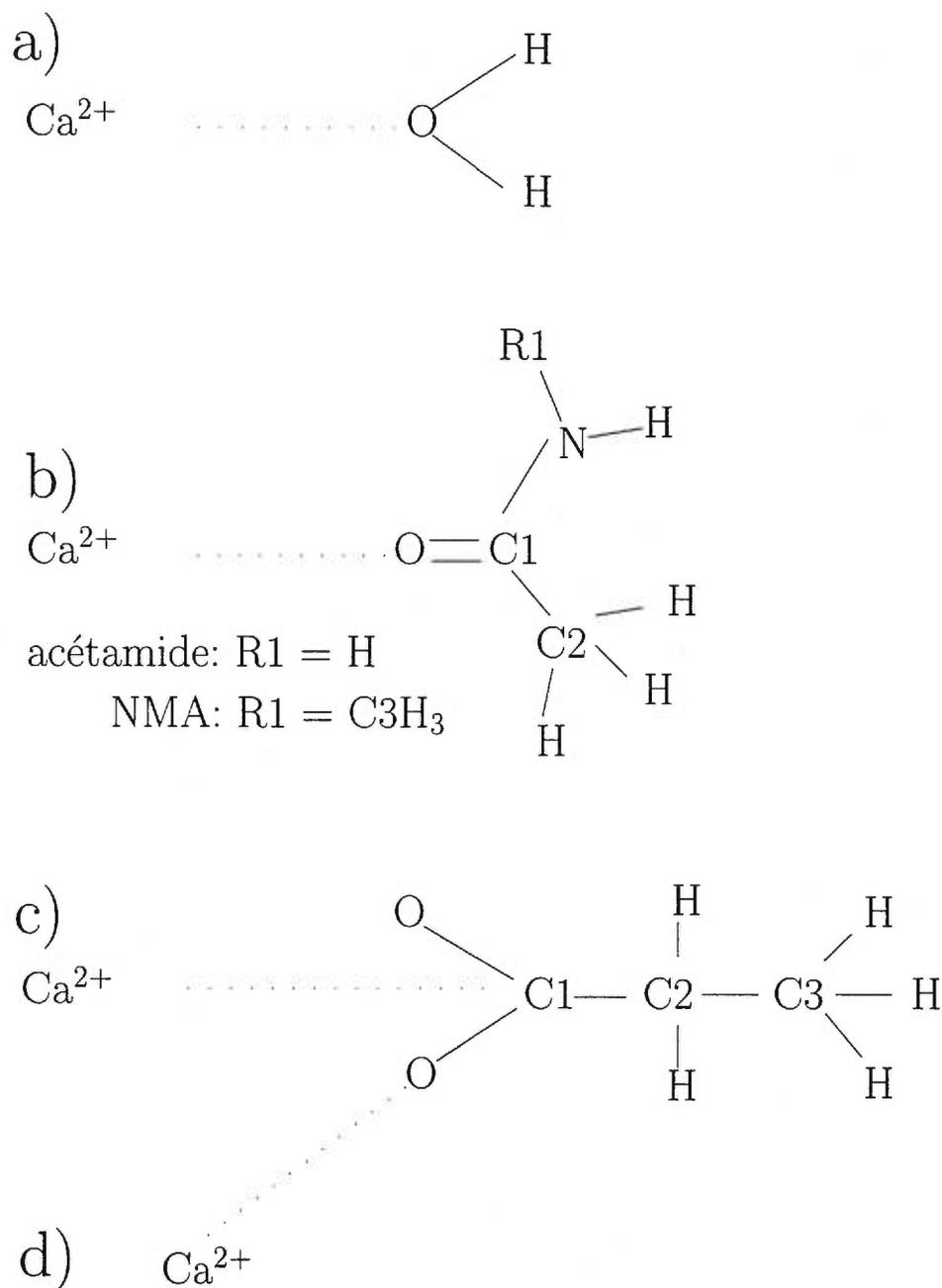


FIGURE 13. Structures et configurations des modèles utilisés pour représenter les interactions  $\text{Ca}^{2+}$ -ligand dans les sites de la calbindine  $\text{D}_{9k}$ : a)  $\text{Ca}^{2+}$ - $\text{H}_2\text{O}$ ; b)  $\text{Ca}^{2+}$ -acétamide et  $\text{Ca}^{2+}$ -NMA; c)  $\text{Ca}^{2+}$ -propionate (configuration *bidentée*); d)  $\text{Ca}^{2+}$ -propionate (configuration *monodentée*). Les lignes pointillées représentent l'axe selon lequel est variée la distance  $\text{Ca}^{2+}$ -O dans les calculs subséquents.

d'interaction.

Le logiciel Gaussian90<sup>50</sup> a été utilisé pour effectuer les calculs par la méthode HF. Une base 6-311G\*\*<sup>82, 83, 84</sup> a été utilisée pour les atomes C, N, O et H, soit une base double dzéta avec fonctions de polarisations. Pour le Ca<sup>2+</sup>, la base de Wachters<sup>87</sup> (14s, 9p) contractée par Gianolo et coll.<sup>88</sup> [6s, 3p] a été retenue et consiste en une base (62221+/621), donc double dzéta avec une fonction diffuse mais sans fonction de polarisation. Pour chaque calcul d'énergie d'interaction, l'énergie totale du système Ca<sup>2+</sup>-ligand est d'abord calculée, puis on doit soustraire l'énergie propre de chacune des espèces en interaction.

Il y a une erreur associée au fait de ne pas utiliser le même ensemble de fonctions de base lors du calcul de l'énergie propre de chacune des espèces, que lors du calcul pour le système total.<sup>89</sup> Le calcul de l'énergie du système total bénéficie d'un développement de fonctions de bases plus étendu (provenant des deux espèces), tandis que le calcul pour chacune des espèces est réalisé seulement à partir de la description de cette espèce. Lors du calcul de l'énergie d'interaction, il serait avantageux d'obtenir des erreurs presque égales lors des calculs de l'énergie du système total et des espèces isolées, car alors ces erreurs pourraient s'annuler plus efficacement lors de la différence et fournir un résultat plus juste. Cette erreur traditionnellement appelée l'erreur associée à la superposition des bases (*Basis Set Superposition Error*) a été corrigée par la méthode *counterpoise*,<sup>89</sup> c'est-à-dire que les énergies propres du Ca<sup>2+</sup> et du ligand ont été respectivement calculées en absence des noyaux du ligand et du Ca<sup>2+</sup> dont les fonctions de base demeurent incluses dans le calcul. Chaque espèce fantôme conserve sa description électronique (sa base) qui est utilisée dans le calcul de l'énergie propre de l'autre espèce afin d'obtenir un calcul de meilleure qualité. Ces énergies d'interaction sont calculées à tout une succession de points sur les axes Ca<sup>2+</sup>-O tels que définis à la figure 13.

Le tableau III donne les résultats des calculs de minimisation de l'énergie

ligand	$r_{min}(\text{Å})$		énergie (kcal/mol)	
	HF	<i>param22</i>	HF	<i>param22</i>
H <sub>2</sub> O	2.3381	2.1802	-50.9392	-50.7460
acétamide	2.1295	2.1075	-82.8432	-66.1340
NMA	2.1840	2.1171	-85.2561	-62.4150
propionate				
<i>monodenté</i>	2.0813	1.9885	-252.0915	-248.6561
<i>bidenté</i>	2.5538	2.3284	-283.7274	-300.3375

TABLEAU III. Distances optimales et énergies minimales de l'interaction Ca<sup>2+</sup>-ligand, calculées par la méthode HF (base 6-311G\*\*<sup>82, 83, 84</sup> pour les atomes C, N, O et H, et base de Wachters<sup>87</sup> contractée par Gianolo et coll.<sup>88</sup> pour le Ca<sup>2+</sup>.) et par le logiciel CHARMM (*param22*<sup>51</sup>).

du système Ca<sup>2+</sup>-ligand en fonction de la distance Ca<sup>2+</sup>-O telle que définie à la figure 13. Il faut rappeler que la contribution de l'énergie de dispersion, de nature électrodynamique, n'est pas incluse dans les calculs d'énergies statiques HF. Malgré ce manque, l'utilisation de cette méthode *ab initio* est considérée valable puisque la contribution de l'énergie de dispersion ne représente qu'entre 1 et 7% de l'énergie totale dans le cas des ions doublement chargés positivement.<sup>90</sup>

Les résultats *ab initio* pour l'interaction Ca<sup>2+</sup>-H<sub>2</sub>O sont en accord avec les résultats obtenus par d'autres auteurs,<sup>79, 91, 92</sup> et les petites différences entre les résultats proviennent de la méthode (HF, MP2 ou MP3: 2-3 kcal/mol, 0.05 Å) et/ou de la base utilisée qui engendre des différences plus considérables tout dépendant de la présence de fonctions de polarisation. Notons que les méthodes MP2 et MP3 incluent la contribution de l'énergie de dispersion. Bien que l'énergie minimale calculée avec *param22* s'accorde avec le résultat HF, la distance optimisée avec *param22* est sous-estimée par 0.158 Å. Cette tendance est maintenue pour l'acétamide (-0.0219 Å), le NMA (-0.0669 Å) et le propionate (-0.0928 et -0.2150 Å), et leur énergie minimale respective est sous-estimée, sauf pour le

cas du propionate dans la configuration *bidentée*. Le meilleur accord après la molécule d'eau est le propionate *monodenté* et le plus mauvais le NMA. Il faut encore rappeler que les résultats HF dépendent beaucoup de la base utilisée.

Des données expérimentales de cristallographie concernant la géométrie des interactions  $\text{Ca}^{2+}$ -groupement carboxyl ont été compilées pour différents types de ligands.<sup>93</sup> De cette compilation, on dégage les tendances générales suivantes concernant la distance  $\text{Ca}^{2+}$ -O dans un contexte où la coordinence du  $\text{Ca}^{2+}$  est de sept ou de huit: 1) la majorité des distances  $\text{Ca}^{2+}$ -O se situe entre 2.3 et 2.5 Å avec des moyennes respectives pour les configurations *mono* et *bidentée* de 2.38 et 2.53 Å. Les configurations *monodentées* ont donc tendance à être plus courtes; 2) cette distance augmente avec la coordinence du  $\text{Ca}^{2+}$ . Les distances optimales calculées par *param22* (tableau III) ont donc de bonnes chances de reproduire les distances  $\text{Ca}^{2+}$ -ligands *in situ*. Une analyse plus approfondie de l'ensemble des paramètres est donnée suite à la présentation aux pages suivantes des courbes d'énergies d'interaction  $\text{Ca}^{2+}$ -ligands.

Le tableau IV donne une liste des charges partielles présentes sur les atomes des molécules modèles en présence et en absence du  $\text{Ca}^{2+}$  lors des calculs *ab initio* (analyse de charges de Mulliken à la distance  $\text{Ca}^{2+}$ -O optimisée), et dans le champ de force *param22*<sup>51</sup> (charge fixe attribuée à chaque atome). Il est important de souligner que les analyses de charges de Mulliken sont très sensibles aux choix de base *ab initio* et qu'une base plus complète donnera de meilleurs résultats.

molécule	atome	HF		<i>param22</i> <sup>51</sup>
		avec Ca <sup>2+</sup>	sans Ca <sup>2+</sup>	
H <sub>2</sub> O	Ca <sup>2+</sup>	1.98	-	2.00
	O	-0.76	-0.50	-0.83
NH <sub>2</sub> C1(C2H <sub>3</sub> )O acétamide	Ca <sup>2+</sup>	1.97	-	2.00
	O	-0.89	-0.59	-0.55
	C1	0.60	0.73	0.55
	C2	-0.29	-0.55	-0.27
NH(C3H <sub>3</sub> )C1OC2H <sub>3</sub> NMA	Ca <sup>2+</sup>	1.97	-	2.00
	O	-0.90	-0.61	-0.51
	C1	0.60	0.72	0.51
	C2	-0.09	-0.19	-0.27
	C3	-0.28	-0.55	-0.11
C3H <sub>3</sub> C2H <sub>2</sub> C1O1O2 propionate <i>monodenté</i>	Ca <sup>2+</sup>	1.95	-	2.00
	O1	-1.01	-0.66	-0.76
	O2	-0.50	-0.64	-0.76
	C1	0.61	0.50	0.62
	C2	-0.29	-0.27	-0.28
	C3	-0.20	-0.20	-0.27
C3H <sub>3</sub> C2H <sub>2</sub> C1O1O2 propionate <i>bidenté</i>	Ca <sup>2+</sup>	1.883	-	2.00
	O1	-0.74	-0.66	-0.76
	O2	-0.72	-0.64	-0.76
	C1	0.46	0.50	0.62
	C2	-0.24	-0.27	-0.28
	C3	-0.20	-0.20	-0.27

TABLEAU IV. Charge partielle présente sur chacun des atomes des molécules modèles à la distance Ca<sup>2+</sup>-O optimale (voir tableau III). Les charges *ab initio* proviennent d'analyses de charges de Mulliken (base 6-311G\*\*<sup>82, 83, 84</sup> pour les atomes C, N, O et H et base de Wachters<sup>87</sup> contractée par Gianolo et coll.<sup>88</sup> pour le Ca<sup>2+</sup>).

Néanmoins, ces résultats suggèrent que les charges attribuées à ces molécules modèles, et donc aux acides aminés qu'elles représentent, ne sont pas parfaitement adaptées au cas où un ion divalent est présent dans le système. Cette observation est d'autant plus valable pour les atomes d'oxygène des molécules modèles. Il ne faut pas non plus négliger la charge du  $\text{Ca}^{2+}$  qui déjà diminue lorsqu'il est en interaction avec un seul ligand. Bien que cette analyse de charges dépende fortement de la base utilisée, il semble tout de même qu'un champ de forces tel que *param22*<sup>51</sup> avec des charges fixes ne tient pas compte d'effets électroniques importants qui surviennent lorsqu'un ion est présent dans le système.

Les courbes d'énergies d'interaction présentées aux figures 14 à 16 donnent un autre aperçu de la qualité des paramètres du  $\text{Ca}^{2+}$ , mais dépendent aussi des autres paramètres du champ de force *param22*<sup>51</sup> qui ont été utilisés. L'interaction  $\text{Ca}^{2+}$ - $\text{H}_2\text{O}$  est celle qui est la mieux représentée, peut-être entre autres pour ces raisons: tout d'abord, les paramètres du  $\text{Ca}^{2+}$  ont été ajustés en fonction de l'énergie libre d'hydratation du  $\text{Ca}^{2+}$ ; et ensuite, la molécule d'eau est la plus petite à interagir avec le  $\text{Ca}^{2+}$ , ce qui minimise l'erreur totale sur la paramétrisation de l'interaction entre ces deux espèces; de plus, le dipôle du modèle TIP3P (2.35 Debye) est surestimé par 27% par rapport à la valeur expérimentale connue (1.855 Debye) afin de compenser l'absence de polarisation induite.

L'acétamide et le NMA sont, tout comme l' $\text{H}_2\text{O}$ , des molécules neutres mais elles ne sont pas aussi bien représentées dans leur interaction avec le  $\text{Ca}^{2+}$ . Les paramètres des acides aminés et des molécules modèles ont été déterminés afin de bien représenter leurs interactions avec d'autres atomes neutres (bien qu'ayant des charges partielles sur certains atomes), comme le témoigne le tableau IV. On note à la figure 15 que la courbe *ab initio* du NMA (bas) présente un puits plus profond que celle de l'acétamide (haut), ce qui n'est pas le cas pour les mêmes courbes produites avec *param22* où l'on observe en fait le contraire. Le puits plus profond pour l'interaction  $\text{Ca}^{2+}$ -acétamide dans le cas de *param22* s'accorde avec

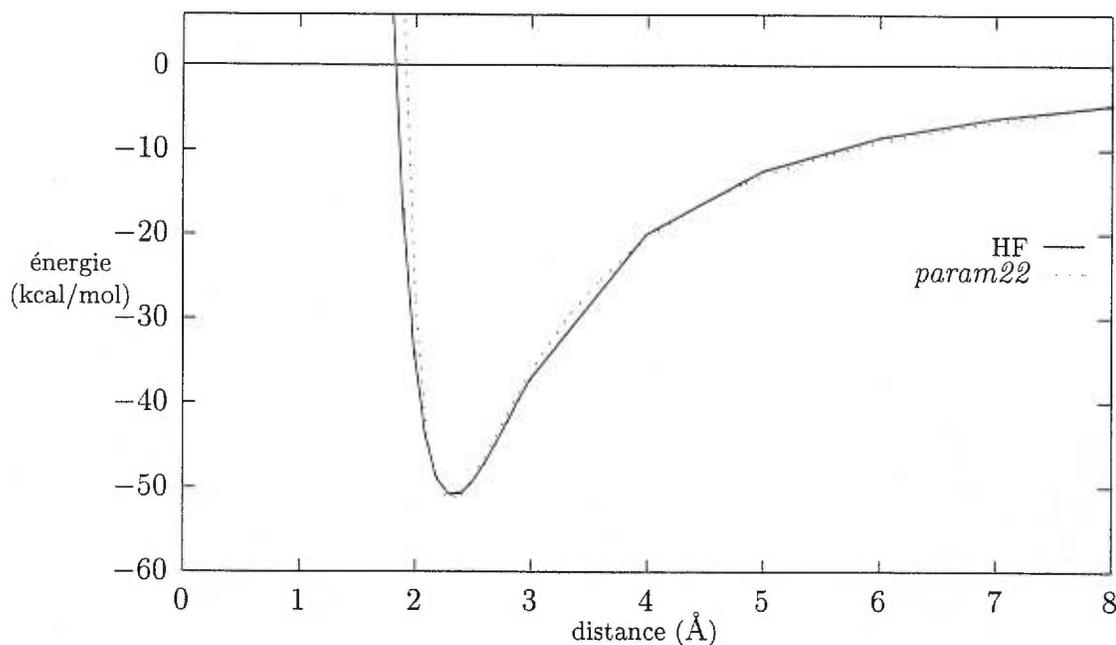


FIGURE 14. Courbes d'énergies d'interaction  $\text{Ca}^{2+}$ - $\text{H}_2\text{O}$  en fonction de la distance telle que définie à la figure 13.

le fait que la charge fixée sur l'oxygène de l'acétamide ( $-0.55e$ ) est légèrement plus élevée que sur le NMA ( $-0.51e$ ), contrairement au résultat *ab initio* où c'est le NMA qui possède une plus forte charge.

Les courbes d'énergie d'interaction  $\text{Ca}^{2+}$ -propionate ne s'étendent pas plus loin que 6 Å. En effet, les calculs *ab initio* effectués au niveau HF pour le cas du propionate en interaction avec le  $\text{Ca}^{2+}$  donnent des énergies positives (répulsion) à plus de 6 Å et la courbe semble alors se comporter comme s'il y avait un problème de dissociation, c'est-à-dire que la répartition spatiale des électrons entre les deux espèces n'est pas effectuée correctement. De fait, à plus de 6 Å, la charge du calcium s'établit à environ  $+1.4e$  tandis que les atomes d'oxygène se répartissent la quantité de charge manquante au calcium de façon différente selon qu'il s'agit de

la configuration *monodentée* ou *bidentée*. Bien que le propionate soit une molécule à couches fermées lorsque liée au  $\text{Ca}^{2+}$ , ce n'est cependant plus le cas à plus de 6 Å, ce qui implique que la méthode HF où l'on impose une forme à la fonction d'onde (HF restreint) ne fonctionne pas bien. Il s'agirait là d'un cas similaire à la dissociation de la molécule  $\text{H}_2$  pour laquelle la méthode HF restreinte donne une distribution des électrons identiques aux deux atomes parce que l'on impose que les deux électrons se trouvent dans la même orbitale. La méthode HF non-restreinte (équations de Pople-Nesbet) apporte habituellement une solution à ce problème en n'imposant pas que la fonction d'onde du complexe soit une fonction propre de l'opérateur de spin, comme c'est le cas pour la méthode restreinte.

Le puits calculé avec *param22* pour la configuration *bidentée* est plus profond que son homologue *ab initio*. Il faut rappeler que les charges attribuées aux atomes du propionate lors des calculs effectués avec *param22* sont les mêmes dans les cas *mono* et *bidenté*. Les courbes d'interactions présentées à la figure 16 montrent que l'interaction *monodentée* avec le propionate est relativement bien représentée et offre un bien meilleur résultat que l'interaction *bidentée*. Ceci est un peu surprenant, puisque dans *param22* les charges de la configuration *monodentée* sont réparties également sur les deux atomes d'oxygène plutôt que d'être polarisées sur l'atome qui interagit avec le  $\text{Ca}^{2+}$ , comme c'est le cas lors des calculs *ab initio*. La figure 17 montre l'interaction  $\text{Ca}^{2+}$ -propionate *bidenté* dont la charge du  $\text{Ca}^{2+}$  a été diminuée à  $+1.88e$  dans *param22* afin représenter un  $\text{Ca}^{2+}$  ayant subi un transfert de charge, et compare ce résultat avec la même courbe *ab initio* présentée à la figure 16 (bas). Bien qu'incomplet, puisque les charges sur les atomes de la molécule modèle n'ont pas été changées, ce résultat montre la sensibilité de la fonction de potentiel aux charges atomiques. Il apparaît qu'un modèle où les charges sont fixes offre peu de possibilités d'amélioration, entre autres pour le  $\text{Ca}^{2+}$  et les atomes du propionate qui devraient en principe avoir des charges différentes pour les configurations *monodentée* et *bidentée*. La section suivante vise donc à définir et à évaluer une méthode d'amélioration du modèle.

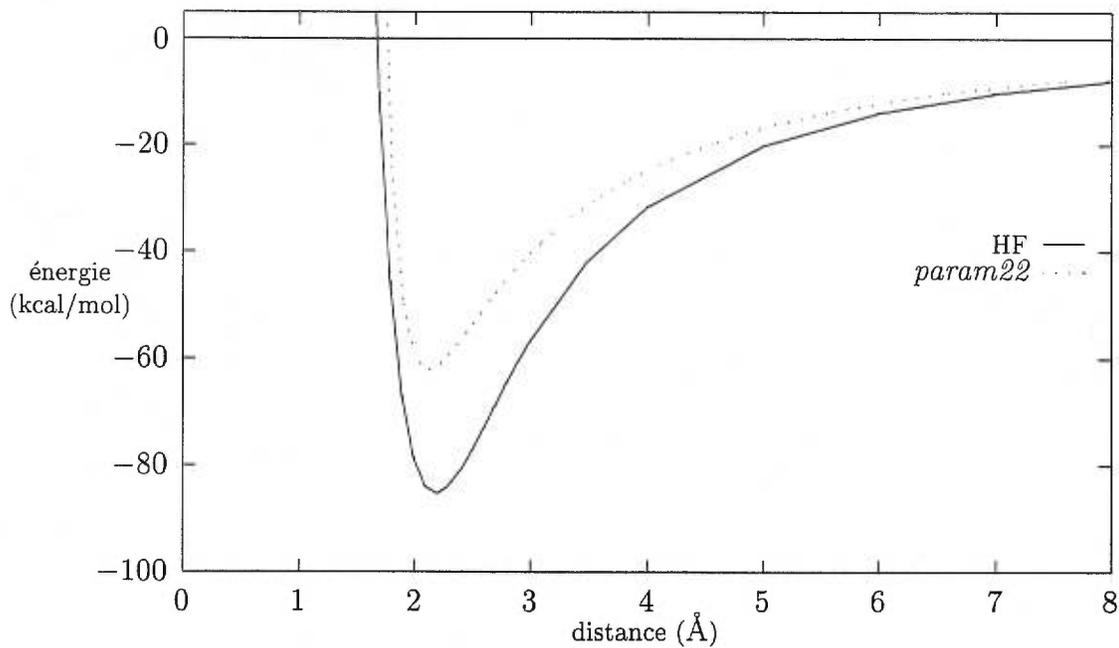
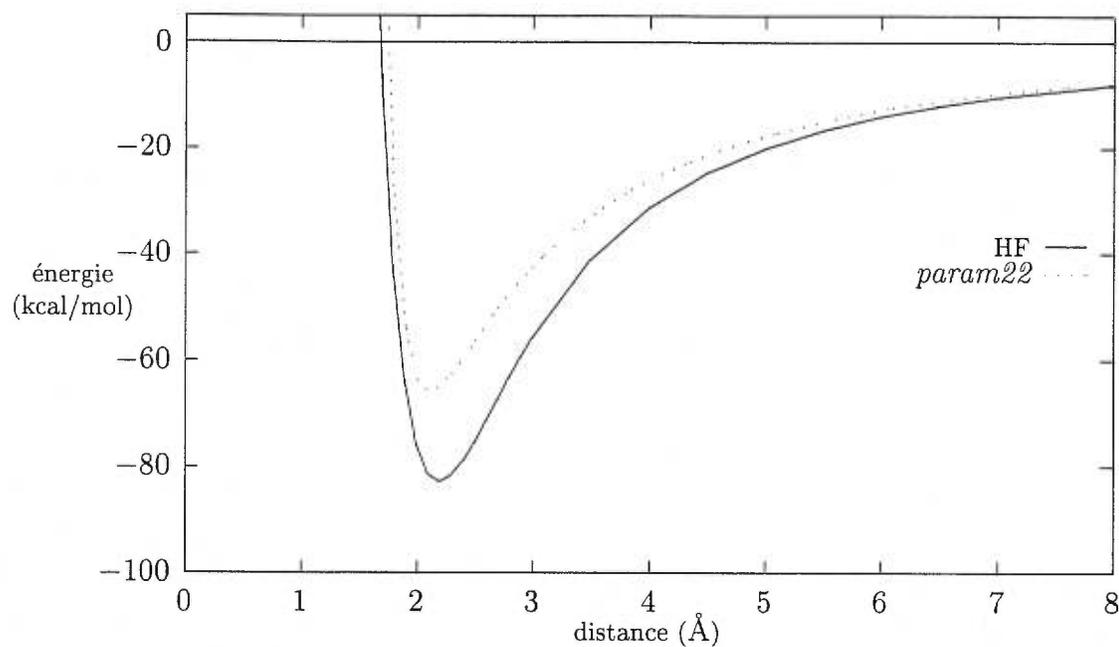


FIGURE 15. Courbes d'énergies d'interaction Ca<sup>2+</sup>-acétamide (haut) et Ca<sup>2+</sup>-NMA (bas) en fonction de la distance telle que définie à la figure 13.

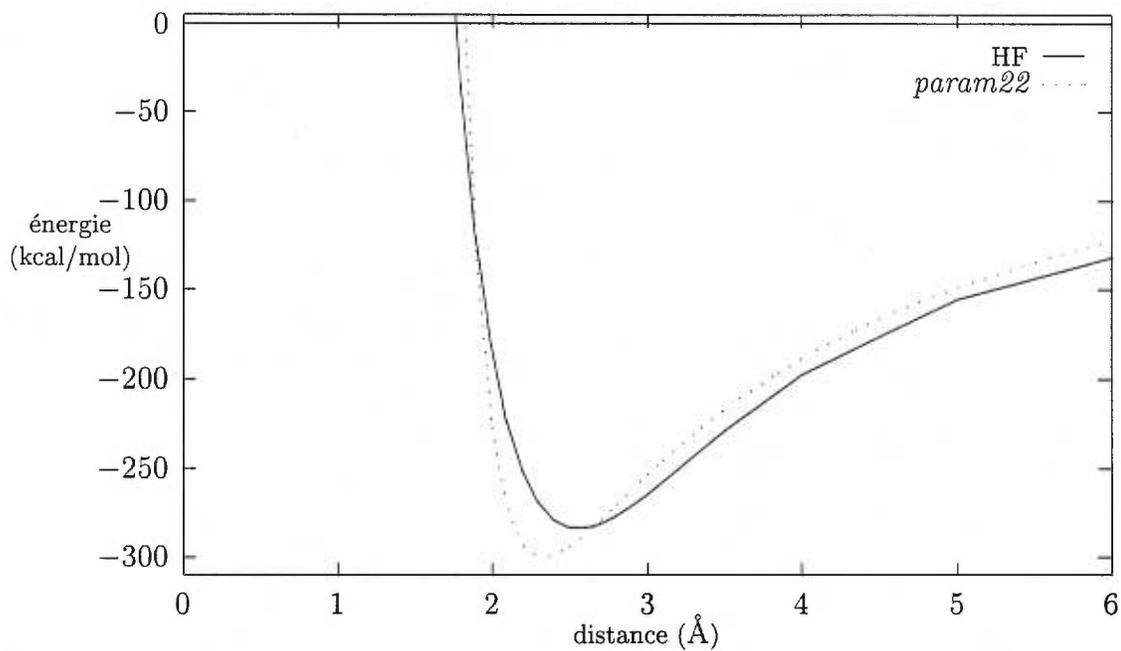
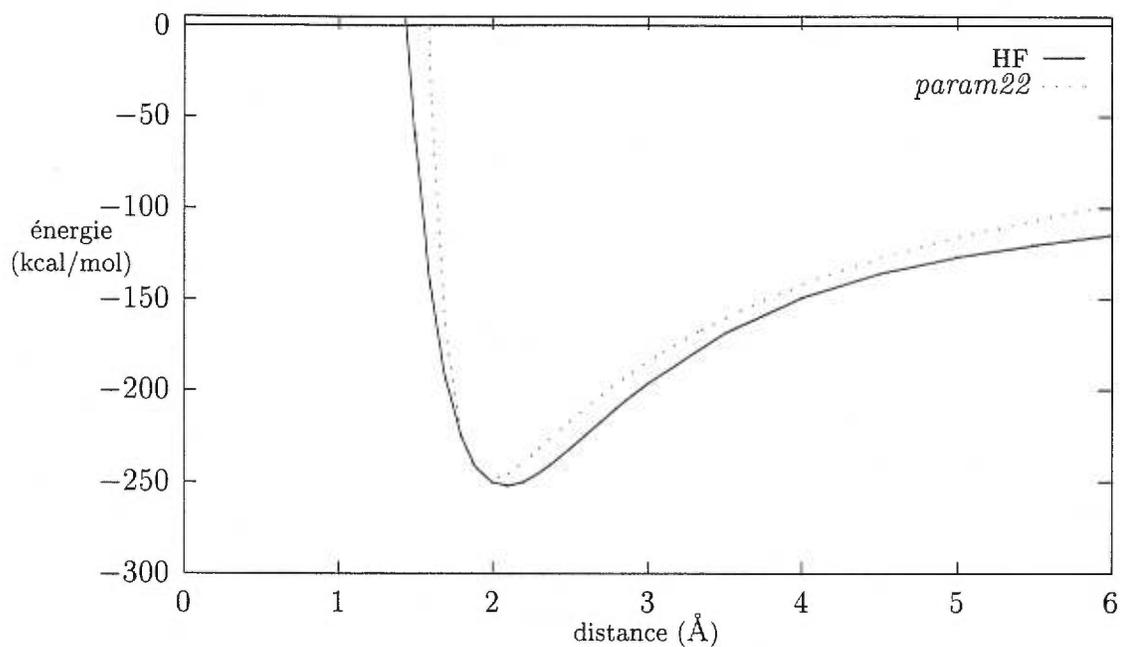


FIGURE 16. Courbes d'énergies d'interaction  $\text{Ca}^{2+}$ -propionate *monodenté* (haut) et *bidenté* (bas) en fonction de la distance telle que définie à la figure 13.

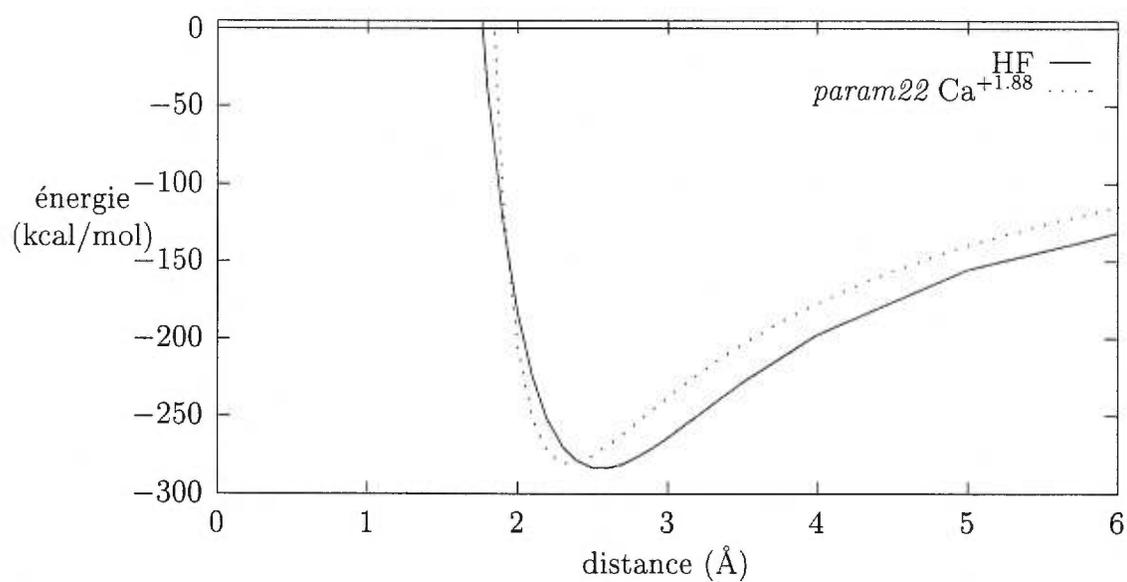


FIGURE 17. Courbes d'énergies d'interaction  $\text{Ca}^{2+}$ -propionate *bidenté*. Dans la courbe de *param22*, le calcium a une charge de  $+1.88e$ . La courbe HF est la même qu'à la figure 16 (bas).

### 3.3 Évaluation de la partie non-additive des interactions $\text{Ca}^{2+}$ - $\text{H}_2\text{O}$ pour des agrégats de deux à six $\text{H}_2\text{O}$

Un modèle idéal, dans le contexte de la DM, serait de rendre tous les atomes lourds polarisables, y compris le  $\text{Ca}^{2+}$ . Cependant, la tentative faite ici d'améliorer le modèle et présentée dans cette prochaine section réduit cette tâche au minimum afin d'évaluer la faisabilité du projet. Les atomes lourds ne sont donc polarisés que par le  $\text{Ca}^{2+}$  qui, pour sa part, conserve sa charge fixée à  $+2e$ .

L'étalon choisi afin d'ajuster le nouveau modèle d'atomes polarisables consiste en la contribution dite non-additive, associée aux effets à plusieurs corps. En fait, il a été vérifié que cette contribution non-additive correspond approximativement à l'énergie de polarisation induite au second ordre pour les interactions ion- $(\text{H}_2\text{O})_n$ <sup>94</sup> et ion-peptide.<sup>95</sup> On vise donc ici à évaluer la contribution non-additive, ou effets à plusieurs corps, des interactions du  $\text{Ca}^{2+}$  avec des agrégats de deux à six molécules d'eau, c'est-à-dire la différence entre l'énergie totale et la somme de toutes les paires d'énergies d'interactions entre atomes. Il a été proposé que ces effets étaient négligeables dans les interactions entre molécules neutres.<sup>96, 97</sup> Cependant, la présence d'un ion monovalent en interaction avec des molécules d'eau<sup>94</sup> ou des groupements peptidiques<sup>95</sup> rend cette approximation inadéquate, notamment parce que les effets à plusieurs corps constituent un facteur qui influence la structure des premières coquilles d'hydratation de l'ion<sup>98, 99</sup> et donc de la coordinence. En particulier, les propriétés qui dépendent de la dynamique ou des interactions à longue portée (constantes diélectriques, différents temps de relaxation) sont particulièrement sensibles aux effets à plusieurs corps.<sup>74, 100</sup> De plus, l'importance de ces derniers augmente avec la charge du cation et du nombre de molécules d'eau qui l'entourent.<sup>100</sup> On s'attend donc à ce que cette contribution non-additive en présence du  $\text{Ca}^{2+}$  soit assez importante.

Bien qu'aucun terme explicite d'énergie n'apparaisse pour le moment dans

la fonction de potentiel utilisée, cette contribution est tout de même incluse de façon moyenne par le biais de paramètres adéquats qui reproduisent les propriétés macroscopiques telles que l'énergie libre d'hydratation du  $\text{Ca}^{2+}$ . Mais toutes les propriétés macroscopiques ne sont pas nécessairement bien reproduites par ce modèle, non plus que la justesse de la réalité physique des interactions au niveau microscopique. L'évaluation de la contribution non-additive se veut un point de départ dans l'élaboration d'un modèle où apparaîtrait explicitement un terme d'énergie qui rendrait compte de cette contribution.

L'énergie d'interaction totale entre les molécules d'un système est donnée par une série de termes d'interactions à deux, trois, quatre, ...  $n$  corps:

$$\begin{aligned} U^{\text{totale}} &= \sum U(i, j) + \sum U(i, j, k) + \sum U(i, j, \dots, n) \\ &= \sum U(2 \text{ corps}) + \sum U(n \text{ corps}) \end{aligned} \quad (3.1)$$

de sorte que la contribution des effets à plusieurs corps (non-additive) peut être évaluée par

$$\sum U(n \text{ corps}) = U^{\text{totale}} - \sum U(2 \text{ corps}) \quad (3.2)$$

et pour le cas du  $\text{Ca}^{2+}$  entouré de  $n$  molécules d'eau, on a

$$\sum U(2 \text{ corps}) = \sum_i^n U(\text{Ca}^{2+}\text{-H}_2\text{O}_i) + \sum_{i>j}^n U(\text{H}_2\text{O}_i\text{-H}_2\text{O}_j) \quad (3.3)$$

Des calculs *ab initio* au niveau de HF avec une base 6-311G\*\*<sup>82, 83, 84</sup> pour les atomes O et H, et la base de Wachters<sup>87</sup> contractée par Gianolo et coll.<sup>88</sup> pour le  $\text{Ca}^{2+}$  ont été utilisés (voir section 3.2.2) afin d'effectuer les calculs pour les systèmes allant de une à six molécules d'eau. L'énergie totale du système est tout d'abord minimisée par rapport aux coordonnées intermoléculaires afin de trouver la géométrie optimale et de calculer  $U^{\text{totale}}$ . Puis, à cette géométrie optimale, les énergies d'interaction  $\sum_i^n U(\text{Ca}^{2+}\text{-H}_2\text{O}_i)$  et  $\sum_{i>j}^n U(\text{H}_2\text{O}_i\text{-H}_2\text{O}_j)$  sont calculées, suite à quoi, la contribution des effets à plusieurs corps est évaluée (équation 3.1). Les

$n$	$r_{\min.}$	$U^{\text{totale}}$	$\sum U(2 \text{ corps})$		$\sum U(n \text{ corps})$
		$\sum_i^n U(\text{Ca}^{2+}-\text{H}_2\text{O}_i)$	$\sum_{i>j} U(\text{H}_2\text{O}_i-\text{H}_2\text{O}_j)$		
1	2.3381	-50.9392	-50.9392	—	—
2	2.3589	-98.2782	-101.6860	1.1002	2.3076
3	2.3757	-141.1456	-152.3868	4.7940	6.4472
4	2.4027	-176.1998	-202.5312	12.0334	14.2980
6	2.4403	-238.5524	-302.3646	31.1118	32.7004

TABLEAU V. Évaluation de la contribution non-additive dans les calculs *ab initio*. La distance  $r_{\min.}$  est en Å et les énergies en kcal/mol.

résultats sont donnés au tableau V. Les deux dernières colonnes de ce tableau correspondent à des interactions répulsives, soit les interactions H<sub>2</sub>O-H<sub>2</sub>O dont les dipôles ne sont pas disposés favorablement, et la contribution non-additive, qui est raisonnablement bien approximée par l'énergie de polarisation induite au second ordre<sup>95</sup> et qui correspond elle aussi à des interactions répulsives reliées à l'induction de dipôles supplémentaires sur les molécules d'eau. Il ressort que l'approximation qui consiste à additionner les interactions électrostatiques par paires néglige cette contribution importante dans le cas du Ca<sup>2+</sup>hydraté. La contribution non-additive pour le cas du Ca<sup>2+</sup> hydraté est environ deux fois celle pour le cas du Na<sup>+</sup> pour lequel il est reconnu qu'une telle approximation n'est pas valable.<sup>100</sup>

Des calculs similaires ont été effectués avec *param22* et les paramètres du Ca<sup>2+</sup> définis à la section précédente, en conservant les mêmes configurations angulaires des molécules d'eau autour du Ca<sup>2+</sup>. La différence avec les calculs *ab initio* réside en ce qu'il n'est pas possible de calculer la même contribution à plusieurs corps avec CHARMM. Les résultats sont présentés au tableau VI. On y voit que la distance optimale calculée est plus courte que pour les calculs *ab initio* et que l'énergie d'interaction est plus favorable dans tous les cas présentés. Ceci est probablement, en partie, une conséquence du fait que

l'énergie d'interaction à deux corps  $\text{Ca}^{2+}\text{-H}_2\text{O}$  est surestimée. La dernière colonne, obtenue par soustraction de la quatrième colonne de la troisième, représente pour un agrégat la somme des interactions  $\text{H}_2\text{O}\text{-H}_2\text{O}$ . Ce résultat est surestimé par rapport au résultat *ab initio* homologue. Mais si l'on se rappelle que le dipôle sur la molécule d'eau TIP3P<sup>60</sup> est surestimé par 27%, ce surplus ne représente que la répulsion supplémentaire due au dipôle induit.

$n$	$r_{\min.}$	$U^{\text{totale}}$	$\sum_i^n U(\text{Ca}^{2+}\text{-H}_2\text{O}_i)$	$\sum_{i>j} U(\text{H}_2\text{O}_i\text{-H}_2\text{O}_j)$
1	2.1802	-52.7460	-52.7460	—
2	2.1840	-103.9087	-105.4858	1.5771
3	2.1910	-151.9044	-158.1927	6.2883
4	2.2131	-192.8637	-210.4632	17.5995
6	2.2438	-269.4158	-313.7772	44.3614

TABLEAU VI. Évaluation des énergies d'interactions à deux corps dans les calculs effectués avec CHARMM et *param22*.

molécule	atome	<i>ab initio</i>
$\text{Ca}^{2+}\text{-H}_2\text{O}$	$\text{Ca}^{2+}$	1.976
	O	-0.758
$\text{Ca}^{2+}\text{-(H}_2\text{O)}_2$	$\text{Ca}^{2+}$	1.958
	O	-0.746
$\text{Ca}^{2+}\text{-(H}_2\text{O)}_3$	$\text{Ca}^{2+}$	1.934
	O	-0.726
$\text{Ca}^{2+}\text{-(H}_2\text{O)}_4$	$\text{Ca}^{2+}$	1.930
	O	-0.70
$\text{Ca}^{2+}\text{-(H}_2\text{O)}_6$	$\text{Ca}^{2+}$	1.900
	O	-0.666

TABLEAU VII. Charges sur les  $\text{Ca}^{2+}$  et sur l'oxygène des molécules d'eau en fonction du nombre de ces dernières.

Afin de mieux comprendre la raison à l'origine de ces différences, une analyse

de charges de Mulliken, effectuée lors des calculs *ab initio*, est présentée au tableau VII. Il semble que la charge attribuée à l'oxygène du modèle de molécule d'eau TIP3P<sup>60</sup> ( $-0.834$ ) soit surestimée et que la charge du  $\text{Ca}^{2+}$  dans *param22* devrait être nettement inférieure à  $+2e$  dans les interactions avec les molécules d'eau. Des calculs homologues ont été effectués par Kaufman-Katz et coll.<sup>91</sup> afin d'évaluer le transfert de charge négative sur le  $\text{Ca}^{2+}$  avec des calculs au niveau MP2 et une base du type Huzinaga<sup>47</sup> avec des fonctions de polarisation et diffuses pour tous les atomes. Ils ont donné une charge encore plus faible sur le  $\text{Ca}^{2+}$  en fonction du nombre de molécules d'eau. Par exemple, la charge du  $\text{Ca}^{2+}$  serait de  $+1.54e$  dans un agrégat de huit molécules d'eau. Bien sûr, ce résultat dépend considérablement de la base de fonctions utilisée mais il demeure qu'un modèle de sphère à charge fixée à  $+2e$  ne semble pas être la façon optimale de représenter le  $\text{Ca}^{2+}$  et il faudrait donc s'attendre à quelques manquements liés au transfert de charge.

A la lumière de ces résultats et de ceux présentés à la section 3.2.2, il apparaît qu'un modèle plus réaliste passerait par l'inclusion de termes d'énergie qui dépendraient de la polarisabilité de chaque atome afin de mieux représenter les interactions électrostatiques. Cela impliquerait, comme le suggèrent les résultats, la détermination d'un nouvel ensemble de paramètres adaptés à la nouvelle fonction de potentiel. Dans le but d'inclure ce terme dans la fonction de potentiel totale, la prochaine section présente une méthode pour déterminer la polarisabilité de chacun des atomes faisant partie des molécules modèles en interaction avec le  $\text{Ca}^{2+}$ .

### 3.4 Évaluation de polarisabilités approchées pour les atomes lourds liés au $\text{Ca}^{2+}$ dans les sites de la calbindine $\text{D}_{9k}$

Le but de cette section est d'obtenir des valeurs de référence pour les polarisabilités des atomes lourds et des groupements chimiques des molécules modèles dans le même contexte où ont été effectués les autres calculs de cette thèse.

La méthode traditionnelle pour décrire les effets de polarisation induite consiste à calculer les variations du moment dipolaire moléculaire en présence d'un champ électrique uniforme  $\mathbf{E}$ :

$$\mu_{\text{mol.}} = \alpha_{\text{mol.}} \mathbf{E}. \quad (3.4)$$

Cependant, cette approche traditionnelle ne permet pas d'attribuer les effets de polarisation aux différents atomes dans une molécule. Par conséquent, elle est d'une utilité limitée dans le contexte qui nous intéresse puisque nous cherchons à développer des polarisabilités atomiques qui seraient généralisables pour une protéine. Pour distinguer les effets de polarisation induite sur chacun des atomes, il est nécessaire de soumettre la molécule à un champ électrique inhomogène. A cette fin, les effets de la réorganisation électronique en présence d'une charge ponctuelle *test* (+1) placée successivement à différents endroits autour de la molécule ont été examinés. S'il n'y avait aucun réarrangement de la structure électronique de la molécule, l'énergie totale serait donnée par:

$$U_{\text{tot.}} = U_{\text{mol.}} + q_{\text{test}} V(\mathbf{r}_{\text{test}}) \quad (3.5)$$

où  $V(\mathbf{r}_{\text{test}})$  est le potentiel électrostatique généré par la molécule au point  $\mathbf{r}_{\text{test}}$  en *absence de charge ponctuelle*  $q_{\text{test}}$ . Ainsi, l'énergie associée à la réorganisation électronique en présence d'une charge ponctuelle correspond à la différence entre l'énergie totale de l'équation 3.5 (cas où la molécule est non-perturbée), et l'énergie totale *ab initio* de la molécule perturbée en présence de la charge ponctuelle. Cette méthode avec charge ponctuelle permet également d'alléger le traitement et les calculs *ab initio*. Elle est approximative et simple, et permet d'obtenir des estimations qui s'accordent assez bien avec d'autres données obtenues d'autres méthodes.<sup>101, 102, 103, 104</sup>

Afin d'extraire des coefficients pour les polarisabilités atomiques, nous postulons que:

$$U_{\text{tot.}} - U_{\text{mol.}} - q_{\text{test}} V(\mathbf{r}_{\text{test}}) = U_{\text{pol.}}^{(1)}. \quad (3.6)$$

où  $V(\mathbf{r}_{\text{test}})$  est cette fois le potentiel électrostatique généré au point où se trouve la charge ponctuelle par toutes les charges de la molécule (noyaux et électrons); la stratégie consiste donc à faire l'hypothèse que les effets de polarisation induite peuvent être concentrés dans le terme d'énergie d'interaction au premier ordre (terme de droite à l'équation 3.6, c'est-à-dire le terme d'énergie des interactions charge-dipôle induit). Ce terme est défini par:<sup>95</sup>

$$U_{\text{pol.}}^{(1)} = -\frac{1}{2} \sum_i \mu_i \cdot \mathbf{E}_i \quad (3.7)$$

Le dipôle induit  $\mu_i$  concentré sur l'atome lourd  $i$  de la molécule lorsque présent dans le champ électrique  $\mathbf{E}$  de la charge ponctuelle est donné par:<sup>95</sup>

$$\mu_i = \alpha_i q_{\text{test}} \frac{\mathbf{r}_{i,\text{test}}}{|\mathbf{r}_{i,\text{test}}|^3} = \alpha_i \mathbf{E}_i \quad (3.8)$$

où  $\alpha_i$  est la polarisabilité concentrée sur l'atome lourd  $i$ ,  $q_{\text{test}}$  est la charge ponctuelle  $+1e$  et  $|\mathbf{r}_{i,\text{test}}|$  est la distance séparant l'atome lourd  $i$  de la charge ponctuelle. Il s'ensuit que

$$U_{\text{pol.}}^{(1)} = -\frac{1}{2} \sum_i \frac{q_{\text{test}}^2 \alpha_i}{|\mathbf{r}_{i,\text{test}}|^4} \quad (3.9)$$

où la polarisabilité atomique  $\alpha_i$  est une valeur moyennée dans l'espace à déterminer dans le contexte particulier de la molécule modèle où l'atome lourd  $i$  se trouve.

Afin d'évaluer  $\alpha_i$ , l'équation ?? est réaménagée comme suit:

$$\sum_i \frac{\alpha_i}{|\mathbf{r}_{i,\text{test}}|^4} = -\frac{2}{q_{\text{test}}^2} (U_{\text{tot.}} - U_{\text{mol.}} - U_{\text{mol.-test}}) \quad (3.10)$$

où  $U_{\text{tot.}}$ ,  $U_{\text{mol.}}$  et  $U_{\text{mol.-test}}$  ( $= q_{\text{test}} V(\mathbf{r}_{\text{test}})$ ) sont des quantités évaluées par des calculs *ab initio* au niveau HF avec une base 6-311G\*\*<sup>82, 83, 84</sup> pour les atomes C, N, O et H. En pratique, chaque molécule modèle est insérée dans une grille de 200 points dont les plus près se trouvent à un minimum de 3 Å d'un atome. Comme il a déjà été mentionné, plutôt que d'utiliser un ion lors de ces calculs, une charge ponctuelle ne nécessitant pas l'usage d'une base de fonctions a été préférée.

Cette stratégie est une option offerte explicitement dans le logiciel Gaussian90 et permet de calculer le potentiel électrostatique  $V(\mathbf{r}_{\text{test}})$  en un point voulu. L'avantage principal demeure qu'aucune erreur de superposition des bases ne doit être corrigée.

A chacun de ces points de grille sont successivement calculées  $U_{\text{tot.}}$  et  $U_{\text{mol.-test}}$ ;  $U_{\text{mol.}}$  a été préalablement calculée au tout début. L'équation 3.10 est l'équation qui a été lissée avec un minimum d'erreurs par le logiciel *Mathematica*<sup>105</sup> aux données obtenues à chaque point de grille et suite auquel on obtient la polarisabilité de chaque groupement chimique concentré dans l'atome lourd du groupement. Les résultats sont présentés au tableau VIII.

La polarisabilité de la molécule d'eau calculée pour cette thèse est inférieure par  $\approx 50\%$  à la donnée expérimentale. Evidemment, le niveau de calcul ainsi que la base de fonctions utilisées influencent fortement ce résultat. Le calcul de polarisabilité de la molécule d'eau a été effectué en présence d'un  $\text{Ca}^{2+}$  plutôt qu'une charge ponctuelle. Avec la base de Wachters<sup>87</sup> contractée par Gianolo et coll.<sup>88</sup> pour représenter le  $\text{Ca}^{2+}$ , les calculs des termes de l'équation 3.10 ont été recalculés et l'erreur de la superposition de la base a été corrigée. Une valeur de  $0.888 \text{ \AA}^3$ , supérieure à la valeur obtenue en présence d'une charge ponctuelle ( $0.785 \text{ \AA}^3$ ) a été obtenue.

Les valeurs calculées pour les molécules entières sont obtenues en additionnant les polarisabilités atomiques concentrées sur les atomes lourds. Cela ne constitue qu'une approximation. Malgré tout, on notera un accord qualitatif certain entre les valeurs calculées ici et les données expérimentales.

	thèse	exp. <sup>103</sup>	autres calculs	
H <sub>2</sub> O	0.785	1.49	1.12 <sup>101</sup>	1.44 <sup>102</sup>
CH <sub>3</sub> OH	2.384	3.32	3.04 <sup>101</sup>	3.34 <sup>102</sup>
CH <sub>3</sub>	1.927		2.24 <sup>106</sup>	
OH	0.457		0.465 <sup>101</sup>	0.637 <sup>104</sup>
CH <sub>3</sub> CH <sub>2</sub> OH	4.2	5.08	5.11 <sup>101</sup>	5.26 <sup>102</sup>
CH <sub>3</sub>	1.934		2.24 <sup>106</sup>	
CH <sub>2</sub>	1.867	1.84 <sup>57</sup>	1.842 <sup>106</sup>	
OH	0.399		0.465 <sup>101</sup>	0.637 <sup>104</sup>
CH <sub>3</sub> CONH <sub>2</sub>	4.142	5.67	5.72 <sup>101</sup>	5.71 <sup>102</sup>
CH <sub>3</sub>	2.748		2.24 <sup>106</sup>	
C	–		0.616 <sup>101</sup>	
O	0.688		0.434 <sup>101</sup>	0.569 <sup>104</sup>
NH <sub>2</sub>	0.706		1.759 <sup>106</sup>	1.738 <sup>104</sup>
CH <sub>3</sub> CONHCH <sub>3</sub>	6.199	7.82	7.91 <sup>101</sup>	7.48 <sup>103</sup>
CH <sub>3</sub>	2.536		2.24 <sup>106</sup>	
C	–		0.616 <sup>101</sup>	
O	0.788		0.434 <sup>101</sup>	0.569 <sup>104</sup>
NH	0.217		1.431 <sup>106</sup>	1.351 <sup>104</sup>
CH <sub>3</sub>	2.526		2.24 <sup>106</sup>	
CH <sub>3</sub> CH <sub>2</sub> COOH	5.781	6.88	7.00 <sup>103</sup>	
CH <sub>3</sub>	2.354		2.24 <sup>106</sup>	
CH <sub>2</sub>	1.842	1.84 <sup>57</sup>	1.842 <sup>106</sup>	
C	–		0.616 <sup>101</sup>	1.027 <sup>106</sup>
OH	0.900		0.637 <sup>104</sup>	
O	0.685		0.569 <sup>104</sup>	

TABLEAU VIII. Polarisabilités en Å<sup>3</sup> de molécules, groupements chimiques et atomes pour les molécules modèles.

Les polarisabilités des autres molécules entières semblent présenter un meilleur accord avec les données expérimentales et les autres méthodes de calcul. Par contre, les valeurs calculées pour les atomes et les groupements chimiques varient d'une méthode à l'autre. Les valeurs obtenues par d'autres méthodes de calculs pour les atomes et les groupements chimiques correspondent à des moyennes du même groupement chimique ou atome retrouvé dans plusieurs molécules différentes. Pour tous les cas, il a fallu poser la valeur de la polarisabilité du carbone lié à l'oxygène à zéro dans le processus de lissage afin de pouvoir déterminer les autres polarisabilités, sinon une valeur négative pour le carbone était obtenue.

### 3.5 De l'inclusion de l'énergie de polarisation induite par le $\text{Ca}^{2+}$ sur une molécule d'eau

Les options permettant d'inclure les énergies de polarisation au premier et second ordre étaient déjà disponibles dans CHARMM.<sup>56</sup> Elles ont été utilisées pour effectuer des tests sur le système du  $\text{Ca}^{2+}$  inclus dans 25 molécules d'eau rapportés dans cette section. Dans ce système, la charge du  $\text{Ca}^{2+}$  est fixée à  $+2e$  et seul l'atome d'oxygène de l' $\text{H}_2\text{O}$  est polarisable par le  $\text{Ca}^{2+}$ .

Reformulons le terme d'énergie de polarisation induite au premier ordre donné à l'équation 3.7:

$$\begin{aligned} U_{\text{pol.}}^{(1)} &= -\frac{1}{8\pi\epsilon_0} \sum_i \frac{q_j^2 \alpha_i}{|\mathbf{r}_{i,j}|^4} \\ &= -\frac{1}{2} \sum_i \mu_{\text{ind.}}^i \cdot \mathbf{E}_j^i \end{aligned} \quad (3.11)$$

où  $\alpha_i$  est la polarisabilité de l'oxygène  $i$  et  $\mu_{\text{ind.}}^i$  est le dipôle induit sur l'oxygène par le  $\text{Ca}^{2+}$ . Le terme au second ordre qui, rappelons-le, représente l'essentiel de la contribution non-additive, est donné par:

$$U_{\text{pol.}}^{(2)} = -\sum_i \sum_j \frac{\mu_{\text{ind.}}^i \cdot \mathbf{r}_{i,j} q_j}{4\pi\epsilon_0 |\mathbf{r}_{i,j}|^3} \quad (3.12)$$

$$= - \sum_i \mu_{\text{ind.}}^i \cdot \mathbf{E}_{\text{total}}^i.$$

Ce terme représente les interactions entre les charges partielles  $q_j$  et les dipôles induits  $\mu_{\text{ind.}}^i$ , séparés par une distance  $\mathbf{r}_{i,j}$ ;  $\mathbf{E}_{\text{total}}^i$  est le champ électrique total créé par toutes les charges partielles au point  $i$  où se trouve le dipôle.

L'énergie libre d'hydratation du  $\text{Ca}^{2+}$  dans ces nouvelles conditions a été évaluée à partir de la trajectoire produite à la section 2.3 avec les paramètres du tableau I. Il s'agit donc d'un calcul de perturbation au premier degré lors duquel on évalue la différence d'énergie libre associée à l'inclusion des termes d'énergies de polarisation au premier et second ordre à l'aide de l'équation suivante:

$$\Delta G_{(\text{sans pol.} \rightarrow \text{avec pol.})} = -k_B T \ln \langle \exp[-\beta(U_{\text{sans pol.}} - U_{\text{avec pol.}})] \rangle_{(\text{sans pol.})} \quad (3.13)$$

Les  $U_{(\text{sans pol.})}$  et  $U_{(\text{avec pol.})}$  sont les énergies configurationnelles calculées respectivement sans et avec les termes d'énergies de polarisation induite à partir des configurations de la trajectoire produite sans terme d'énergie de polarisation induite;  $\beta = 1/k_B T$  où  $k_B$  est la constante de Boltzmann et  $T$  la température.

Rappelons que l'énergie libre d'hydratation du  $\text{Ca}^{2+}$  mesurée expérimentalement est d'environ  $-380.8 \text{ kcal/mol}$ ; <sup>70</sup> l'inclusion du terme au premier ordre sans modification des paramètres donnerait, comme il est facile de le prévoir, une énergie libre encore plus négative ( $\approx -547.8 \text{ kcal/mol}$ ) à cause des interactions favorables entre le  $\text{Ca}^{2+}$  et les dipôles induits des molécules d'eau. Par contre, l'inclusion du terme au second ordre avec les mêmes paramètres apporte une contribution répulsive et donnerait une énergie libre d'hydratation aux environs de  $-454.3 \text{ kcal/mol}$ , puisque les dipôles induits sont parallèles aux dipôles permanents.

La courbe d'énergie d'interaction  $\text{Ca}^{2+}\text{-H}_2\text{O}$  recalculée elle aussi avec les termes au premier et second ordre est présentée à la figure 18. La polarisabilité de l'oxygène est celle déterminée à la section précédente, soit  $0.785 \text{ \AA}^3$ . La profondeur du puits est surestimée par *param22* résultant aussi en une distance minimale d'approche trop courte; et les interactions à courte distance sont décrites par

un mur de répulsion trop abrupt. Le fait que le dipôle permanent du modèle de molécule d'eau TIP3P<sup>60</sup> soit au départ surestimé par 27%, ainsi que le fait que le  $\text{Ca}^{2+}$  reste à  $+2e$  jouent probablement une grande part dans ces aspects à corriger. Les courbes d'énergie d'interaction entre le  $\text{Ca}^{2+}$  et les molécules modèles

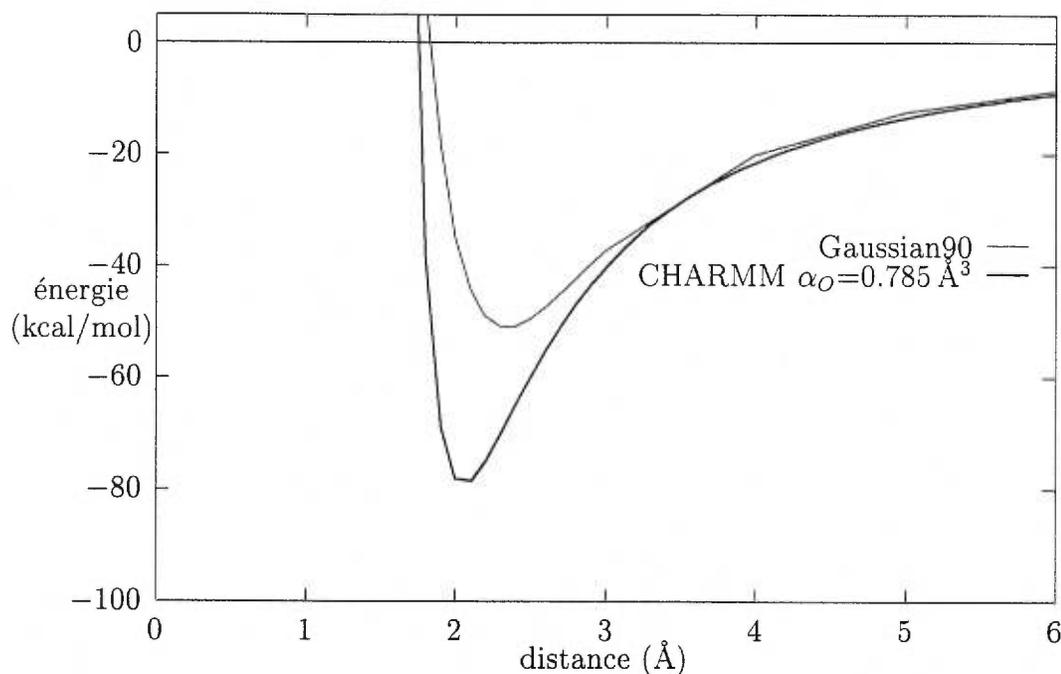


FIGURE 18. Courbes de l'énergie d'interaction  $\text{Ca}^{2+}\text{-H}_2\text{O}$  en fonction de la distance telle que définie à la figure 13 et incluant les termes d'énergies de polarisation au premier et second ordre.

ont aussi été testées de la même façon que pour l' $\text{H}_2\text{O}$  et les mêmes défauts du modèle ont été reconnus.

Bien que la méthode utilisée pour calculer la différence d'énergie libre associée à l'inclusion des termes d'énergies de polarisation induite comporte plusieurs

approximations, elle semble indiquer que des propriétés macroscopiques telle l'énergie libre d'hydratation ne peuvent être reproduites par un simple ajout de termes d'énergies de polarisation induite au premier et second ordre sans modifications soit des paramètres du  $\text{Ca}^{2+}$  et du modèle de molécule ou acide aminé, ou soit de la forme fonctionnelle du potentiel de Lennard-Jones. Il serait peut-être aussi nécessaire de décrire par de différents potentiels les interactions  $\text{Ca}^{2+}$ - $\text{H}_2\text{O}$  à courte distance et celles qui sont à moyennes et longues distances.<sup>94</sup> Une renormalisation des polarisabilités déterminées à la section précédente n'est pas non plus à exclure.

### 3.6 Conclusion

Dans ce chapitre, un ensemble de paramètres de Lennard-Jones pour le  $\text{Ca}^{2+}$  a été déterminé afin de reproduire l'énergie libre d'hydratation du  $\text{Ca}^{2+}$  et les propriétés statiques telles la structure immédiate des molécules d'eau autour du  $\text{Ca}^{2+}$  (fonction de distribution radiale) et la coordinence du  $\text{Ca}^{2+}$ . Cependant les courbes d'interaction *ab initio* ne sont pas reproduites de façon identique. Une raison à l'origine de cette différence est que les charges attribuées aux atomes des acides aminés du champ de force *param22* sont fixes et n'ont pas la possibilité de se réajuster afin de bien représenter des interactions avec des ions divalents et positifs. En effet, le  $\text{Ca}^{2+}$  étant un ion polarisant, il subit un transfert de charge considérable qui devrait être tenu en compte.

La contribution non-additive dans les interactions  $\text{Ca}^{2+}-(\text{H}_2\text{O})_n$  a été évaluée par des méthodes *ab initio* puis comparée avec les calculs homologues effectués avec *param22*. Bien que les paramètres du  $\text{Ca}^{2+}$  reproduisent son énergie libre d'hydratation, la fonction d'énergie potentielle totale néglige cependant la contribution non-additive reliée aux effets à plusieurs corps qui est considérable dans le cas du  $\text{Ca}^{2+}$ . Aussi, un transfert de charge plus important que dans l'interaction du  $\text{Ca}^{2+}$  avec un seul ligand survient.

Cette contribution non-additive joue un rôle dans les propriétés structurales et dynamiques des liquides.<sup>74, 99</sup> Sachant que cette contribution est bien représentée par l'énergie de polarisation induite au second ordre pour les interactions  $\text{Ca}^{2+}-\text{H}_2\text{O}$ <sup>94</sup> et  $\text{Ca}^{2+}$ -peptide,<sup>95</sup> des polarisabilités pour les atomes et groupements chimiques de molécules représentant les ligands du  $\text{Ca}^{2+}$  dans les sites de la calbindine  $\text{D}_{9k}$  ont été déterminées. Afin de déterminer ces polarisabilités, une méthode approximative et simple a été élaborée dans laquelle seuls les atomes lourds de molécules modèles d'acides aminés peuvent être polarisés par une charge ponctuelle. Les résultats obtenus s'accordent raisonnablement bien avec les données

provenant d'expériences ou d'autres calculs.

Ensuite, une tentative d'inclure les termes d'énergie de polarisation au premier et second ordre pour le cas du  $\text{Ca}^{2+}$  hydraté a été accomplie. Les paramètres déterminés à la première section de ce chapitre ont été utilisés. Dès le départ, cette tentative démontre la sensibilité des résultats aux paramètres et suggère fortement l'existence de difficultés inhérentes au développement d'une fonction de potentiel qui tiendrait compte des effets de polarisation induite. Devant l'ampleur du travail qu'impliquerait la reparamétrisation d'un champ de forces pouvant représenter la calbindine  $\text{D}_{9k}$ , qui compte 75 acides aminés, sans garantie de succès à court et moyen termes, et étant donné que des propriétés macroscopiques peuvent être reproduites avec les paramètres associés, il a été préféré d'utiliser tel quel le champ de force *param22* pour les protéines, d'approximer que le  $\text{Ca}^{2+}$  était une sphère de charge  $+2e$  et que ses ligands n'étaient pas polarisables. Et bien que l'ensemble de paramètres utilisés pour les calculs subséquents ne reproduisent pas parfaitement les courbes d'énergie d'interaction *ab initio*, il ne faut pas perdre de vue que ces paramètres devraient idéalement s'ajuster à des données qui représentent les interactions qui se passent à l'intérieur de la protéine totale hydratée.

## CHAPITRE 4

### Simulations de dynamique moléculaire de la calbindine $D_{9k}$ dans quatre de ses états d'occupation

#### 4.1 Introduction

Il y a une vingtaine d'années, avec les premières simulations de dynamique moléculaire (DM),<sup>53, 107</sup> il est apparu que la vision plutôt statique que l'on avait alors des protéines devait être révisée en faveur d'un modèle tenant compte de l'aspect dynamique.<sup>54</sup> Ainsi, la relation structure-fonction qui présidait à beaucoup de recherches sur les protéines a été reformulée en relation structure-dynamique-fonction. Ce changement de vision, initié et formulé clairement suite aux premières expériences de simulations de DM, est le point de départ de toute une avenue de recherches à la fois dans les domaines de la méthodologie et de la résolution de mécanismes d'actions des protéines.

Un des grands avantages des simulations de DM par rapport aux méthodes expérimentales de recherche est de pouvoir générer et disposer d'un ensemble de configurations qui se prête à une analyse statistique des propriétés structurales et dynamiques d'un système donné. Cependant, les résultats obtenus dépendent directement et fortement de la fonction de potentiel utilisée ainsi que des paramètres du modèle. Les résultats de simulations de DM doivent donc être analysés et interprétés de pair avec les données expérimentales homologues.

La calbindine  $D_{9k}$  offre un très bon exemple de l'importance de l'aspect

dynamique dans la fonction des protéines. Comme il a déjà été mentionné au premier chapitre, la liaison du  $\text{Ca}^{2+}$  à la calbindine  $\text{D}_{9k}$  ne provoque pas de changement conformationnel majeur de la chaîne principale mais engendre tout de même la réorganisation des chaînes latérales et de petites réorganisations subtiles dans l'orientation relative des hélices entre elles et par rapport aux sites.<sup>25, 33, 32</sup> Par contre, les changements de dynamique entraînés par la liaison de  $\text{Ca}^{2+}$  sont plus importants que les changements structuraux et la liaison d'un premier  $\text{Ca}^{2+}$  provoque de plus grands changements de dynamique que la liaison du second.<sup>24, 28, 33, 39</sup> Cela dit, il est fondamental de se demander comment la coopérativité dans la liaison du  $\text{Ca}^{2+}$  à la calbindine  $\text{D}_{9k}$  peut-elle se produire avec seulement de petits changements structuraux? Et quel est le rôle fonctionnel associé aux différents types de fluctuations qui ont lieu dans les divers états d'occupation?

Dès lors, la DM apparaît comme un instrument puissant de recherche car elle permet de suivre le déplacement de chaque atome dans le temps et donc d'analyser, dans l'ombre bien sûr du modèle utilisé, la structure et la dynamique d'un système biomoléculaire complexe à un niveau de résolution difficilement atteignable par des méthodes expérimentales. Des simulations de DM ont déjà été effectuées sur la calbindine  $\text{D}_{9k}$ ,<sup>108, 109, 110</sup> ainsi que sur la parvalbumine<sup>71</sup> qui est la plus proche voisine de la calbindine  $\text{D}_{9k}$  avec 108 acides aminés, deux sites de liaison pour le  $\text{Ca}^{2+}$  et six hélices  $\alpha$ .<sup>4</sup>

Une première simulation de la parvalbumine a été effectuée par Ahlström et coll. en 1987<sup>71</sup> avec le logiciel MUMOD.<sup>111</sup> La structure cristallographique de la parvalbumine liée à deux  $\text{Ca}^{2+}$  a été utilisée comme point de départ.<sup>4</sup> Lors de la simulation, tous les atomes de la parvalbumine ont été décrits explicitement, sauf les groupements  $\text{CH}$ ,  $\text{CH}_2$  et  $\text{CH}_3$  qui ont été traités comme une seule particule ou pseudoatome. Une fonction de potentiel standard a servi à simuler trois systèmes différents: 1) la parvalbumine liée à deux  $\text{Ca}^{2+}$  dans le vide; 2) la parvalbumine

liée à deux  $\text{Ca}^{2+}$  et incluse dans une boîte périodique contenant 2327 molécules d'eau et trois ions  $\text{Na}^+$ ; et 3) la parvalbumine sans  $\text{Ca}^{2+}$  dans le vide. Toutes les trois simulations ont été effectuées à 300 K et les interactions non-liées ont été coupées à 9 Å. Des trajectoires de 80 ps (systèmes 1 et 3) et 70 ps (système 2) ont été récoltées et analysées.

Les deux simulations dans le vide amènent la structure à différer considérablement de la structure cristallographique initiale, tandis que la simulation dans l'eau engendre des changements beaucoup moins drastiques. La coordinence du  $\text{Ca}^{2+}$  est affectée dans les deux sites, que ce soit pour la simulation dans le vide ou dans l'eau. Alors que les  $\text{Ca}^{2+}$  dans les sites 1 et 2 possèdent respectivement six et huit ligands dans la structure cristallographique, les simulations dans le vide et dans l'eau donnent huit ligands pour chaque  $\text{Ca}^{2+}$  et les ligands ne sont pas tous les mêmes que dans la structure cristallographique. La présence de l'eau dans la simulation affecte considérablement la dynamique de la parvalbumine, et ce autant pour les acides aminés à la surface qu'à l'intérieur de la protéine.

La calbindine  $\text{D}_{9k}$  liée à deux  $\text{Ca}^{2+}$  a été simulée pour la première fois en 1989 par Ahlström et coll.<sup>108</sup> à partir de la structure cristallographique de Szebenyi et coll.<sup>112</sup> De même que pour la parvalbumine, le logiciel MUMOD<sup>111</sup> a été utilisé et les groupements  $\text{CH}$ ,  $\text{CH}_2$  et  $\text{CH}_3$  ont été traités comme des pseudoatomes. La calbindine  $\text{D}_{9k}$ , ainsi que les 36 molécules d'eau faisant partie de la structure cristallographique ont été immergées dans une boîte périodique de 2 212 molécules d'eau supplémentaires. Les interactions non-liées ont été coupées à 10 Å et la simulation a été effectuée à 300 K pendant 124 ps au total.

L'analyse de cette trajectoire a révélé que la protéine s'est légèrement contractée par rapport à la structure cristallographique de départ. Aussi, le  $\text{Ca}^{2+}$  dans le site I a perdu un atome d'oxygène d'un groupement carbonyle (de Ala14) et a acquis en revanche deux atomes d'oxygène de groupements carboxyles (de Glu60), pour une coordinence totale de huit ligands, soit deux de plus que dans la structure

cristallographique. Le site II a acquis un atome d'oxygène d'un groupement carboxyl (Asp54 passe de *mono* à *bidenté*) et il y a eu échange de la molécule d'eau présente initialement dans la structure cristallographique pour une autre provenant des 2212 rajoutées. Le fait notable de cette simulation est que les deux atomes d'oxygène de l'acide aminé Glu60 (site II), se sont étirés vers le site I où ils se sont liés au  $\text{Ca}^{2+}$  pendant presque toute la durée de la simulation. Cet acide aminé n'est nullement lié au  $\text{Ca}^{2+}$  du site I dans les données cristallographiques. Une coordinence de huit est aussi observée dans ce cas, soit un ligand de plus que dans la structure cristallographique.

Suite à cette simulation, Kördel et Teleman<sup>109</sup> ont procédé à une série de trois améliorations du modèle U (nomenclature donnée par ces auteurs) utilisé par Ahlström et coll.<sup>108</sup> Une de ces améliorations consiste à représenter tous les atomes explicitement (modèle A); deux autres consistent à ajouter un terme d'énergie afin de maintenir les ponts hydrogènes, et de redistribuer les vitesses en fonction du type de mouvements accomplis par l'atome (modèle N). Encore une fois, la structure cristallographique de l'état D de Szebenyi et coll.<sup>112</sup> et le logiciel MUMOD<sup>111</sup> ont été utilisés. Les autres paramètres des simulations n'étaient pas significativement différents de ceux du modèle U de Ahlström et coll.<sup>108</sup>

Une analyse comparée de l'aspect dynamique des modèles U, A et N a révélé qu'au mieux, la flexibilité de la calbindine  $\text{D}_{9k}$  était surestimée par 10%. Le modèle d'atomes explicites (A) se comporte nettement mieux que les modèles U et N, et le modèle N, bien qu'il surestime d'autant plus la flexibilité de la calbindine  $\text{D}_{9k}$ , reproduit qualitativement mieux la flexibilité relative des différentes sous-unités de la protéine. De plus, on suggère d'effectuer des simulations plus longues ou à tout le moins de ne commencer à échantillonner que lorsque le système est équilibré, ainsi que d'augmenter le rayon de coupure des interactions non-liées.

La dernière simulation à être rapportée est celle de Mehler et coll. (1993)<sup>110</sup> où encore une fois la structure cristallographique de l'état D de Szebenyi et

coll.<sup>112</sup> est utilisée comme point de départ. Tous les atomes sont représentés explicitement lors des calculs. Cette fois, le logiciel CHARMM<sup>56</sup> et le champ de force *param19* ont servi à simuler la calbindine D<sub>9k</sub> qui est testée dans deux modèles d'hydratation différents: 1) les 36 molécules d'eau présentes dans la structure cristallographique sont enlevées et on place ensuite 1 191 nouvelles molécules d'eau formant une coquille de 9 Å autour de la calbindine D<sub>9k</sub>. Rien n'est spécifié quant aux conditions appliquées à la frontière du solvant; 2) les 36 molécules d'eau présentes dans la structure cristallographique sont conservées et le solvant (eau) autour est modélisé par un milieu diélectrique qui dépend linéairement de la distance, c'est-à-dire que les interactions électrostatiques sont divisées par une constante diélectrique de plus en plus grande à mesure que l'on s'éloigne du centre du système. Les systèmes ont été simulés à 300 K et les interactions non-liées éteintes à 8 Å. Tous les états d'occupation ont été simulés (A, S<sub>I</sub>(D), S<sub>II</sub> et D) à partir de la structure cristallographique de l'état D, en effaçant successivement et dans divers ordres les Ca<sup>2+</sup> présents dans les deux sites. Des trajectoires allant de 200 ps jusqu'à 600 ps ont été récoltées et analysées.

En général, les simulations où la calbindine D<sub>9k</sub> se trouve dans un milieu diélectrique uniforme éprouvent les mêmes difficultés que les simulations de protéines dans le vide, c'est-à-dire qu'elles ne reproduisent pas correctement les données expérimentales de structure et de dynamique. Par contre, les simulations effectuées dans la sphère de 1 191 molécules d'eau reproduisent mieux les mêmes données expérimentales, bien que des différences notables apparaissent, entre autres en ce qui concerne la coordinence du Ca<sup>2+</sup>. Dans la simulation avec le solvant explicite, la coordinence du Ca<sup>2+</sup> demeure de sept dans chacun des sites mais la nature des ligands change. Le Ca<sup>2+</sup> dans le site I échange un ligand (Glu27) pour un atome d'oxygène provenant d'un acide aminé du site II (Glu60), tout comme lors de la simulation de Ahlström et coll.,<sup>108</sup> où par contre les deux atomes d'oxygène du groupement carboxylique venaient lier le Ca<sup>2+</sup> du site I. Le Ca<sup>2+</sup> dans le site II du système simulé demeure lié aux mêmes atomes que dans la structure cristal-

lographique de départ, mais les distances  $\text{Ca}^{2+}$ -ligand sont modifiées, et sont en général plus grandes. Malgré ces différences, un accord qualitatif est obtenu avec les données expérimentales et il est vérifié que le retrait du  $\text{Ca}^{2+}$  des sites de la calbindine  $\text{D}_{9k}$  affecte peu sa structure, mais affecte davantage sa dynamique.

En résumé, les grandes tendances observées expérimentalement quant aux propriétés structurales et dynamiques sont qualitativement reproduites par les simulations effectuées en représentant explicitement tous les atomes de la protéine et du solvant, alors que les modèles de pseudoatomes ou de solvant modélisé par un diélectrique continu apparaissent inadéquats. Même avec la présence de molécules d'eau explicites autour de la protéine, des problèmes tels que la détérioration de la structure secondaire (réseau de ponts hydrogènes) ainsi que des changements de coordinence du  $\text{Ca}^{2+}$  peuvent survenir. De plus, la flexibilité de la protéine est dans la plupart des cas surestimée. Il faut rappeler que les sites de la calbindine  $\text{D}_{9k}$  sont distants d'environ  $12 \text{ \AA}$ , avec ou sans  $\text{Ca}^{2+}$ , et que couper les interactions non-liées, que ce soit à  $8$  ou  $10 \text{ \AA}$ , est un compromis des moins souhaitables étant donné que les interactions électrostatiques ont une longue portée (par exemple, l'interaction entre deux  $\text{Ca}^{2+}$  dans le vide distants par  $12 \text{ \AA}$  est de plus de  $110 \text{ kcal/mol}$ ).

Le but de ce chapitre est donc d'effectuer une étude des propriétés structurales et dynamiques des quatre états d'occupation (A,  $\text{S}_I(\text{D})$ ,  $\text{S}_{II}$ , D) de la calbindine  $\text{D}_{9k}$  par des méthodes de simulations de dynamique moléculaire. Contrairement aux autres simulations effectuées sur la calbindine  $\text{D}_{9k}$ , il a été possible de bénéficier de structures expérimentales de départ pour tous les états, sauf pour l'état  $\text{S}_I(\text{D})$ . De plus, aucun rayon de coupure n'a été appliqué aux interactions non-liées et un modèle d'hydratation qui représente à la fois les effets directs des molécules d'eau ainsi que les effets d'un solvant infini a été utilisé. Les résultats sont comparés avec les données expérimentales lorsque disponibles et il est discuté de quelques éléments pouvant jouer un rôle clef dans la coopérativité.

## 4.2 Méthode

### 4.2.1 Préparation des structures de départ

La provenance des structures expérimentales de départ a été donnée à la section 2.2.3 (point 1). La structure A du mutant P43G provenant d'expériences de RMN a été utilisée telle quelle sans aucun changement avant d'être soumise au protocole d'équilibration et de production.

La structure  $S_{II}$  du mutant P43G, provenant elle aussi d'expériences de RMN, ne contenait pas de coordonnées pour l'ion, bien que la structure ait été déterminée en présence d'un ion  $Cd^{2+}$  dans le site II. Afin d'attribuer des coordonnées au  $Ca^{2+}$ , la structure du site II de l'état  $S_{II}$  a été superposée à la structure du site II de l'état D, déterminée par diffraction des rayons X,<sup>15</sup> de sorte à minimiser la différence de distance entre les positions atomiques des deux structures (voir équation 4.1), en excluant les atomes d'hydrogène. Une valeur minimale de 1.17 Å a été obtenue. Puis, les coordonnées de l'ion  $Ca^{2+}$  de l'état D ont été copiées dans la structure de l'état  $S_{II}$  de même que celles de la molécule d'eau liée au  $Ca^{2+}$  dans le site II de l'état D. Enfin, une minimisation de l'énergie ( $\approx 200$  pas) de tous les atomes compris dans un rayon de 5 Å du centre du  $Ca^{2+}$  a été effectuée, avec une contrainte harmonique (constante de force = 1 kcal/(mol. Å<sup>2</sup>)) imposée aux atomes d'oxygène liés au  $Ca^{2+}$ .

La structure de la protéine sauvage de l'état D déterminée par diffraction des rayons X à une résolution de 1.6 Å<sup>15</sup> a été utilisée plutôt que la structure déterminée par RMN parce que cette dernière ne fournissait pas les coordonnées des  $Ca^{2+}$  dans les sites. De plus, la structure cristallographique avait l'avantage de fournir les coordonnées de 57 molécules d'eau attachées à la protéine. Afin de pouvoir comparer cet état aux états A et  $S_{II}$ , il a fallu remplacer la proline en position 43 par une glycine afin d'obtenir le même mutant (P43G). La structure

cristallographique de l'état D a tout d'abord été superposée à la structure du mutant P43G de l'état D déterminée par RMN. Puis, la différence entre les deux structures (sauf les atomes d'hydrogène, voir équation 4.1) a été minimisée (valeur finale = 1.91 Å). Ensuite, les coordonnées des atomes C, C $\alpha$ , N et O de la Gly43 du mutant P43G (RMN) ont été transférées sur la structure cristallographique afin de remplacer les coordonnées de la Pro43. Enfin, une minimisation de l'énergie ( $\approx 200$  pas) de la Gly43 et de ses deux acides aminés adjacents a été effectuée tout en maintenant fixe le reste de la protéine.

Comme aucune structure expérimentale de l'état S<sub>I</sub>(D) n'était disponible, une structure a été préparée à partir de la structure cristallographique de l'état D déjà insérée dans une sphère de 2 285 molécules d'eau, d'où le Ca<sup>2+</sup> du site II a été effacé. Puis, avec une contrainte harmonique (constante de force = 1 kcal/(mol. Å<sup>2</sup>)) appliquée à tous les atomes, la structure a été soumise à 200 pas de minimisation suivis par 5 ps d'équilibration. Ce cycle a été répété trois fois. La contrainte a ensuite été enlevée et le cycle a été répété une dernière fois. La structure qui a émergé de ce processus a été utilisée comme structure de départ pour les périodes d'équilibration et de production de la trajectoire.

De plus, un autre état A (appelé A(S<sub>II</sub>) dans le reste du texte) a été construit rétrospectivement à l'analyse des résultats de l'état A. Cette structure A(S<sub>II</sub>) a été préparée à partir de la structure équilibrée de l'état S<sub>II</sub> de laquelle on a effacé le Ca<sup>2+</sup> pour ensuite répéter le protocole d'équilibration.

Chacune des structures des états A, S<sub>II</sub>, et D a ensuite été placée dans une sphère d'environ 27 Å contenant 2 285 molécules d'eau avant que ne s'amorce les étapes d'équilibration et de production de trajectoires. Un cliché de la structure hydratée de l'état D est présentée à la figure 19.

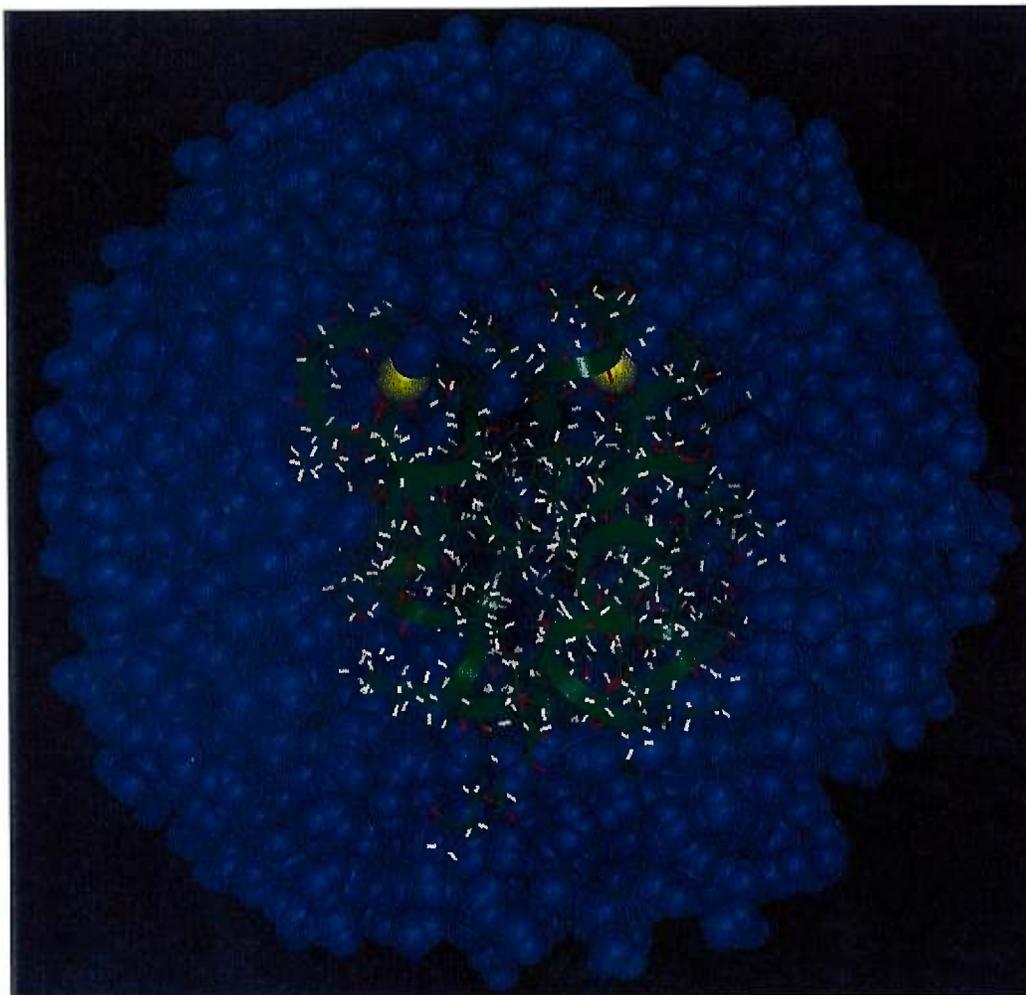


FIGURE 19. Structure de l'état D de la calbindine D<sub>9k</sub> entourée par 2 285 molécules d'eau.

## 4.2.2 Protocoles d'équilibration et de production des trajectoires

- Équilibration

Tous les systèmes ont été simulés à  $\approx 300$  K en effectuant de la dynamique de Langevin sur toutes les molécules d'eau auxquelles un coefficient de friction  $\gamma_i/m_i = 0.05$  ps a été appliqué seulement aux oxygènes, ce qui correspond à un temps caractéristique de relaxation de  $20$  ps<sup>-1</sup>. Un pas d'intégration de  $0.002$  ps a été choisi, ainsi que l'algorithme SHAKE<sup>65</sup> qui maintient fixes tous les liens chimiques avec des atomes d'hydrogène. Une contrainte harmonique a été appliquée sur le centre de masse de la protéine de chacun des systèmes, avec une force de  $10$  kcal/(mol. Å<sup>2</sup>) (états A et D) et de  $10\ 000$  kcal/(mol. Å<sup>2</sup>) (états S<sub>I</sub>(D) et S<sub>II</sub>). Une force beaucoup plus grande a été imposée au centre de masse des structures des états S<sub>I</sub>(D) et S<sub>II</sub> parce que ces dernières se déplaçaient de façon assez marquée vers la frontière de la sphère, c'est-à-dire vers le diélectrique continu ( $\epsilon = 80$ ). Il est possible qu'une inhomogénéité momentanée de la densité des molécules d'eau ait suffi à provoquer la dérive de la protéine vers la frontière.

Avant de débiter l'équilibration proprement dite, une minimisation de l'énergie des  $2\ 285$  molécules H<sub>2</sub>O a été effectuée ( $500$  pas) en maintenant fixes la protéine, les Ca<sup>2+</sup> (états S<sub>I</sub>(D), S<sub>II</sub>, et D) ainsi que les  $57$  molécules d'eau de la structure cristallographique de l'état D. L'équilibration a débuté avec  $5$  ps de simulation de DM, toujours en fixant la protéine, les Ca<sup>2+</sup> et les  $57$  molécules d'eau cristallographiques de l'état D. Pour l'état D,  $5$  ps supplémentaires ont été effectuées après avoir relâché la contrainte sur les  $57$  molécules d'eau cristallographiques. Puis, la contrainte fixant la protéine et les Ca<sup>2+</sup> a été changée en contrainte harmonique (constante de force =  $1$  kcal/(mol. Å<sup>2</sup>)) et  $100$  pas de minimisation de l'énergie ont été suivis par  $25$  ps de simulation de DM. La constante de force a été diminuée à  $0.5$  kcal/(mol. Å<sup>2</sup>) et la même séquence de minimisation et d'équilibration a été répétée. Finalement, toutes les contraintes ont été relâchées et encore

une fois, 100 pas de minimisation de l'énergie suivis par 25 ps de simulation de DM ont cette fois complété la période d'équilibration.

La structure A(S<sub>II</sub>) a subi un protocole un peu différent: une contrainte harmonique (constante de force = 1 kcal/(mol. Å<sup>2</sup>)) a d'abord été appliquée à tous les atomes et une minimisation de l'énergie (200 pas) suivi par 5 ps de simulation de DM ont été effectués. Cette séquence de minimisation et d'équilibration a été répétée trois fois, puis une période d'équilibration de 75 ps sans contrainte a mis fin au processus d'équilibration.

- Production des trajectoires

Une trajectoire de 200 ps a été produite pour chacune des structures équilibrées des états A, A(S<sub>II</sub>), S<sub>I</sub>(D), S<sub>II</sub> et D. De plus, une période supplémentaire de production de 200 ps a été cumulée pour l'état D afin d'évaluer la convergence des résultats.

## 4.3 Résultats

### 4.3.1 Analyse comparée des propriétés structurales à l'équilibre

#### 4.3.1.1 Évolution temporelle de la différence structurale entre les structures expérimentales de départ et les structures simulées

L'évolution temporelle de la différence entre les structures simulées et les structures expérimentales de départ est présentée à la figure 20 afin de valider les ensembles de configurations générés et de justifier le calcul de propriétés moyennes à l'équilibre. La différence structurale entre deux structures est ici évaluée par:

$$D_{\text{struc.}} = \sqrt{\frac{\sum_i m_i |\mathbf{r}_i^{\text{DM}} - \mathbf{r}_i^{\text{exp.}}|^2}{\sum_i m_i}}, \quad (4.1)$$

où  $\mathbf{r}_i^{\text{DM}}$  et  $\mathbf{r}_i^{\text{exp.}}$  sont la position de l'atome  $i$  respectivement dans la structure simulée et dans la structure expérimentale par rapport à une origine commune.

Chaque différence de distance est pondérée par la masse de l'atome concerné. A chaque 5 ps de simulation, un cliché a été enregistré et comparé à la structure expérimentale de référence *via* l'équation 4.1.

La figure 20 montre que l'état D est celui dont la structure simulée demeure la plus proche de la structure expérimentale de départ (déterminée par diffraction des rayons X), suivi de près par l'état S<sub>II</sub> (déterminée par RMN); les simulations de l'état A sont celles qui dévient le plus de la structure expérimentale (déterminée par RMN).

Pour tous les atomes lourds de la structure D simulée, la différence structurale demeure stable autour de 1.8 Å tout au long de la simulation, après une légère augmentation de 0.84 Å qui se produit dans les premières 100 ps. Pour le cas où seule la chaîne principale est considérée, on note aussi une légère augmentation de 0.6 Å lors des premières 100 ps, et la différence se stabilise à 1.41 Å.

La structure simulée S<sub>II</sub> présente aussi une augmentation de la différence structurale lors des premières 75 ps ( $\approx 1.17$  Å pour tous les atomes et  $\approx 0.95$  Å pour la chaîne principale) qui se stabilise autour de 2.43 Å et 1.9 Å respectivement pour tous les atomes lourds et ceux de la chaîne principale seulement. La structure expérimentale de départ a été choisi arbitrairement parmi l'ensemble de 33 structures S<sub>II</sub> déterminées, dont l'écart type de la structure moyenne est de 1.25 Å pour tous les atomes lourds, et de 0.86 Å pour la chaîne principale.<sup>33</sup> Ces écarts types, qui peuvent être interprétés comme une mesure de précision de la structure, suggèrent que la structure simulée est demeurée raisonnablement proche de la structure expérimentale.

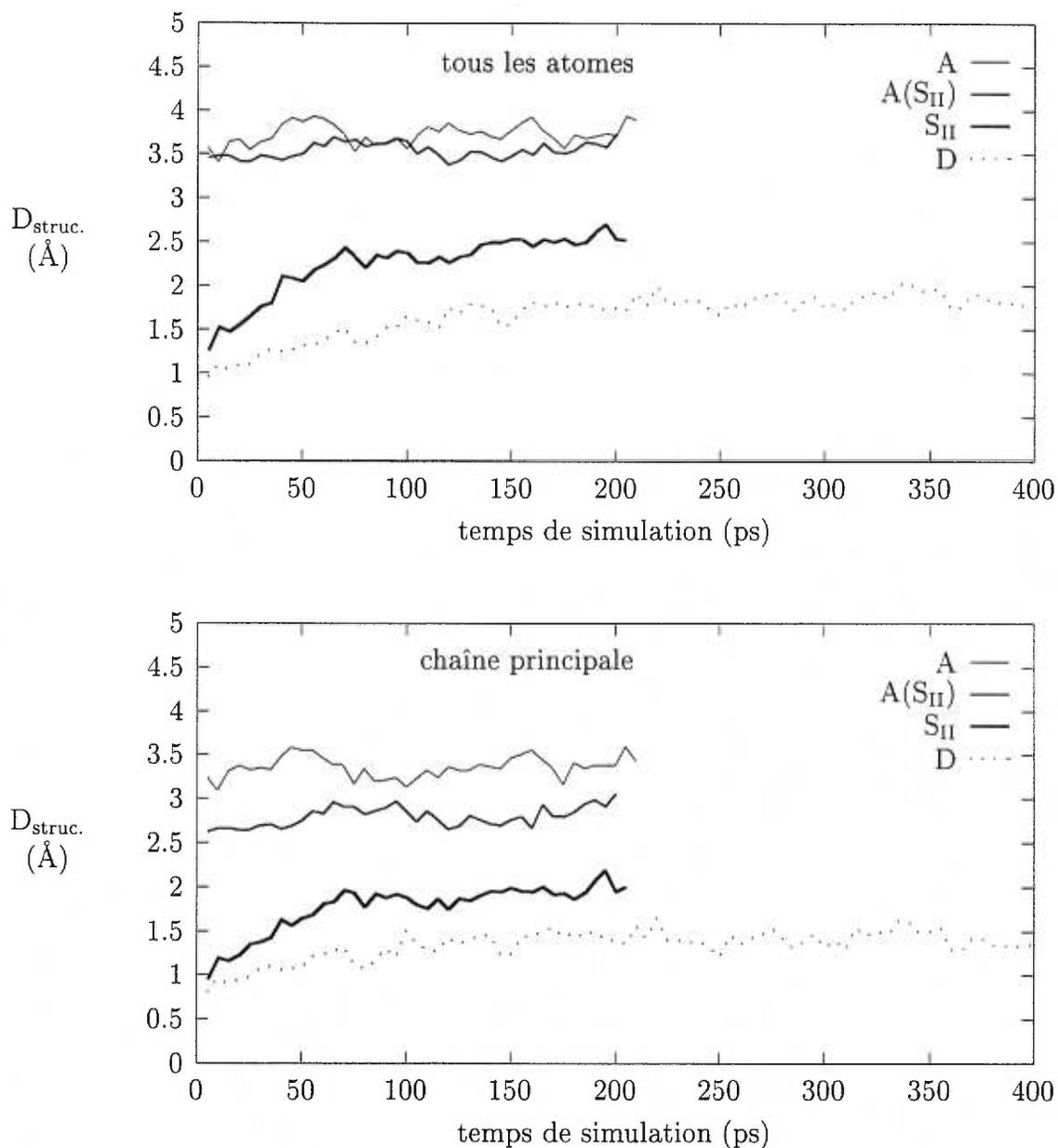


FIGURE 20. Différences entre les structures expérimentales et les structures simulées ( $D_{\text{struc.}}$  en fonction du temps de simulation (production de la trajectoire) en considérant tous les atomes lourds de la calbindine  $D_{9k}$  (haut) ou tous les atomes lourds de la chaîne principale, en excluant le  $\text{Ca}^{2+}$  dans les deux cas. Les  $D_{\text{struc.}}$  sont évaluées selon l'équation 4.1.

Les structures simulées A et A(S<sub>II</sub>) sont toutes deux comparées à la même structure RMN de l'état A de départ. La figure 20 montre que la structure simulée A(S<sub>II</sub>), provenant de la structure RMN de l'état S<sub>II</sub>, affiche moins de différences que la structure simulée à partir de cette structure RMN de l'état A. Les différences par rapport à la structure RMN de l'état A s'établissent à 3.73 Å et 3.36 Å respectivement pour les simulations A et A(S<sub>II</sub>) en considérant tous les atomes lourds. Si l'on ne considère que les atomes de la chaîne principale, on trouve aussi que la simulation A(S<sub>II</sub>) demeure plus près de la structure RMN ( $\approx 2.80$  Å) que la simulation A ( $\approx 3.36$  Å). Il ressort donc, étonnamment, que la simulation de l'état A faite à partir de la structure expérimentale de l'état S<sub>II</sub> de laquelle on a retiré l'ion, se comporte mieux que la simulation faite à partir de la structure expérimentale d'origine. La structure expérimentale a ici aussi été choisie parmi un ensemble de 33 structures dont l'écart type de la moyenne est, pour tous les atomes lourds, de 1.59 Å, tandis que l'écart type de la moyenne pour la chaîne principale seulement est de 1.17 Å.<sup>32</sup> Bien que plus élevés que pour la simulation S<sub>II</sub>, ces chiffres indiquent que les structures simulées de l'état A demeurent stables et que les trajectoires forment des ensembles de configurations à l'équilibre.

Puisque chaque structure simulée est demeurée raisonnablement proche de son homologue expérimental et que l'ensemble des configurations est représentatif d'un système à l'équilibre, des calculs de propriétés moyennes ont été effectués tels que décrits dans les prochaines sous-sections, ainsi que des études comparatives des différents résultats.

#### 4.3.1.2 Différences structurales entre la structure expérimentale de départ et la structure simulée, et des structures simulées entre elles

Pour chaque état d'occupation, les différences structurales entre la structure expérimentale et chacune des 20 000 configurations simulées ont été calculées selon l'équation 4.1 et les résultats sont pondérés par la masse de chaque atome. Les différences structurales moyennées sur les 20000 configurations ont été calculées pour chaque élément de structure secondaire, tels que définis dans les données expérimentales homologues, afin de pouvoir établir des comparaisons. Toutes les comparaisons excluent les atomes d'hydrogène et les  $\text{Ca}^{2+}$ , et les résultats sont présentés au tableau IX.

Les deux simulations de l'état A sont celles qui diffèrent le plus de la structure expérimentale correspondante, la simulation A( $\text{S}_{\text{II}}$ ) demeurant tout de même globalement plus près que la simulation A. Si l'on ne considère que la chaîne principale de la simulation A( $\text{S}_{\text{II}}$ ), on obtient des éléments de structure secondaire qui demeurent beaucoup plus près de la structure expérimentale, sauf pour le site I et le segment 36-45. Dans la simulation A, le site II est l'élément de structure secondaire le plus différent de la structure expérimentale, suivi par le segment 36-45. Ces deux régions sont celles qui sont les moins bien déterminées expérimentalement, en particulier pour les acides aminés 1, 41-43 (segment 36-45), 56-57 (site II) et 75.<sup>32</sup> Une comparaison visuelle entre la structure simulée A et la structure expérimentale a révélé que la petite boucle formée par le segment 36-45 s'est légèrement déformée durant la simulation, ainsi que l'hélice C (acides aminés 46-53) qui présente un relâchement. De plus, l'orientation relative de l'hélice C par rapport aux autres éléments de structure secondaire a été modifiée, entraînant comme conséquences la distortion du site II et une distance plus grande entre les deux sites. Des résultats expérimentaux sur la structure A, on note que l'orientation relative de l'hélice C par rapport à l'hélice D (acides aminés 63-73)

		acides aminés							
		1-75	3-15	16-24	25-35	36-45	46-53	54-62	63-73
			hélice A	site I	hélice B	segment	hélice C	site II	hélice D
A	tout	3.76	2.02	1.93	1.52	3.33	2.91	4.26	1.54
	c.p.	3.31	1.06	1.07	0.48	2.69	1.67	3.65	0.60
	C $_{\alpha}$	3.30	0.89	0.92	0.48	2.71	1.61	3.65	0.60
A(S <sub>II</sub> )	tout	3.53	1.52	2.85	1.68	5.12	1.71	3.45	1.80
	c.p.	2.78	0.49	1.54	0.46	4.16	0.50	2.55	0.59
	C $_{\alpha}$	2.79	0.42	1.37	0.47	4.29	0.48	2.48	0.33
S <sub>II</sub>	tout	2.27	1.35	1.61	0.99	1.38	1.54	1.60	1.36
	c.p.	1.75	0.82	1.02	0.38	0.84	0.57	1.06	0.89
	C $_{\alpha}$	1.72	0.74	1.12	0.36	0.82	0.53	0.81	0.87
D	tout	1.52	1.17	0.95	0.76	1.22	1.04	0.77	1.37
	c.p.	1.21	0.59	0.58	0.33	1.05	0.39	0.66	0.76
	C $_{\alpha}$	1.23	0.56	0.48	0.28	1.11	0.35	0.49	0.77

TABLEAU IX. Différences structurales moyennes ( $\text{\AA}$ ) entre la structure expérimentale de départ et les structures simulées, calculées pour chaque élément de structure secondaire. L'abréviation c.p. signifie chaîne principale; tous les résultats excluent les atomes d'hydrogène et les Ca $^{2+}$ .

n'est pas bien déterminée.<sup>32</sup>

La structure simulée  $S_{II}$  demeure assez près de la structure expérimentale déterminée par RMN, c'est-à-dire qu'elle excède au plus par 1.02 Å l'écart type expérimental. Pour sa part, la structure simulée D est celle qui demeure la plus près de la structure expérimentale de départ, déterminée par diffraction de rayons X, et offre donc un modèle particulièrement bon pour les analyses subséquentes. Après avoir évalué les différences structurales engendrées par le modèle de DM et vérifié que ces différences se situaient à l'intérieur de l'écart type moyen de l'ensemble des structures expérimentales ou d'un intervalle raisonnable, il est alors permis de calculer les différences structurales entre les quatre états d'occupation simulés, toujours selon l'équation 4.1 et en pondérant les résultats par la masse de chaque atome et en excluant les atomes d'hydrogène.

Le nombre de configurations pour chaque état (20 000) nous contraint à calculer des différences entre structures moyennes, plutôt que de calculer les différences moyennes, tout comme au tableau IX. Le calcul de différences entre structures moyennes donnera une valeur inférieure au calcul de différences moyennes, c'est-à-dire:

$$\| \langle \mathbf{r}_1 \rangle - \langle \mathbf{r}_2 \rangle \| \leq \| \langle \mathbf{r}_1 - \mathbf{r}_2 \rangle \| \quad (4.2)$$

qui consiste en l'inégalité du triangle.<sup>113</sup>

Une structure moyenne a donc été calculée pour chaque trajectoire comptant 20 000 configurations et les différences structurales présentées au tableau X ont été calculées entre ces structures moyennes. L'analyse qui suit, basée sur des comparaisons de structures moyennes, ne peut inclure de commentaires approfondis sur les chaînes latérales puisque ces dernières, surtout à la surface de la protéine, fluctuent considérablement.

Les comparaisons impliquant la structure simulée A affichent les plus grandes différences des autres structures simulées. Toutefois, il faut se rappeler que la

		acides aminés							
		1-75	3-15	16-24	25-35	36-45	46-53	54-62	63-73
			hélice A	site I	hélice B	segment	hélice C	site II	hélice D
A-A(S <sub>II</sub> )	tout	3.81	1.37	2.01	1.53	4.21	2.04	4.24	1.42
	c.p.	3.39	0.86	1.63	0.25	3.49	1.38	2.78	0.67
A-S <sub>I</sub>	tout	4.19	1.79	1.83	1.69	3.55	2.49	3.90	2.29
	c.p.	3.76	1.19	0.72	0.26	2.86	1.65	3.02	1.40
A(S <sub>II</sub> )-S <sub>I</sub>	tout	2.78	1.19	1.84	0.96	2.03	1.23	2.77	1.50
	c.p.	2.18	0.62	1.65	0.33	1.57	0.51	1.72	1.05
A-S <sub>II</sub>	tout	3.34	1.39	1.97	1.37	2.90	1.85	3.23	1.66
	c.p.	2.86	0.73	1.55	0.20	2.39	1.34	2.62	0.88
A(S <sub>II</sub> )-S <sub>II</sub>	tout	2.57	0.98	1.48	0.83	2.56	0.82	3.33	1.07
	c.p.	2.23	0.49	0.93	0.22	1.81	0.37	2.07	0.63
A-D	tout	3.75	1.59	1.93	1.66	3.56	2.47	3.61	2.09
	c.p.	3.36	1.15	0.73	0.23	3.11	1.61	2.62	1.37
A(S <sub>II</sub> )-D	tout	2.64	1.12	2.03	1.21	1.62	1.26	3.24	1.50
	c.p.	2.12	0.63	1.64	0.29	1.15	0.41	2.03	1.05
S <sub>I</sub> -S <sub>II</sub>	tout	2.82	1.65	1.75	0.81	2.18	1.35	2.42	1.64
	c.p.	2.38	0.89	1.50	0.23	1.59	0.51	1.30	0.89
S <sub>I</sub> -D	tout	1.36	0.58	1.03	0.58	1.13	0.65	2.13	0.86
	c.p.	1.14	0.28	0.34	0.15	0.85	0.31	1.40	0.26
S <sub>II</sub> -D	tout	2.32	1.37	1.62	0.95	1.98	1.35	1.39	1.43
	c.p.	1.83	0.80	1.39	0.24	1.42	0.40	0.80	0.84

TABLEAU X. Différences structurales ( $\text{\AA}$ ) entre les quatre états d'occupation de la calbindine D<sub>9k</sub>, calculées à partir de structures simulées moyennes. L'abréviation c.p. signifie chaîne principale; tous les résultats excluent les atomes d'hydrogène et les Ca<sup>2+</sup>.

structure simulée A affiche elle-même les plus grandes différences par rapport à sa structure expérimentale de départ. Seules les comparaisons entre la simulation A(S<sub>II</sub>) et les autres structures simulées sont donc commentées, puisque la chaîne principale de la simulation A(S<sub>II</sub>) dévie moins (sauf pour le site I) par rapport à la structure expérimentale A, quoique cette déviation demeure supérieure à l'écart type de la structure expérimentale moyenne.

Les données présentées au tableau X s'accordent avec les données expérimentales<sup>32, 33</sup> et montrent que la calbindine D<sub>9k</sub> ne subit pas de changement conformationnel majeur suite à la liaison de un ou deux Ca<sup>2+</sup>, sans regard à l'ordre de liaison du Ca<sup>2+</sup>. Toujours en accord avec les mêmes données expérimentales, il ressort que les structures S<sub>I</sub>(D), S<sub>II</sub> et D affichent des différences de chaîne principale qui demeurent en deçà de l'écart type de la structure expérimentale moyenne. Ceci n'étant pas vrai pour toutes les comparaisons impliquant la structure A(S<sub>II</sub>), cette dernière se révèle donc comme différente des structures S<sub>I</sub>(D), S<sub>II</sub> et D qui ressortent similaires entre elles. Notons encore que la structure simulée S<sub>I</sub>(D) se révèle plus près de la structure simulée D que la structure simulée S<sub>II</sub>. Ceci s'accorde avec les résultats d'une étude effectuée sur un modèle de l'état S<sub>I</sub>(D) (mutant N56A) par Wimberly et coll.<sup>114</sup> qui ont montré que la structure de l'état S<sub>I</sub>(D) était presque identique à la structure de l'état D. Mais le détail des différences par rapport à l'état S<sub>II</sub> n'était pas encore disponible au moment de la rédaction de cette thèse.

Au niveau des éléments de structure secondaire, les hélices demeurent très similaires d'une structure simulée à l'autre avec des différences en-dessous de 1 Å pour la chaîne principale, sauf pour l'hélice D qui varie légèrement plus. Cette tendance est aussi en accord avec les données expérimentales.<sup>32, 33</sup> La région des acides aminés 36 à 45 (segment 36-45), qui correspond à la chaîne peptidique entre les deux mains EF, présente des différences structurales supérieures à celles des hélices. Toutefois, ce lien affiche les plus grands écarts types par rapport à

la position moyenne de chaque acide aminé dans les données expérimentales et est apparu mobile plus que tout autre élément de structure secondaire avec les extrémités de la protéine. Par exemple, le segment 36-45 de la structure A(S<sub>II</sub>) semble plus près de la structure simulée S<sub>II</sub> que de la structure expérimentale A. Les positions moyennes de ce segment sont donc à considérer avec circonspection.

Ce même commentaire vaut aussi pour les deux sites de liaison qui, comme le montrent les données expérimentales, fluctuent plus que les hélices, mais tout de même moins que le segment 36-45. Ceci est d'autant plus pertinent que les différences structurales qui contribuent à distinguer les différents états sont presque toutes attribuables aux sites de liaison. La liaison d'un Ca<sup>2+</sup> au site I affecte la conformation de la chaîne principale du site I (1.65 Å), mais aussi celle du site II (1.72 Å). De même, la liaison d'un Ca<sup>2+</sup> au site II affecte la chaîne principale de ce site (2.07 Å), ainsi que la chaîne principale du site I mais dans une moindre mesure (0.93 Å). Bien qu'il semble que la liaison d'un Ca<sup>2+</sup> à un site modifie la structure du site resté vide, la comparaison A(S<sub>II</sub>)-D montre que la modification de structure du site resté vide n'est pas complète. En effet, le site I dont la chaîne principale est différente par 0.93 Å dans la comparaison A(S<sub>II</sub>)-S<sub>II</sub> affiche des différences presque égales dans les comparaisons A(S<sub>II</sub>)-S<sub>I</sub>(D) (1.65 Å) et A(S<sub>II</sub>)-D (1.64 Å). Cela signifie qu'une modification supplémentaire de la chaîne principale d'environ 0.71 Å survient au moment de la liaison d'un Ca<sup>2+</sup> dans le site I. Le même phénomène est observé pour le site II dont la chaîne principale est différente par 1.72 Å de la structure A(S<sub>II</sub>) lorsqu'un Ca<sup>2+</sup> est présent dans le site I, et cette différence augmente à 2.07 Å et 2.03 Å respectivement dans les comparaisons A(S<sub>II</sub>)-S<sub>II</sub> et A(S<sub>II</sub>)-D. La venue d'un Ca<sup>2+</sup> dans le site II engendre donc une modification structurale supplémentaire du site II d'environ 0.33 Å.

La structure simulée S<sub>I</sub>(D), dont la structure expérimentale n'est pas disponible au moment de la rédaction de cette thèse, permet une comparaison structurale des deux états liés à un seul Ca<sup>2+</sup>. La comparaison S<sub>I</sub>(D)-S<sub>II</sub> montre que ces

deux structures sont aussi différentes entre elles que chacune des deux l'est de la structure A(S<sub>II</sub>), et que ces différences demeurent majoritairement concentrées dans les sites. Les comparaisons S<sub>I</sub>(D)-D et S<sub>II</sub>-D révèlent, comme il a déjà été mentionné, que la structure simulée S<sub>I</sub>(D) se montre plus près que la structure simulée S<sub>II</sub> de la structure simulée D. La différence S<sub>I</sub>(D)-D est en grande partie concentrée dans le site II et les différences pour toutes les hélices sont très faibles.

#### 4.3.1.3 Différences de coordinence et de disposition des ligands entre les structures expérimentales de départ et les structures simulées

Afin d'avoir une idée plus juste de la qualité du modèle de simulations, les ligands de chaque Ca<sup>2+</sup> dans chaque système simulé ont été déterminés et sont définis comme tels si l'atome d'oxygène se trouve à un rayon de moins de 3 Å. La coordinence, la disposition et les distances Ca<sup>2+</sup>-ligands dans les sites sont comparés aux données expérimentales. Les distances moyennes entre le Ca<sup>2+</sup> et ses ligands ont été calculées pour tous les états d'occupation, et sont comparées avec les distances initialement présentes dans la structure expérimentale correspondante.

Les résultats pour l'état D sont présentés au tableau XI. La différence majeure par rapport à la structure expérimentale est la présence d'une molécule d'eau supplémentaire dans le site I, pour un total de huit ligands. Pour le site II, la coordinence et la disposition des ligands restent en accord avec la structure expérimentale. Dans la majorité des cas, les distances Ca<sup>2+</sup>-O sont plus courtes, sauf pour Asp19 et Gln22 dont les atomes d'oxygène se trouvent  $\approx 0.2$  Å plus loins. Notons que ces résultats représentent une amélioration par rapport aux simulations effectuées par d'autres auteurs sur la calbindine D<sub>9k</sub>: lors de la simulation de Ahlström et coll.<sup>108</sup> le Ca<sup>2+</sup> dans le site I perdait la coordinence avec l'atome d'oxygène provenant de la chaîne principale du Ala14, et gagnait deux atomes d'oxygène du groupement carboxyle du Glu60, pour une coordinence de huit;

état D		
site I	exp.	simul.
Ala14 O	2.33	2.28±0.08
Glu17 O	2.47	2.36±0.12
Asp19 O	2.25	2.46±0.23
Gln22 O	2.36	2.53±0.22
Glu27 OE1	2.42	2.18±0.06
Glu27 OE2	2.60	2.23±0.08
H <sub>2</sub> O O	2.43	2.29±0.56
H <sub>2</sub> O O	absent	2.54±0.77
site II		
Asp54 OD1	2.41	2.10±0.05
Asp56 OD1	2.34	2.28±0.09
Asp58 OD1	2.38	2.16±0.06
Glu60 O	2.39	2.28±0.08
Glu65 OE1	2.54	2.20±0.07
Glu65 OE2	2.54	2.30±0.10
H <sub>2</sub> O O	2.54	2.26±0.07

TABLEAU XI. Distances ( $\text{\AA}$ ) entre le  $\text{Ca}^{2+}$  et ses ligands dans l'état D; exp. signifie dans la structure expérimentale de départ; simul. est la moyenne des distances évaluées dans les 20 000 configurations de l'ensemble. Les atomes de type OE\* ou OD\* appartiennent à une chaîne latérale et les atomes O appartiennent à la chaîne principale.

le  $\text{Ca}^{2+}$  dans le site II gagnait pour sa part un atome d'oxygène du groupement carboxyl de l'Asp54. Lors de la simulation de Mehler et coll.,<sup>110</sup> le  $\text{Ca}^{2+}$  dans le site I échangeait aussi un ligand (Glu27) pour un atome d'oxygène provenant de Glu60, tandis que le  $\text{Ca}^{2+}$  dans le site II demeurait lié aux mêmes atomes que dans la structure de départ, mais avec des distances  $\text{Ca}^{2+}$ -ligand en général plus grandes. Les fluctuations des distances apparaissent dans la colonne des résultats de simulation et sont très faibles, bien qu'un peu plus élevées les molécules  $\text{H}_2\text{O}$ . La plus grande distance entre le  $\text{Ca}^{2+}$  et l'atome d'oxygène des Asp19 et Gln22, de pair avec leurs plus grandes fluctuations, jouent certainement un rôle dans la présence d'une molécule d'eau supplémentaire dans le site I. Les deux atomes d'oxygène du Glu27 semblent être configurés de façon légèrement différente dans les structures expérimentale et simulée. La répartition des deux atomes d'oxygène est plus à caractère *bidenté* dans la simulation que dans l'expérience.

L'état  $S_I(D)$ , dont les résultats sont présentés au tableau XII, conserve tous les mêmes ligands de la structure expérimentale D dont elle est issue. Cette structure n'ayant pas encore été déterminée expérimentalement, il n'est donc pas possible d'évaluer la qualité du résultat simulé, sinon qu'en comparant avec la structure expérimentale D. Notons qu'ici, la distance  $\text{Ca}^{2+}$ -O de l'Asp19 de la structure  $S_I(D)$  simulée demeure plus près de la distance expérimentale et fluctue moins que dans la simulation de l'état D. L'oxygène du Gln22 se situe par contre à peu près à la même distance que dans la simulation de l'état D et affiche des fluctuations d'amplitude similaire. Ce rapprochement de l'Asp19 et la plus grande stabilité de cette distance suffisent peut-être à ce qu'il n'y ait pas de molécule d'eau supplémentaire dans le site.

Les données expérimentales concernant la nature et la disposition des ligands du  $\text{Ca}^{2+}$  dans l'état  $S_{II}$  ne sont pas disponibles puisque la structure expérimentale a été déterminée par RMN. Les distances montrées dans la colonne exp. représentent les distances entre le  $\text{Ca}^{2+}$  et ses ligands tout de suite après avoir copié les

état S <sub>I</sub> (D)		
site I	exp.	simul.
Ala14 O	n.d.	2.26±0.08
Glu17 O	n.d.	2.31±0.10
Asp19 O	n.d.	2.28±0.09
Gln22 O	n.d.	2.51±0.26
Glu27 OE1	n.d.	2.18±0.07
Glu27 OE2	n.d.	2.22±0.08
H <sub>2</sub> O O	n.d.	2.25±0.07

TABLEAU XII. Distances entre le Ca<sup>2+</sup> et ses ligands dans l'état S<sub>I</sub>(D). n.d. signifie non déterminé. Voir le tableau XI pour les détails.

coordonnées du Ca<sup>2+</sup> de la structure expérimentale D à la structure expérimentale S<sub>II</sub> avant que ne soit effectuée la minimisation de l'énergie du site mais après que la position de l'ion dans le site rigide ait été optimisée. (voir section 4.2.1). Il ne s'agit donc pas des distances expérimentales mais d'une estimation de la nature et de la position des ligands déduite d'après la structure expérimentale. On voit néanmoins que la structure simulée S<sub>II</sub> acquiert un ligand supplémentaire par rapport au nombre initial (Asp54 qui devient *bidenté*) et qu'on retrouve une disposition des ligands plus symétrique autour de l'ion grâce au processus de minimisation de l'énergie et d'équilibration. C'est peut-être à cause du manque de symétrie après copiage de l'ion qu'un ligand supplémentaire saisit l'opportunité de venir rapidement s'insérer dans le patron de coordinence. Peut-être qu'une période d'équilibration avec contrainte de plus longue durée aurait suffi à conserver la structure des ligands autour de l'ion.

Aussi, des différences par rapport à la structure simulée D ressortent à l'étude du tableau XIII, soit la présence d'un deuxième atome d'oxygène provenant du groupement carboxylique de Asp54 pour un total de huit ligands. Il ressort aussi

que certaines distances dans la structure simulée  $S_{II}$  sont plus grandes que dans la structure simulée D. De plus, les fluctuations sont systématiquement plus grandes dans la simulation  $S_{II}$  (sauf pour Asp58), avec une molécule d'eau qui entre et qui sort du site (fluctuations de 2.96 Å au tableau XIII). Encore une fois, ces observations s'apparentent probablement au fait que l'Asp54 passe de *mono* à *bidenté*. Rappelons que ce ligand supplémentaire provenant de Asp54 survenait aussi lors de la simulation de l'état D réalisée par Ahlström et coll.<sup>108</sup>

état $S_{II}$		
site II	exp.	simul.
Asp54 OD1	2.07	2.31±0.12
Asp54 OD2	absent	2.21±0.25
Asp56 OD1	2.15	2.24±0.13
Asp58 OD2	2.11	2.16±0.06
Glu60 O	2.19	2.35±0.11
Glu65 OE1	2.15	2.24±0.11
Glu65 OE2	(3.30)	2.47±0.35
H <sub>2</sub> O O	2.53	2.32±2.96

TABLEAU XIII. Distances entre le  $Ca^{2+}$  et ses ligands dans l'état  $S_I(D)$ . Voir le tableau XI pour les détails.

Le fait que la plupart des ligands se retrouvent plus près du  $Ca^{2+}$  dans la structure simulée D qu'à l'expérience, ainsi que la présence de ligands supplémentaires dans certains cas, s'accorde avec les résultats du chapitre 3 et est probablement une conséquence du fait que la polarisation induite est négligée (les charges partielles attribuées aux atomes sont des paramètres fixes). En réalité, le  $Ca^{2+}$  polarise la structure électronique de ses ligands, résultant en une densité accrue de charges négatives autour du  $Ca^{2+}$ . Par contre, dans le modèle de simulation, les charges partielles attribuées aux atomes ne pouvant pas s'ajuster à

la présence du  $\text{Ca}^{2+}$ , on assiste plutôt au déplacement d'atomes supplémentaires dans l'environnement du  $\text{Ca}^{2+}$ , et donc d'une coordinence accrue. La déformation des sites dans les structures simulées par rapport aux structures expérimentales est aussi probablement une conséquence de ces modifications des distances  $\text{Ca}^{2+}$ -ligands par rapport aux valeurs expérimentales.

#### 4.3.1.4 Différences de disposition des ponts hydrogènes entre les structures expérimentales et les structures simulées, et des structures simulées entre elles

Les distances moyennes entre les atomes d'hydrogène liés à l'atome d'azote de la chaîne principale (appelé HN) et les atomes d'oxygène de la chaîne principale (appelé O) ont été calculées pour les cinq structures simulées. Les distances HN-O inférieures à 3 Å ont été retenues et considérées comme formant un pont hydrogène. Pour chaque structure simulée, le réseau de ponts hydrogènes a été comparé à celui retrouvé dans la structure expérimentale,<sup>33</sup> sauf pour l'état  $S_I(D)$  où ces résultats n'étaient pas connues au moment de la rédaction de cette thèse. Dans ce cas, les autres structures expérimentales sont utilisées comme référence. Malgré le fait que la structure expérimentale de l'état D provienne de données cristallographiques, les ponts hydrogènes déterminés dans la structure RMN ont été utilisés pour fins de comparaison puisque le milieu est le même que dans la simulation.

De façon générale, le réseau de ponts hydrogènes de la chaîne principale des structures simulées est demeuré le même que dans les structures expérimentales; les atomes HN et O appartenant respectivement aux acides aminés  $i$  et  $i-4$  forment le patron général des ponts hydrogènes, sauf pour l'hélice D des états  $S_I(D)$ ,  $S_{II}$  et D où une hélice<sub>3/10</sub> est formée par quelques liens HN-O appartenant aux acides aminés  $i$  et  $i-3$ . Ce passage d'une hélice  $\alpha$  régulière à une hélice<sub>3/10</sub> en allant de l'état A vers les états  $S_I(D)$ ,  $S_{II}$  et D fait partie des différences associées à la

liaison de  $\text{Ca}^{2+}$ .

Le tableau XIV donne la liste des ponts hydrogènes qui sont présents dans la structure expérimentale correspondante mais absents dans la structures simulée. La structure simulée A( $\text{S}_{\text{II}}$ ) n'est pas mentionnée dans le tableau XIV puisqu'elle a retenu tous les ponts hydrogènes de la structure expérimentale A. La structure simulée A, pour sa part, subit quelques modifications: les ponts hydrogènes annotés P au tableau XIV sont remplacés par des ponts hydrogènes avec des molécules d'eau tandis que les autres ponts hydrogènes forment des ponts hydrogènes avec l'acide aminé précédent ou suivant.

A	$\text{S}_{\text{I}}(\text{D})$	$\text{S}_{\text{II}}$	D
(HN-O)	(HN-O)	(HN-O)	(HN-O)
(Leu6, Ser2)			
(Lys7, Pro3)			
(Ala15, Glu11) <sup>P</sup>	(Ala15, Glu11) <sup>P</sup>		(Ala15, Glu11)
(Phe50, Leu46)	(Gln22, Asp19) <sup>P</sup>		(Lys16, Tyr13) <sup>P</sup>
(Glu52, Glu48) <sup>P</sup>	(Gln33, Lys29)	(Lys16, Tyr13)	(Lys16, Tyr13) <sup>P</sup>
(Leu53, Leu49)	(Asp54, Phe50)	(Glu52, Glu48)	(Lys55, Leu53) <sup>P</sup>
(Asp54, Phe50)	(Lys55, Leu53)		(Val68, Glu65) <sup>P</sup>
*(Val61, Leu23) <sup>P</sup>			
(Val68, Glu64)			
(Lys72, Leu69)			

TABLEAU XIV. Ponts hydrogènes reliant les atomes HN à O de la chaîne principale qui sont présents dans les structures expérimentales (RMN) mais qui sont absents dans les structures simulées. La structure simulée A( $\text{S}_{\text{II}}$ ) est absente de ce tableau parce qu'elle a retenu tous les ponts hydrogènes de la structure expérimentale A. L'annotation p signifie que le pont hydrogène a été perdu; les autres ponts hydrogènes de la liste établissent des ponts hydrogènes avec un acide aminé adjacent. L'astérisque indique que ce pont hydrogène relie les deux sites.

Pour l'état  $S_I(D)$ , les ponts hydrogènes communs aux états A,  $S_{II}$  et D ont été retenus pour la comparaison. Les ponts hydrogènes n'apparaissant que dans l'état D et absents dans la simulation  $S_I(D)$  n'ont pas été considérés manquants. Un total de deux ponts hydrogènes, faisant partie du site I, ont été perdus au profit de ponts hydrogènes avec des molécules d'eau, les autres ont été remplacés par des liaisons impliquant un acide aminé adjacent.

Aucun pont hydrogène n'a été perdu dans la structure simulée  $S_{II}$ , bien que deux soient remplacés par des liaisons avec d'autres acides aminés adjacents. La structure simulée D perd trois ponts hydrogènes dont un dans le site I et deux dans le site II, et qui sont aussi remplacés par des ponts hydrogènes avec des molécules d'eau, en plus d'un autre faisant partie du site I qui est modifié.

La majorité des ponts hydrogènes perdus (P) se trouvent directement dans un site, ou très près d'un site. L'origine de ces bris ne peut être déterminée avec certitude, mais le fait qu'il y ait déformation du site dans les structures simulées avec  $Ca^{2+}$ , pourrait impliquer des tensions ayant causé le bris de certains ponts hydrogènes. Par contre, il en va autrement de la structure simulée A pour laquelle cette hypothèse ne tient pas mais dont il sera discuté à la section 4.3.2.

#### 4.3.1.5 Différences de compacité entre les structures simulées

Le rayon de giration et la surface accessible à l'eau ont été calculés pour chaque structure simulée et comparés avec la valeur obtenue à partir de la structure expérimentale de départ. Ceci constitue une mesure de la compacité relative de chaque structure simulée ainsi qu'une autre façon d'évaluer de combien une structure simulée s'est éloignée de la structure expérimentale.

Le rayon de giration calculé pour la calbindine  $D_{9k}$  est défini par

$$r_G = \sqrt{\frac{\sum_i m_i r_i^2}{\sum_i m_i}} \quad (4.3)$$

	A	A(S <sub>II</sub> )	S <sub>I</sub> (D)	S <sub>II</sub>	D
$r_G$ (Å) (simul.)	12.47	11.65	11.49	11.46	11.29
(exp.)	11.31	–	–	11.08	11.26
$\sigma_{acc.}$ (Å <sup>2</sup> ) (simul.)	5750.21	5056.35	4852.81	4867.60	4636.44
(exp.)	4694.44	–	–	4425.72	4610.81

TABLEAU XV. Compilation des rayons de giration moyens ( $r_G$ ) et des surfaces accessibles moyennes ( $\sigma_{acc.}$ ) pour les cinq structures simulées.

où  $m_i$  est la masse de l'atome  $i$ ,  $r_i$  est la distance de cet atome au centre de masse de la protéine, et le calcul inclus tous les atomes.

Le rayon de giration moyen de chaque structure simulée a été calculé à partir de chaque trajectoire respective. Les résultats sont donnés au tableau XV; on y trouve aussi les valeurs calculées pour chaque structure expérimentale de départ. Il semble que les structures simulées aient un rayon de giration qui varie inversement au niveau d'occupation du  $Ca^{2+}$ , observation qui n'est cependant pas appuyée par les données expérimentales, bien qu'il ne s'agisse que d'une seule mesure prise sur la structure de départ. Il ressort que la structure A se relâche pendant la simulation, et que la structure simulée A(S<sub>II</sub>) se révèle plus près de la structure expérimentale A. Les structures simulées S<sub>I</sub>(D) et S<sub>II</sub> ont des rayons de giration très similaires, et bien qu'il n'y ait pas de donnée expérimentale pour la S<sub>I</sub>(D), on observe que la structure S<sub>II</sub> s'est légèrement décontractée au cours de la simulation, et l'on peut supposer qu'il en est de même pour la structure simulée S<sub>I</sub>(D). Dans le cas de la structure simulée D, le rayon de giration de départ est presque parfaitement maintenu jusqu'à la fin de la simulation.

Le tableau XV donne aussi la valeur de la surface accessible moyenne  $\sigma_{acc.}$  pour chaque structure simulée et de chaque structure expérimentale de départ. La surface accessible dite de Lee et Richards<sup>115</sup> est calculée en faisant rouler une sonde sphérique d'un rayon de 1.4 Å sur la surface de la protéine de sorte à

ce que la sonde explore toute la surface accessible et en faisant la somme. Les résultats nous montrent la même tendance que pour le rayon de giration, c'est-à-dire que la structure simulée A s'éloigne le plus de sa structure expérimentale de départ, mais que la structure simulée A(S<sub>II</sub>) en est aussi proche que les structures simulées S<sub>I</sub>(D) et S<sub>II</sub> le sont de la structure expérimentale S<sub>II</sub>. Les deux structures simulées S<sub>I</sub>(D) et S<sub>II</sub> ont des surfaces accessibles très similaires entre elles mais légèrement différentes de la structure expérimentale S<sub>II</sub>. Enfin, la structure simulée D demeure presque en parfait accord avec la valeur obtenue à partir de la structure expérimentale D.

Une mesure très révélatrice de la compacité consiste à compiler toutes les paires d'atomes de la chaîne principale dont la distance moyenne (sur toute la trajectoire) est plus petite qu'une valeur maximale (ici 5 Å). Les figures 21 à 23 montrent ces résultats pour chaque structure simulée. De même que les résultats du rayon de giration et de la surface accessible, les matrices des contacts suggèrent que les deux liaisons successives de Ca<sup>2+</sup> engendrent une contraction de la protéine.

La structure simulée A est celle qui présente le moins de contacts. Elle est encore une fois différente de la A(S<sub>II</sub>) qui présente des contacts supplémentaires entre les sites I et II et entre les bouts terminaux N et C, mais quelques contacts en moins entre les atomes du site I. Les deux structures ne présentent aucun contact interne au site II.

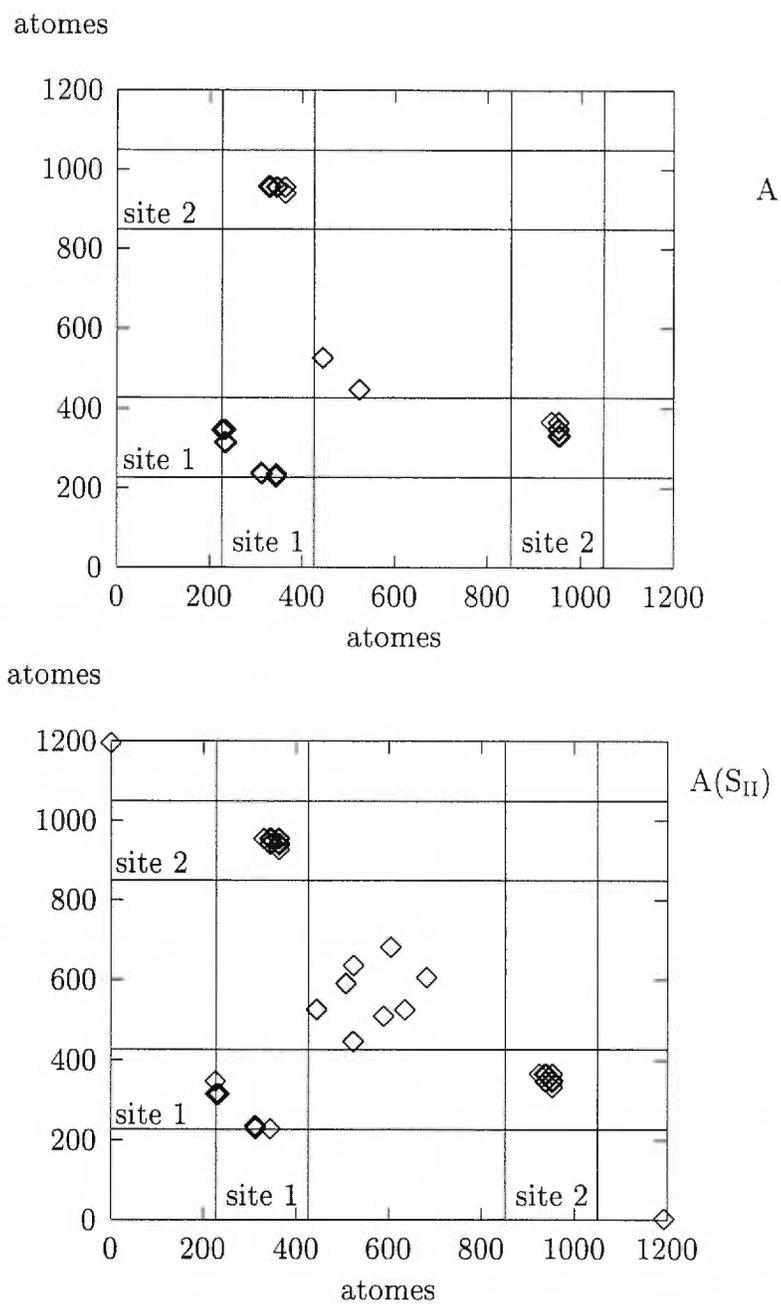


FIGURE 21. Matrices des contacts des structures simulées A et A(S<sub>II</sub>): compilation des atomes lourds de la chaîne principale se trouvant à moins de 5 Å d'un autre atome lourd de la chaîne principale. Les deux atomes doivent aussi être séparés par au moins 75 atomes dans la séquence primaire.

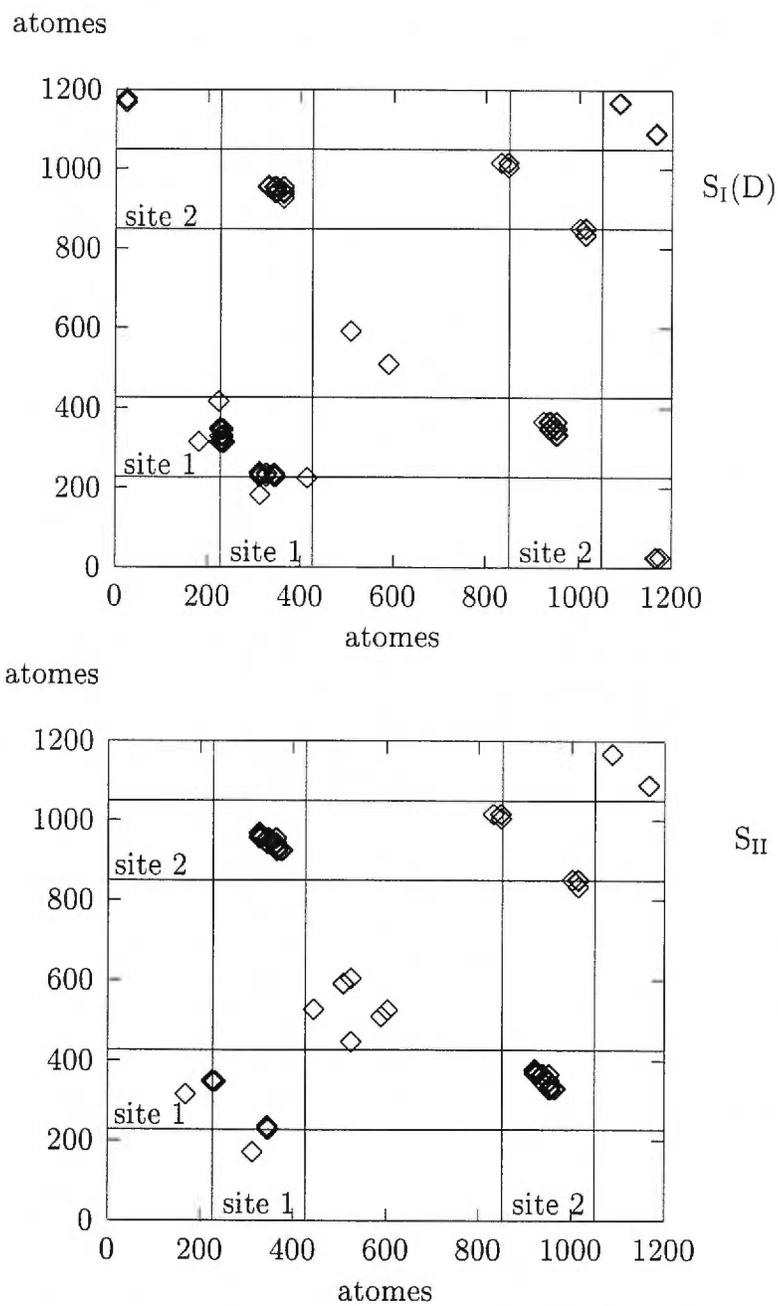


FIGURE 22. Matrices des contacts des structures simulées  $S_I(D)$  et  $S_{II}$ . Voir figure 21.

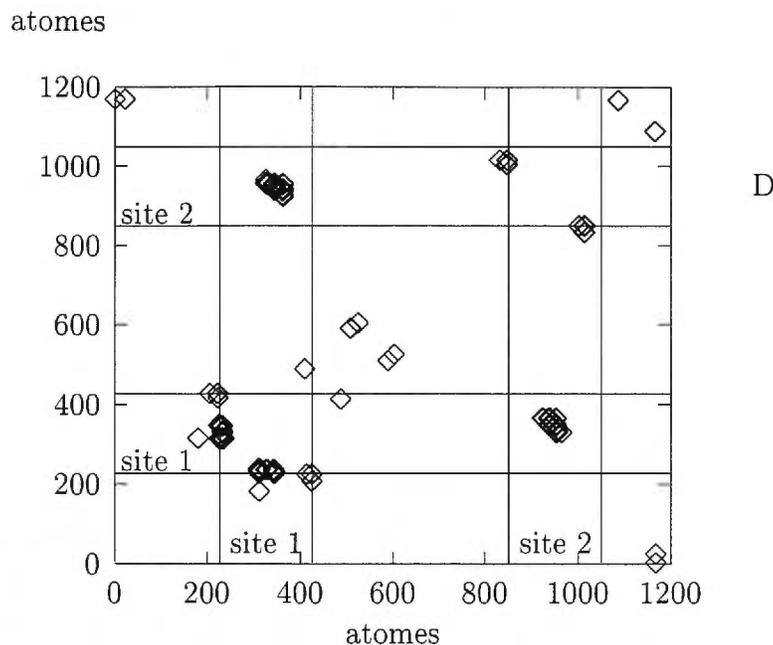


FIGURE 23. Matrice des contacts de la structure simulée D. Voir figure 21.

La structure simulée  $S_I(D)$  présente beaucoup plus de contacts internes au site I, entre les deux sites et, contrairement aux structures simulées A et  $A(S_{II})$ , des contacts internes au site II. Quelques contacts supplémentaires apparaissent aussi entre les bouts terminaux N et C. La structure simulée  $S_{II}$  présente quelques contacts internes au site II de plus que la structure  $S_I(D)$ , et davantage de contacts entre les sites I et II, mais beaucoup moins de contacts internes au site I. Des contacts entre les bouts terminaux ne sont pas présents. La structure simulée D retient les contacts supplémentaires des structures simulées  $S_I(D)$  et  $S_{II}$  qui apparaissent lorsque comparées aux structures simulées A et  $A(S_{II})$ , c'est-à-dire tous les changements structuraux engendrés par la liaison des deux  $Ca^{2+}$ . Ces observations sont en accord avec les résultats du rayon de giration et de la surface accessible au solvant.

#### 4.3.1.6 Différences d'hydratation des chaînes principales, des chaînes latérales et des sites entre les structures simulées

L'hydratation moyenne des chaînes principale et latérale de chaque acide aminé a été évaluée pour toutes les structures simulées en calculant le nombre de molécules d'eau contenues dans un rayon de 4 Å des chaînes principale et latérale de chaque acide aminé. Ces résultats sont montrés aux figures 24 et 25.

Une première caractéristique remarquable est l'alternance de grande et faible hydratation à environ tous les deux ou trois acides aminés dans les hélices. Ceci correspond probablement à la position de la chaîne latérale qui pointe vers l'extérieur de la protéine directement dans le solvant, ou vers l'intérieur de la protéine, selon la position de l'acide aminé dans l'hélice. L'alternance observée provient du nombre d'acides aminés par tour d'hélice. Ces alternances ne sont d'ailleurs observées que dans une très faible mesure dans les sites qui sont positionnés vers l'extérieur de la protéine et donc plus exposés au solvant.

La figure 24 du haut montre que la structure simulée A se distingue des autres par une plus grande hydratation du segment 36-45 et des acides aminés 23 et 24, mais que son site II est légèrement moins hydraté que celui de la structure simulée A(S<sub>II</sub>). Etant donné que la structure simulée A(S<sub>II</sub>) est globalement demeurée plus près de la structure expérimentale A, les comparaisons suivantes sont effectuées principalement avec la structure A(S<sub>II</sub>), mais rappelons que le site I de la structure simulée A est demeurée plus près de la structure expérimentale (voir tableau IX).

La figure 24 du bas montre que le profil d'hydratation de la chaîne principale est très similaire pour toutes les structures simulées. Il apparaît cependant que le site II des structures simulées A(S<sub>II</sub>) et S<sub>I</sub>(D) est légèrement plus hydraté que pour les deux autres (S<sub>II</sub> et D) qui ont un Ca<sup>2+</sup> lié à majoritairement à des atomes d'oxygène des chaînes latérales dans ce site. En particulier, il semble que la liaison d'un premier Ca<sup>2+</sup> dans le site II déshydrate la chaîne principale des

acides aminés 21 à 25. Bien que ce soit un effet modéré, ce résultat suggère que la liaison d'un  $\text{Ca}^{2+}$  dans le site II a pour effet, dans une certaine mesure, de préparer le site I de sorte à ce que l'oxygène de la chaîne principale de l'acide aminé 22 ait préalablement été libéré des molécules d'eau environnantes. Les résultats pour les chaînes latérales sont présentées à la figure 25.

Au haut, on note cette fois que la structure simulée A affiche toujours quelques différences par rapport à la structure  $A(S_{II})$ , surtout dans le segment 36-45 ainsi que pour quelques acides aminés du site I. La figure du bas montre que les chaînes latérales du site II sont particulièrement affectées par la présence du  $\text{Ca}^{2+}$ , puisque ce dernier est lié par les chaînes latérales dans le site II. Pour le site I, il semble que la chaîne latérale de l'acide aminé 18 des structures simulées  $S_I(D)$  et D est plus hydratée que celle des deux autres ( $A(S_{II})$  et  $S_{II}$ ) qui n'ont pas de  $\text{Ca}^{2+}$  dans le site I. Peut-être que la présence d'un  $\text{Ca}^{2+}$  dans le site I, lié principalement à des oxygènes de la chaîne principale, force la chaîne latérale de la glycine à se réorienter vers l'extérieur de la protéine, directement dans le solvant.

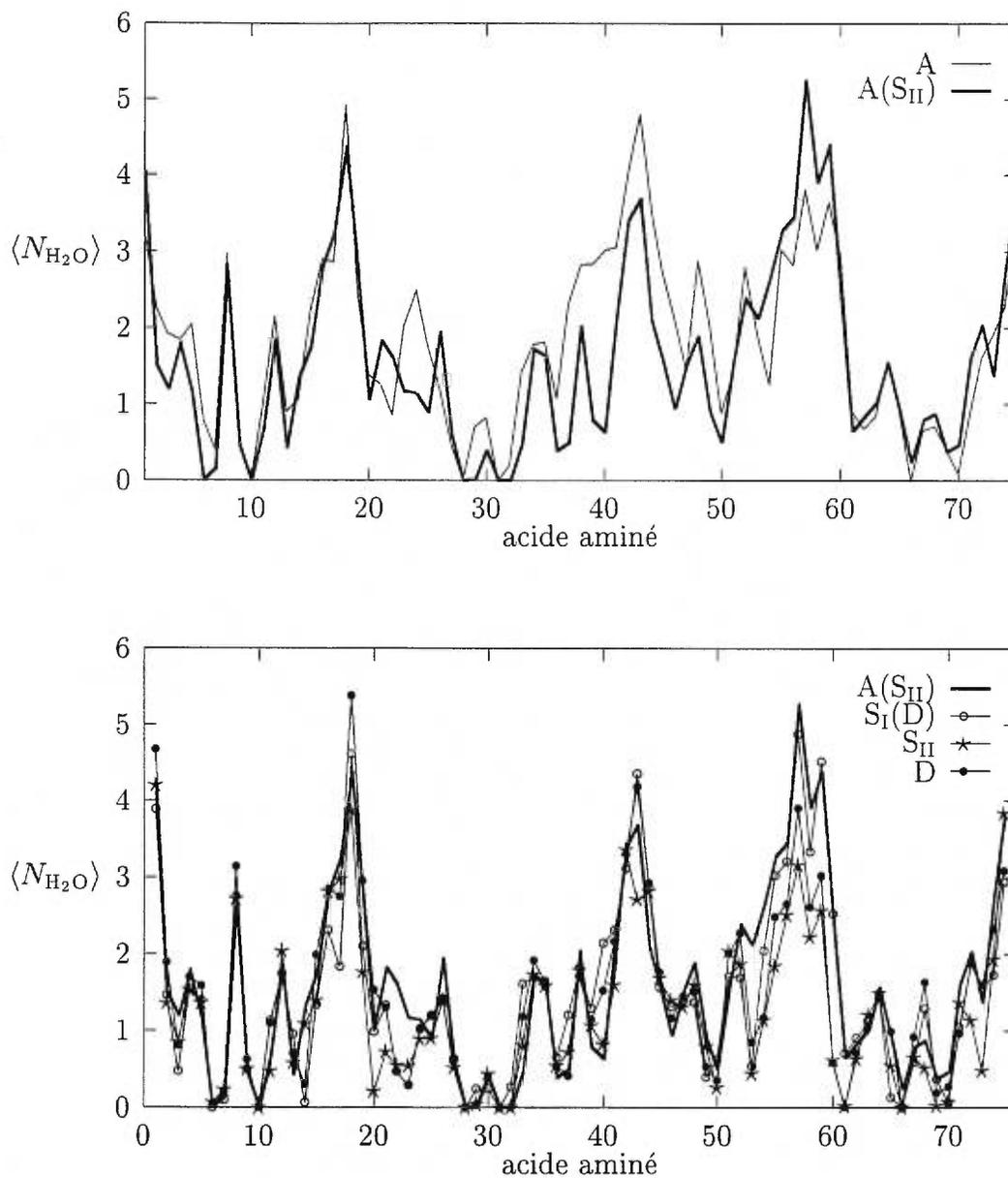


FIGURE 24. Hydratation de la chaîne principale des structures simulées. L'ordonnée donne le nombre de molécules d'eau dans un rayon de 4 Å du centre de masse de la chaîne principale d'un acide aminé.

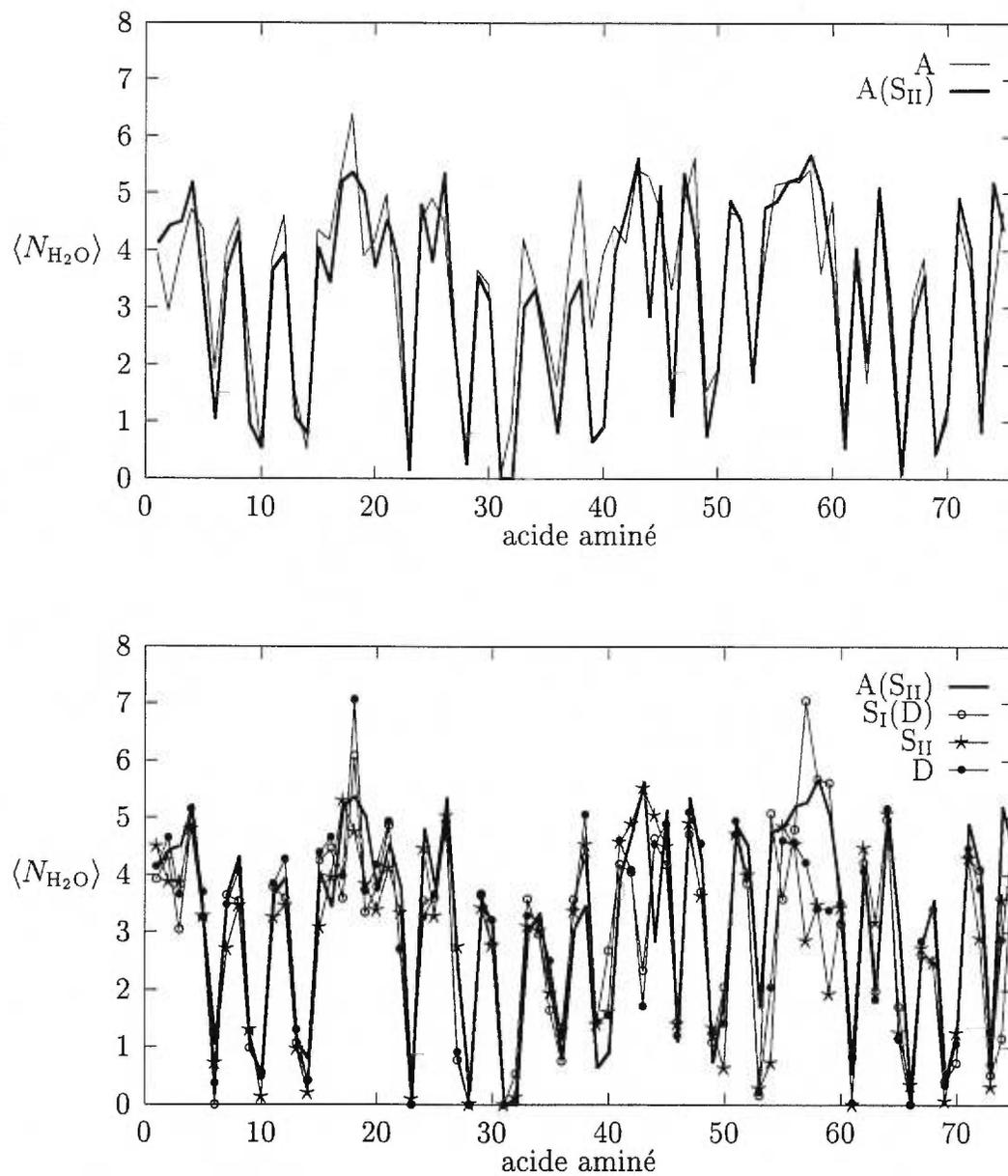


FIGURE 25. Hydratation de la chaîne latérale de chaque acide aminé pour toutes les structures simulées. (Voir figure 24.)

### 4.3.2 Analyse comparée des propriétés dynamiques à l'équilibre

#### 4.3.2.1 Fluctuations moyennes autour de la position moyenne pour chaque acide aminé

Lors du calcul des structures simulées moyennes, les fluctuations moyennes (écart type) autour de la position moyenne de chaque acide aminé ont été calculées en ne tenant compte que des atomes lourds. Les résultats sont présentés à la figure 26.

On note ici aussi le même genre d'alternance en fonction du numéro d'acide aminé qu'aux figures 24 et 25 et qui dépend probablement de sa position vers l'extérieur (plus grandes fluctuations) ou vers l'intérieur de la protéine (moins grandes fluctuations). L'état A présente le plus de fluctuations en comparaison avec les états  $S_I(D)$ ,  $S_{II}$  et D dont les acides aminés fluctuent nettement moins. De plus, la structure simulée A fluctue considérablement plus que la structure simulée  $A(S_{II})$  dans les régions du bout N-terminal, du segment 36-45 et du site II, là où elle diffère le plus de la structure expérimentale. Malgré les plus faibles fluctuations dans la structure simulée  $A(S_{II})$ , il demeure que les bouts N et C-terminaux et les sites I et II se distinguent des hélices par la plus grande amplitude de leurs fluctuations. Il est étonnant de constater que le site II de la structure simulée  $A(S_{II})$  fluctue moins que la structure simulée  $S_{II}$  dont elle est issue, mais cela a peut-être à voir avec le fait qu'un pont hydrogène (Glu52, Glu48) a été modifié (mais non perdu) dans le site II de la  $S_{II}$ , mais pas dans la  $A(S_{II})$ .

La présence d'un  $Ca^{2+}$  dans un des deux sites fait diminuer le niveau des fluctuations du site où il est directement lié, mais aussi dans une certaine mesure le site resté inoccupé, et plus généralement le niveau de base de fluctuations des hélices. Ceci est particulièrement remarquable pour la structure simulée  $S_I(D)$  dont le niveau de fluctuations est plus faible que pour la structure simulée  $S_{II}$  pour tous les éléments de structure secondaire, sauf le segment 36-45.

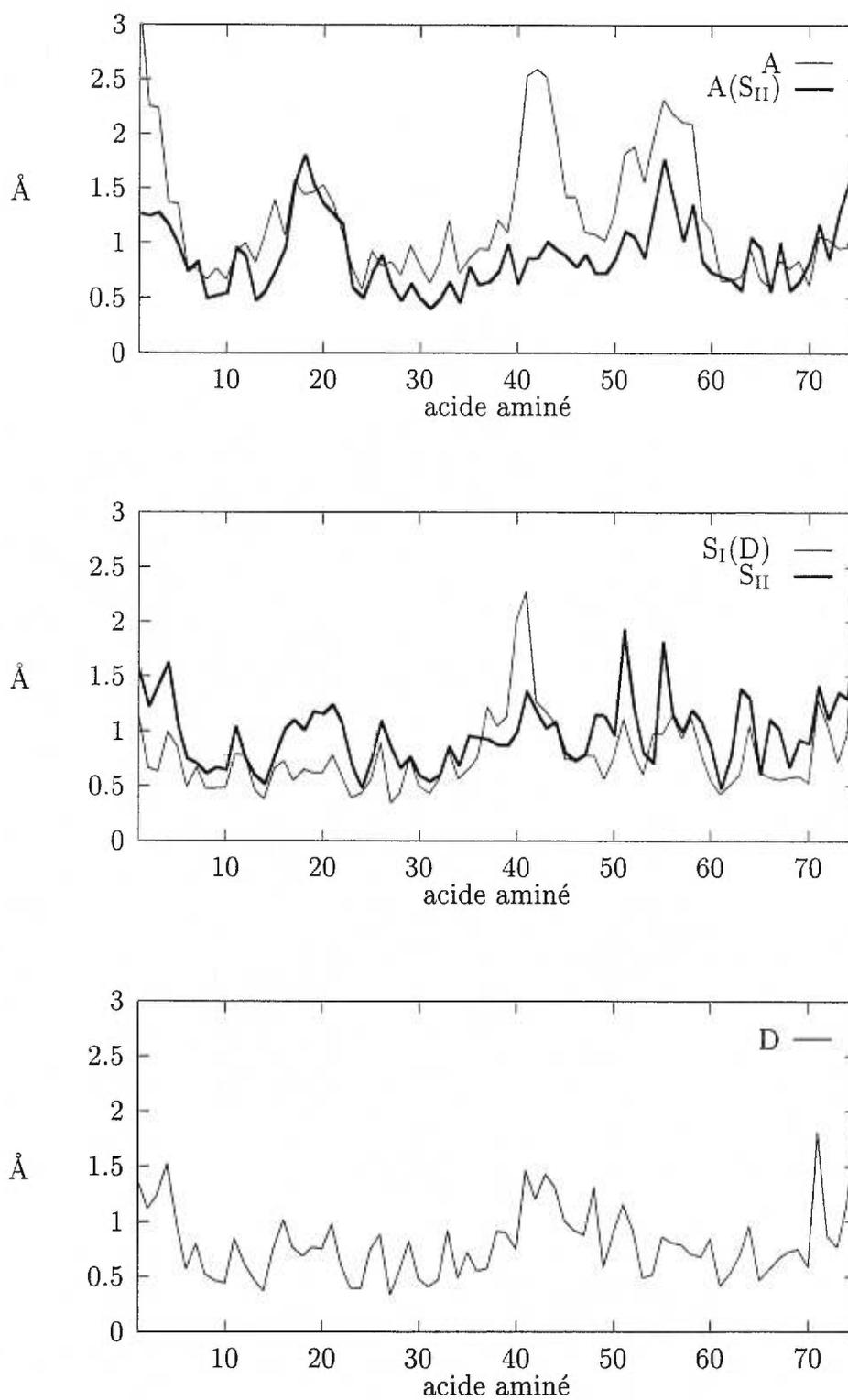


FIGURE 26. Fluctuations moyennes ( $\text{\AA}$ ) autour de la position moyenne de chaque acide aminé pondérées par la masse de chaque atome lourd inclus dans le calculs.

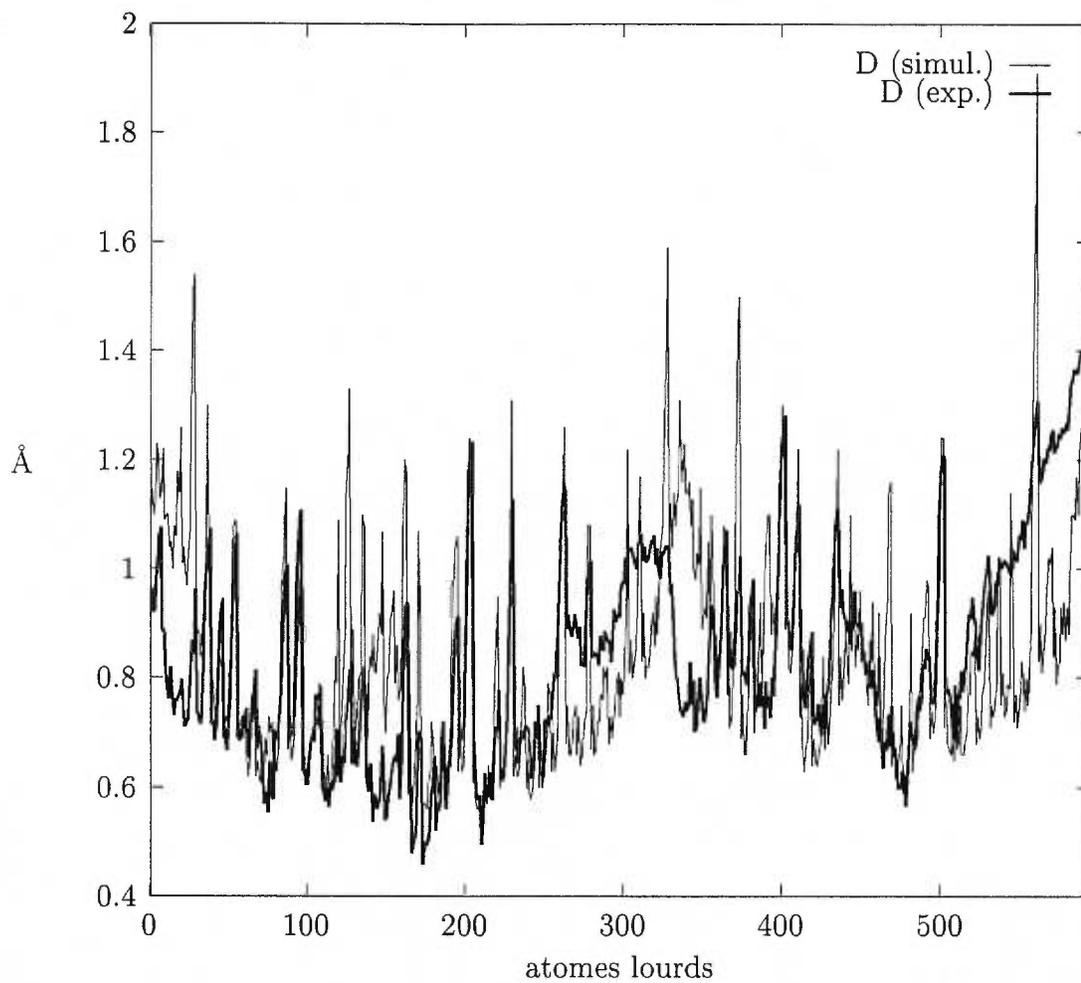


FIGURE 27. Comparaison entre les fluctuations moyennes ( $\text{\AA}$ ) de chaque atome lourd de la calbindine  $D_{9k}$  calculées à partir de la trajectoire (simul.) et calculées à partir des facteurs  $B$  déterminés par diffraction des rayons X (exp.).<sup>15</sup>

Dans le cas de la structure simulée  $S_{II}$ , la présence d'un  $Ca^{2+}$  dans le site II fait diminuer les fluctuations par rapport à la structure simulée A. Mais cette remarque n'est cependant pas aussi vraie lorsque la comparaison est faite avec la structure simulée  $A(S_{II})$ : en effet les acides aminés 51 et 55 du site II de la  $S_{II}$  continuent de fluctuer autant que dans la structure simulée  $A(S_{II})$ . De façon générale, les fluctuations de la structure simulée  $S_I(D)$  se retrouvent à un niveau très similaire à celles de la structure simulée D, sauf pour le segment 36-45 qui fluctue plus dans la structure simulée  $S_I(D)$ . En fait, il est étonnant de constater que le niveau des fluctuations dans la structure simulée  $S_I(D)$  est plus bas que dans la structure simulée D. Néanmoins, la structure simulée D affiche un niveau de fluctuations assez bas, en comparaison aux autres structures simulées et, mis à part les bouts terminaux N et C ainsi que le segment 36-45, il ne semble pas y avoir de région qui se démarque par un niveau plus élevé de fluctuations.

Il ressort que la structure de départ d'une structure simulée joue un rôle important dans le comportement dynamique de la structure simulée, à tout le moins pour le cas de l'état A. Il serait donc hasardeux de poser des conclusions plus détaillées quant aux conséquences du niveau d'occupation de  $Ca^{2+}$  de la calbindine  $D_{9k}$ , bien qu'il semble se dégager que ce dernier affecte de façon générale à la baisse le niveau des fluctuations, à commencer par celui des sites et du segment 36-45.

Les fluctuations des atomes lourds de la structure simulée D et calculées à partir de la trajectoire (les valeurs obtenues avant le calcul des moyennes par acide aminé) peuvent être comparées aux fluctuations atomiques reliées aux facteurs  $B$ , dits de température ou de Debye-Waller, qui sont déterminés par diffraction des rayons X. Ces facteurs  $B$  étant disponibles pour la structure expérimentale D,<sup>15</sup> ils ont été convertis en fluctuations atomiques  $\langle \Delta r_i^2 \rangle_{\text{dynamique}}$  approximées comme étant harmoniques et isotropes, et auxquelles ils sont reliés par:<sup>54</sup>

$$\sqrt{\langle \Delta r_i^2 \rangle_{\text{dynamique}}} = \sqrt{\frac{3}{8\pi^2} B_i - \langle \Delta r_i^2 \rangle_{\text{désordre}}} \quad (4.4)$$

où  $B_i$  est le facteur B de l'atome  $i$  et  $\langle \Delta r_i^2 \rangle_{\text{désordre}}$  est la contribution provenant du désordre du réseau cristallin et autres effets. Comme ces différentes contributions n'étaient pas disponibles dans le cas de la structure de l'état D de la calbindine D<sub>9k</sub>, nous avons donc posé que  $\langle \Delta r_i^2 \rangle_{\text{désordre}} = 0$ .

La figure 27 compare les résultats de la simulation D aux données expérimentales.<sup>15</sup> La limite inférieure des deux enveloppes de fluctuations est très similaire, et correspond aux fluctuations des atomes lourds de la chaîne principale. Une analyse détaillée montre que les bouts terminaux N et C, les acides aminés 18, 19 et 20 du site I ainsi que les acides aminés 42 à 45 (segment 36-45) fluctuent légèrement plus dans la simulation, tandis que les acides aminés 33 à 41 (segment 36-45) fluctuent moins. Les chaînes latérales sont en moyenne plus flexibles dans la simulation. Ceci est peut-être relié au fait qu'il n'y a pas de contraintes imposées par le réseau cristallin dans le cas de la simulation de la protéine en solution.

#### 4.3.2.2 Fluctuations moyennes des angles dièdres $\phi$ et $\psi$

La définition d'un angle dièdre a été donnée à la figure 9. L'angle dièdre  $\phi$  est obtenu en remplaçant les atomes fictifs (1, 2, 3 et 4) par les atomes C, N, C $_{\alpha}$  et C de la chaîne principale, tandis que l'angle dièdre  $\psi$  est obtenu en les remplaçant par les atomes N, C $_{\alpha}$ , C et N de la chaîne principale. Comme le premier atome lourd d'une protéine est un atome N (bout N-terminal) et le dernier un atome C (bout C-terminal), il n'y a pas d'angle dièdre  $\phi$  associé au premier acide aminé, de même qu'il n'y a pas d'angle dièdre  $\psi$  associé au dernier acide aminé. Ces fluctuations sont d'un intérêt particulier, puisqu'elles ne représentent que des fluctuations internes, contrairement aux fluctuations des acides aminés qui peuvent inclure des mouvements de corps rigides.

Les fluctuations moyennes des angles  $\phi$  et  $\psi$  ont été calculées à partir des trajectoires pour les cinq simulations et les résultats sont présentés aux figures

28 et 29. Cette fois, les fluctuations sont concentrées dans trois régions distinctes qui tranchent du reste de la protéine, soit le site I, le segment 36-45 et le site II. Le reste de la protéine, c'est-à-dire les quatre hélices, fluctue autour de  $9^\circ$ . Les structures simulées A et A(S<sub>II</sub>) affichent globalement le plus de fluctuations, suivies par les structures simulées S<sub>I</sub>(D) et S<sub>II</sub>, et finalement la structure simulée D. La liaison du Ca<sup>2+</sup> est donc associée à une diminution graduelle des fluctuations internes de la protéine dans les sites de liaison et le segment 36-45. Plus en détails, pour les angles dièdres  $\phi$  la structure simulée A fluctue globalement plus que la structure simulée A(S<sub>II</sub>), sauf pour les acides aminés 16, 17, 18, 55, 58 et 59 qui fluctuent plus dans la structure simulée A(S<sub>II</sub>). Une différence notable entre les deux simulations est la plus grande stabilité du segment 36-45 dans la simulation A(S<sub>II</sub>). Le profil des fluctuations de la simulation A(S<sub>II</sub>) ressemble beaucoup à celui de la simulation S<sub>II</sub> dont elle est issue, cette dernière affichant cependant moins de fluctuations dans le site II (où un ion est présent), ainsi que dans le site I dans une certaine mesure. Pour la structure simulée S<sub>II</sub>, bien qu'un ion soit présent dans le site II, les dièdres des acides aminés 57 à 60 continuent de fluctuer.

Le site I de la structure simulée S<sub>I</sub>(D) fluctue moins que dans toutes les autres structures. Par contre, le dièdre  $\phi$  de l'acide aminé 59 (site II) fluctue en moyenne de  $133^\circ$ . Les dièdres du segment 36-45 fluctuent moins que dans la structure simulée D dont elle est issue, et cette dernière présente des sites I et II qui fluctuent respectivement légèrement plus et moins que dans les structures simulées S<sub>I</sub>(D) et S<sub>II</sub>. L'analyse des angles dièdres  $\psi$  mène qualitativement aux mêmes conclusions, bien que le comportement en détail soit quelque peu différent. Les différences notables par rapport aux fluctuations des dièdres  $\phi$  sont celles-ci: les niveaux des fluctuations du segment 36-45 des structures simulées A et A(S<sub>II</sub>) se ressemblent davantage, et le dièdre  $\psi$  de la structure simulée S<sub>I</sub>(D) fluctue plus que le dièdre  $\phi$ , ainsi que dans la structure simulée D où en particulier le dièdre  $\psi$  de l'acide aminé 43 fluctue en moyenne de  $68^\circ$ .

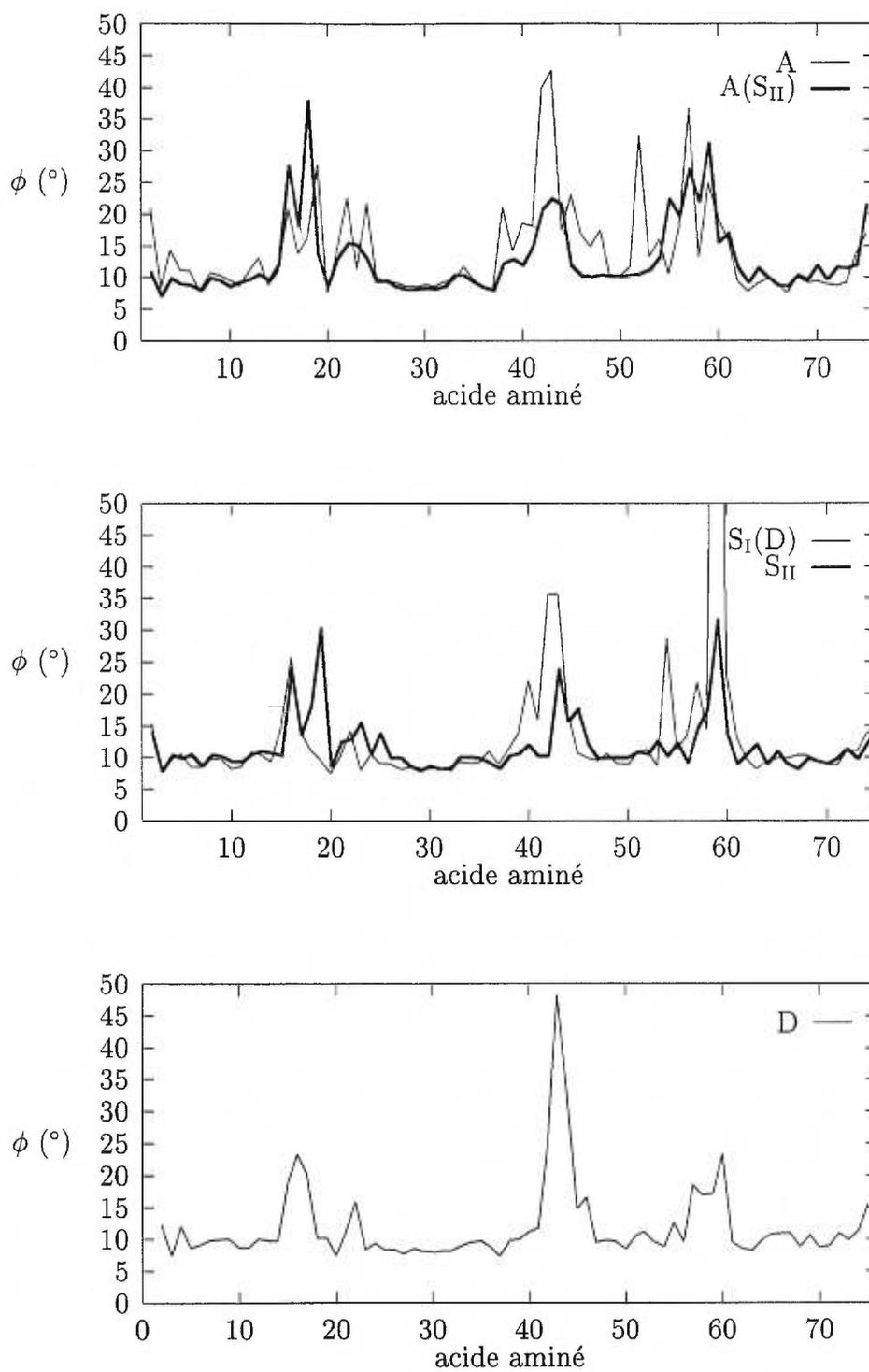


FIGURE 28. Fluctuations moyennes autour de la valeur moyenne de chaque angle dièdre  $\phi$ .

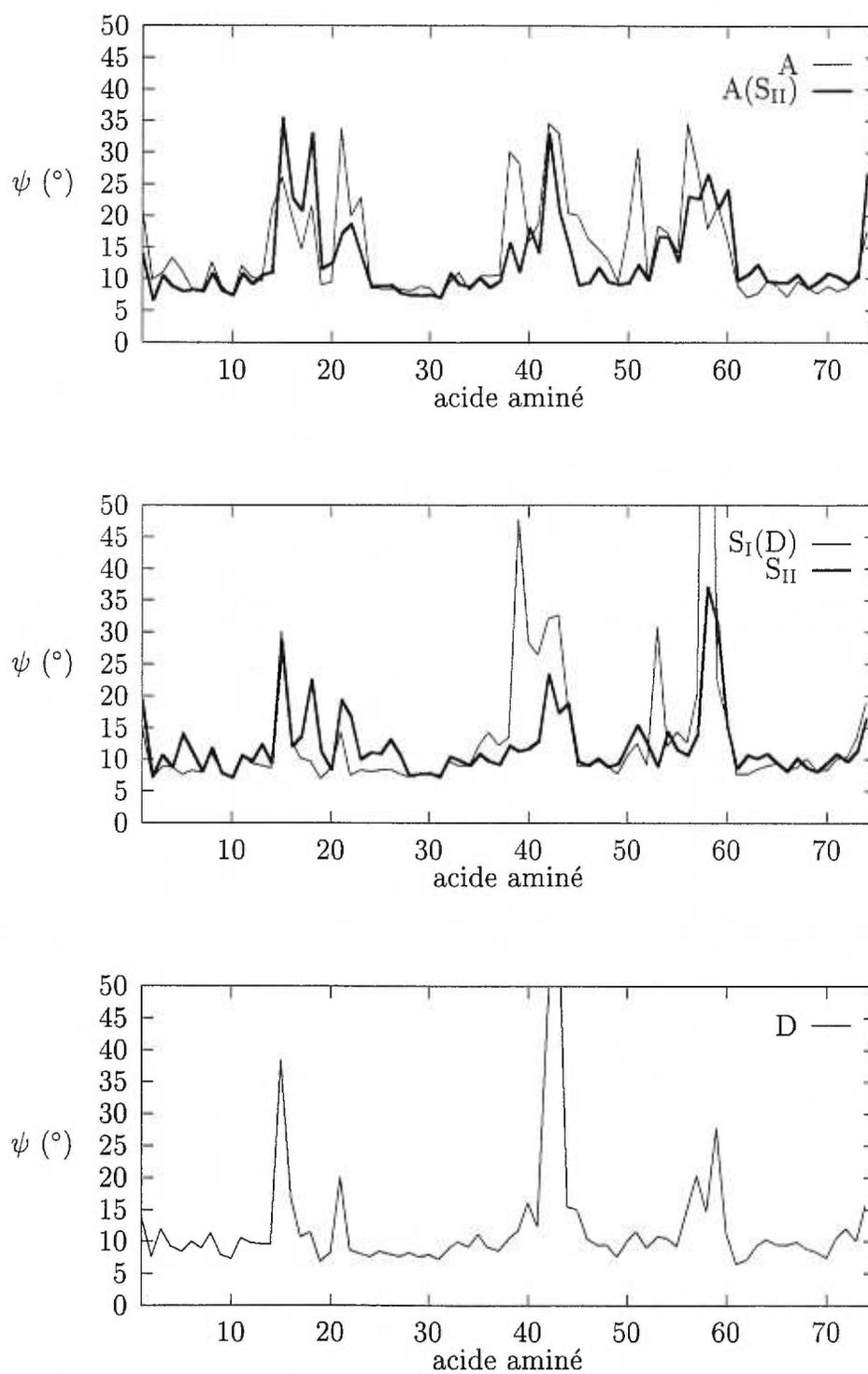


FIGURE 29. Fluctuations moyennes autour de la valeur moyenne de chaque angle dièdre  $\psi$ .

Bien qu'une tendance générale ressorte de ces comparaisons et que l'on puisse dire qualitativement que la liaison du  $\text{Ca}^{2+}$  aux sites de la calbindine  $\text{D}_{9k}$  diminue globalement les fluctuations internes dans le site II et le segment 36-45, il n'en demeure pas moins, en contrepartie, que certains acides aminés isolés subissent une hausse du niveau de leurs fluctuations du dièdre  $\psi$  et/ou  $\phi$ . Les fluctuations du site I, qui sont plus considérables dans la structure simulée D que dans la  $\text{S}_1(\text{D})$  surprennent un peu et demeurent pour l'instant sans explication, sinon que le comportement d'une structure simulée dépend de sa structure utilisée au départ. Il est possible que seul un nombre réduit de dièdres peuvent fluctuer librement avec davantage d'amplitude, puisqu'ils sont les seuls encore capables d'exprimer des fluctuations.

#### 4.3.2.3 Paramètres d'ordre $\langle S^2 \rangle$ (RMN) des vecteurs parallèles aux liens NH et $\text{C}_\alpha\text{H}_\alpha$ ; comparaisons entre les valeurs expérimentales et calculées

Le paramètre d'ordre est déterminé par des expériences de RMN et constitue une mesure de restriction spatiale des mouvements internes rapides d'un vecteur reliant deux noyaux, soit dans notre cas les vecteurs liens N-H et  $\text{C}_\alpha\text{H}_\alpha$ , à une échelle de temps de l'ordre des picosecondes.

Pour des spins nucléaires 1/2 (ex:  $^{15}\text{N}$  avec  $^1\text{H}$  ou  $^{13}\text{C}_\alpha$  avec  $^1\text{H}$ ), la relaxation suite à l'application d'un champ magnétique est dominée par l'interaction dipolaire avec les noyaux voisins et cette interaction spin-spin est tributaire des mouvements présents dans l'environnement de la paire de spins.<sup>116, 46</sup> L'étude de la relaxation spin-spin apporte donc de l'information au sujet de la dynamique du vecteur lien concerné à une échelle de temps des picosecondes, qui sont considérés comme des mouvements rapides.

Les paramètres de relaxation traditionnellement mesurés sont le temps de

relaxation longitudinale ( $T_1$ ), le temps de relaxation transverse ( $T_2$ ) et le NOE, c'est-à-dire le changement d'intensité du pic de résonance dû au transfert de magnétisation impliqué lors du phénomène de relaxation. Ces trois paramètres dépendent d'une densité spectrale  $J(\omega)$  donnée par

$$J(\omega) = 2 \int_0^{\infty} \cos(\omega t) C(t) dt \quad (4.5)$$

qui est la transformée de Fourier en cosinus de la fonction de corrélation  $C(t)$  de l'interaction spin-spin. Cette fonction de corrélation peut être factorisée en une partie reliée à la rotation diffusionnelle globale de la protéine  $C_r(t)$  et une partie reliée aux mouvements internes  $C_i(t)$ :

$$C(t) = C_r(t) C_i(t) \quad (4.6)$$

où

$$C_r(t) = \frac{1}{5} \exp(-t/\tau_r) \quad (4.7)$$

avec le temps caractéristique de relaxation rotationnelle  $\tau_r$ . La partie pour les mouvements internes seulement est donnée par:

$$C_i(t) = \langle P_2(\hat{e}(0) \cdot \hat{e}(t)) \rangle \quad (4.8)$$

où  $\hat{e}(0)$  est le vecteur unitaire reliant les deux noyaux pointant dans la direction du lien N-H ou  $C_\alpha H_\alpha$  au temps zéro, et  $\hat{e}(t)$  est le même vecteur à un temps  $t$  subséquent;  $P_2(x)$  est le polynôme de Legendre au second ordre,  $(3x^2 - 1)/2$ .

Comme cette densité spectrale est difficilement évaluable expérimentalement, une stratégie développée par Lipari et Szabo<sup>117</sup> consiste à modéliser la fonction de corrélation  $C_i(t)$  des mouvements internes par une autre fonction qui représente bien le comportement de la fonction de corrélation, notamment aux limites  $t \rightarrow 0$  et  $t \rightarrow \infty$ . Cette fonction de corrélation modèle est donnée par

$$C_i(t) = S^2 + (1 - S^2) \exp(-t/\tau_i) \quad (4.9)$$

dont la limite  $t \rightarrow \infty$  est  $S^2$ , le paramètre d'ordre généralisé, et  $\tau_i$  est le temps effectif de corrélation des mouvements internes. Le paramètre d'ordre varie de 0, dans le cas d'un mouvement isotrope, à 1 dans le cas d'un mouvement complètement restreint. Le paramètre d'ordre est donc une mesure de l'espace angulaire exploré par le vecteur colinéaire au lien entre les deux noyaux.

Dans notre cas, les paramètres d'ordre sont calculés à partir de la limite  $t \rightarrow \infty$  de l'expression analytique exacte de la fonction de corrélation, c'est-à-dire la moyenne du plateau de la fonction suivante (reliée à l'équation 4.8 par le théorème d'addition des harmoniques sphériques):

$$\begin{aligned} \langle S^2 \rangle &= \frac{4\pi}{5} \sum_{m=-2}^2 |\langle Y_{2m}(\Theta, \Phi) \rangle|^2 \\ &= \frac{3}{4} |\langle Y_2(\Theta, \Phi) \rangle|^2 + 3 |\langle Y_1(\Theta, \Phi) \rangle|^2 + \frac{1}{4} |\langle Y_0(\Theta, \Phi) \rangle|^2 \end{aligned} \quad (4.10)$$

où  $\langle \dots \rangle$  indique une moyenne d'ensemble. Les  $Y_{2m}(\Theta, \Phi)$  sont les harmoniques sphériques de rang 2 avec les angles  $\Theta$  et  $\Phi$  indiquant l'orientation du vecteur lien unitaire dans le cadre de référence de la molécule fixe. Cette expression est exacte; ce sont cependant les trajectoires générées par le modèle de simulation qui constituent l'approximation dans les calculs de paramètres d'ordre, puisqu'elles déterminent les trajectoires  $\Theta(t)$  et  $\Phi(t)$ .

La figure 30 compare les paramètres d'ordre des vecteurs N-H calculés à partir de la trajectoire à ceux obtenus expérimentalement.<sup>24, 109</sup> Rappelons qu'il n'y a pas de vecteur N-H pour les prolines, qui sont au nombre de quatre dans la calbindine D<sub>9k</sub>, soit les acides aminés 1, 3, 20 et 37. De plus, les paramètres d'ordre expérimentaux des acides aminés 15, 17, 23 et 24 de l'état S<sub>II</sub> n'ont pu être déterminés. Les paramètres d'ordre expérimentaux de l'état S<sub>I</sub>(D) n'étaient pas encore connus au moment de la rédaction de cette thèse.

La figure 30 montre qu'il existe un bon accord qualitatif entre nos résultats

et les résultats expérimentaux. Les paramètres d'ordre calculés sont généralement supérieurs aux valeurs expérimentales pour les hélices et ce niveau augmente avec la venue du  $\text{Ca}^{2+}$  dans un site. Brüshweiler<sup>118</sup> a montré que les paramètres d'ordre associés aux mouvements très rapides calculés à partir de trajectoires générées par des méthodes classiques étaient surestimés par  $\approx 7-9\%$  par rapport à des calculs quantiques. Mais les paramètres d'ordre sont par contre inférieurs lorsqu'il s'agit des sites, particulièrement pour le site I. Remarquons aussi que dans la région des hélices, on observe une alternance (comme aux figures 24, 25 et 26) de la valeur du paramètre d'ordre en fonction du numéro d'acide aminé de faible amplitude, et comparable aux données expérimentales. Pour les sites, ces alternances sont d'amplitude beaucoup plus grande, ce qui donne pour certains acides aminés des paramètres d'ordre beaucoup plus faibles que les valeurs expérimentales.

Pour la structure simulée A, les paramètres d'ordre des vecteurs N-H des acides aminés 15-22, 38-46, 52, 55 et 57 sont considérablement inférieurs aux données expérimentales; la structure simulée A( $S_{II}$ ) se comporte notablement mieux avec seulement les acides aminés 16-19 présentant des paramètres d'ordre inférieurs aux valeurs expérimentales et les acides aminés 40-43 présentant le contraire.

Pour la structure simulée  $S_{II}$ , on peut supposer que les paramètres d'ordre des vecteurs N-H des acides aminés 16 et 19 sont inférieurs à ce que serait la valeur expérimentale (qui ne sont pas connues), tandis que ceux des acides aminés 22, 59 et 60 sont inférieurs aux valeurs expérimentales, et que ceux des acides aminés 40-43 sont supérieurs aux valeurs expérimentales. Bien que les valeurs expérimentales pour l'état  $S_I(D)$  ne soient pas connues, il est fort probable que les paramètres d'ordre des acides aminés 16, 40-42, 55 et 60 seraient inférieurs aux valeurs expérimentales. Mais on sait, par des expériences de mesures du taux d'échange des protons appartenant à des groupements amides du modèle N56A de l'état  $S_I(D)$ , que sa dynamique est presque identique à celle de l'état D.<sup>114</sup>

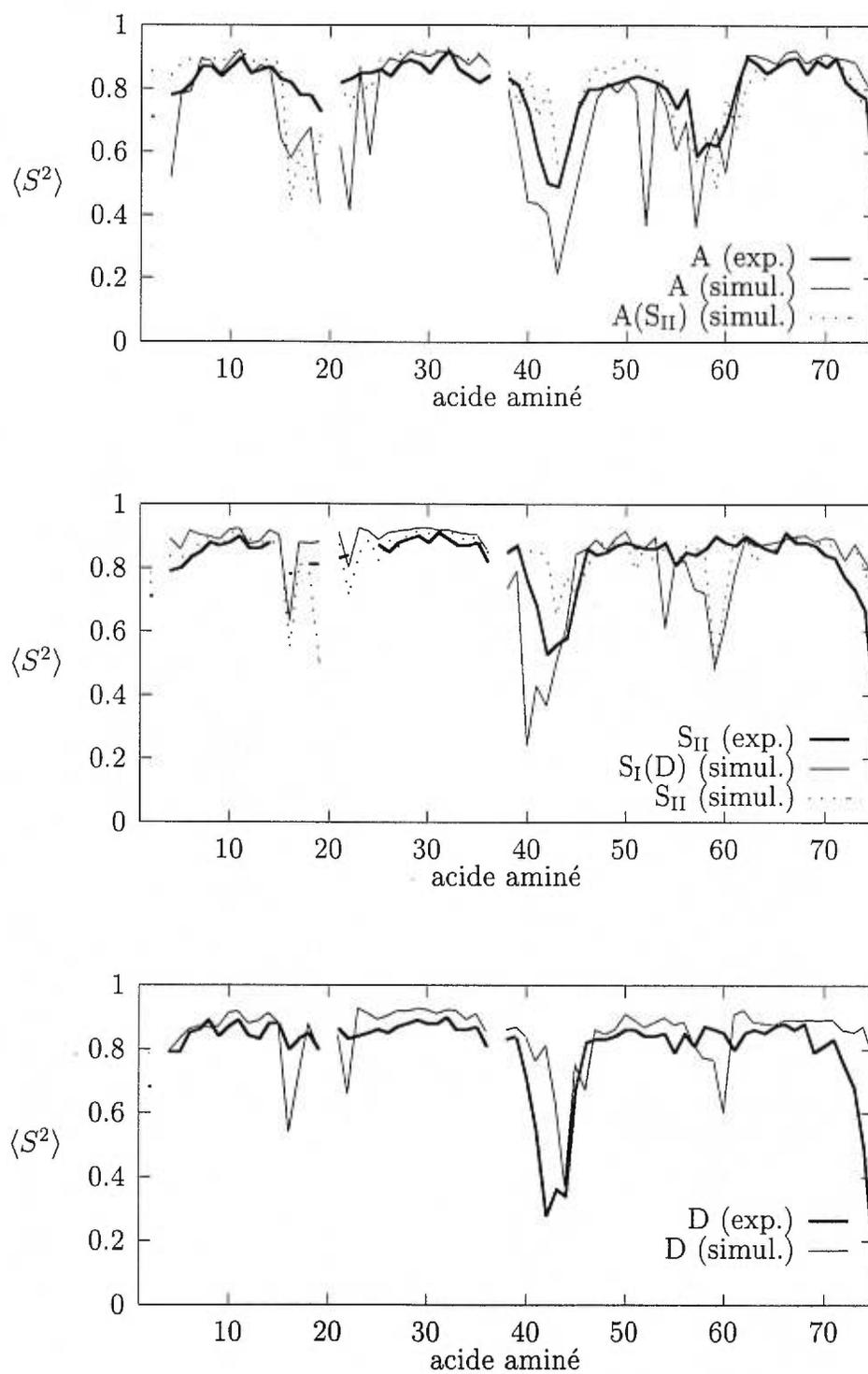


FIGURE 30. Comparaisons des paramètres d'ordre ( $\langle S^2 \rangle$ ) des vecteurs N-H provenant de données expérimentales avec ceux calculés à partir des trajectoires.

La structure simulée D est celle dont les valeurs calculées sont les plus en accord avec les valeurs expérimentales. Les paramètres d'ordre des vecteurs N-H des acides aminés 16, 17, 22 et 58-60 sont inférieurs aux valeurs expérimentales, tandis que ceux des acides aminés 40-42 leur sont supérieurs. Mais un accord qualitatif demeure, et il ressort que l'état A est le plus flexible, suivi par les états  $S_{II}$  et  $S_I(D)$ , et finalement l'état D. La liaison du  $Ca^{2+}$  dans un site affecte donc la flexibilité des deux sites, mais aussi du segment 36-45 qui devient plus flexible selon les données expérimentales, mais qui se rigidifie selon nos calculs.

Il est intéressant de relever que le profil des paramètres d'ordre en fonction des acides aminés affiche une correspondance presque parfaite avec les fluctuations des angles dièdres  $\phi$  ou  $\psi$ . En fait, si on tente de superposer les paramètres d'ordre N-H aux figures 28 ou 29 (on doit inverser la feuille) on voit que les acides aminés ayant des paramètres d'ordre trop faibles sont ceux dont les angles dièdres fluctuent de façon singulière et excessive. Ce thème est repris et analysé en détail dans le chapitre suivant.

Le tableau XVI donne les valeurs moyennes des paramètres d'ordre des vecteurs N-H par élément de structure secondaire. Les mêmes acides aminés que dans les moyennes expérimentales<sup>24</sup> ont été retenus pour effectuer les moyennes des valeurs provenant des simulations.

Ces résultats suggèrent que les hélices sont très bien représentées, puisque leurs valeurs moyennes rejoignent celles des valeurs expérimentales sauf pour quelques exceptions légèrement supérieures. Ces moyennes suggèrent aussi que les sites sont bien représentés, sauf pour le site I des structures simulées A et A( $S_{II}$ ). Il faut cependant remarquer que l'écart type dans le cas des sites est plus élevé que pour les hélices, ce qui rappelle qu'il y a de plus grandes fluctuations des valeurs des paramètres d'ordre dans les sites.

Les valeurs expérimentales des paramètres d'ordre des vecteurs  $C_\alpha H_\alpha$  n'étant

		acides aminés					
		4-14	16, 18-22	25-35	46-54	55-62	63-73
		hélice A	site I	hélice B	hélice C	site II	hélice D
A	exp.	$0.85 \pm 0.01$	$0.80 \pm 0.02$	$0.87 \pm 0.01$	$0.82 \pm 0.01$	$0.72 \pm 0.04$	$0.86 \pm 0.01$
	simul.	$0.83 \pm 0.10$	$0.55 \pm 0.10$	$0.90 \pm 0.02$	$0.73 \pm 0.14$	$0.64 \pm 0.15$	$0.90 \pm 0.01$
A(S <sub>II</sub> )simul.		$0.90 \pm 0.02$	$0.63 \pm 0.15$	$0.91 \pm 0.01$	$0.86 \pm 0.03$	$0.66 \pm 0.10$	$0.87 \pm 0.03$
S <sub>II</sub>	exp.	$0.85 \pm 0.01$	$0.82 \pm 0.01$	$0.880 \pm 0.004$	$0.863 \pm 0.005$	$0.86 \pm 0.01$	$0.84 \pm 0.02$
	simul.	$0.87 \pm 0.03$	$0.74 \pm 0.14$	$0.91 \pm 0.03$	$0.73 \pm 0.04$	$0.85 \pm 0.13$	$0.86 \pm 0.02$
D	exp.	$0.84 \pm 0.01$	$0.82 \pm 0.02$	$0.862 \pm 0.004$	$0.838 \pm 0.005$	$0.83 \pm 0.01$	$0.82 \pm 0.02$
	simul.	$0.87 \pm 0.04$	$0.75 \pm 0.13$	$0.91 \pm 0.01$	$0.86 \pm 0.07$	$0.82 \pm 0.10$	$0.88 \pm 0.01$

TABLEAU XVI. Comparaisons des valeurs moyennes (par élément de structure secondaire) des paramètres d'ordre des vecteurs N-H expérimentales<sup>24</sup> avec celles calculées à partir des trajectoires.

pas connues, il n'est pas possible d'effectuer une analyse aussi exhaustive que dans le cas des vecteurs N-H. Néanmoins, on s'attend à ce que ces paramètres d'ordre ( $C_\alpha H_\alpha$ ) soient globalement supérieurs aux paramètres d'ordre des vecteurs N-H, ceci parce que le plan peptidique O-C-N-H forme un plan plutôt rigide avec de part et d'autre une liberté de rotation autour des liaisons simples avec les  $C_\alpha$  adjacents. Il n'est pas possible au vecteur  $C_\alpha H_\alpha$  de se déplacer aussi librement que le vecteur N-H. La figure 31 nous montre pourtant des résultats qui s'accordent avec les résultats obtenus pour les vecteurs N-H. Certains acides aminés des sites semblent toujours problématiques lors des simulations, mais il n'est pas possible de vérifier dans quelle mesure. Notons que contrairement aux résultats expérimentaux des vecteurs N-H, où la flexibilité du segment 36-45 s'accroît avec le niveau d'occupation du  $Ca^{2+}$ , le segment 36-45 semble se rigidifier avec la liaison du  $Ca^{2+}$  selon les paramètres d'ordre des vecteurs  $C_\alpha H_\alpha$ .

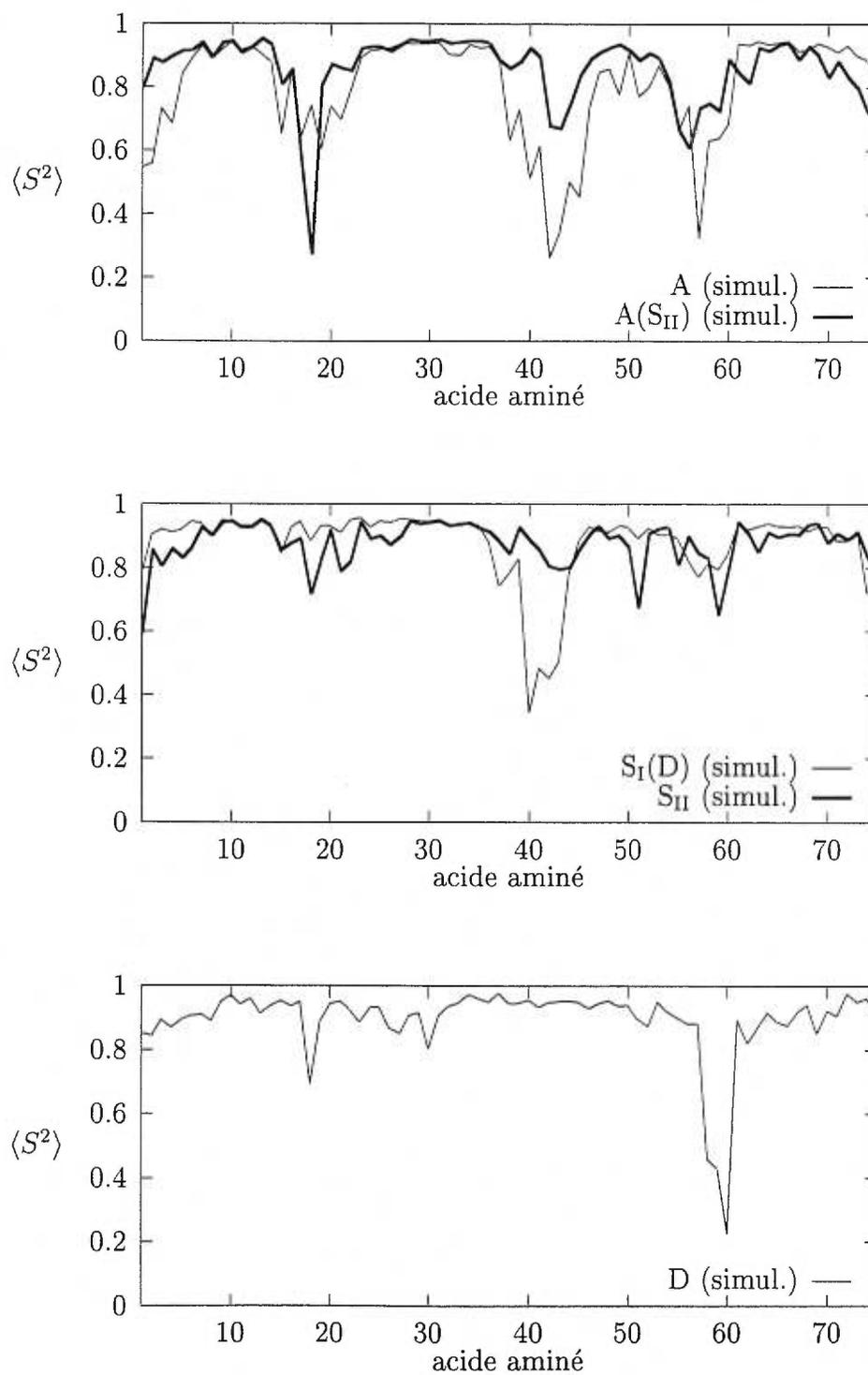


FIGURE 31. Paramètres d'ordre ( $\langle S^2 \rangle$ ) des vecteurs  $C_\alpha H_\alpha$  calculés à partir des trajectoires.

Des comparaisons similaires entre paramètres d'ordre mesurés expérimentalement, et calculés à partir de trajectoires ont déjà été faites.<sup>109, 119, 120, 121</sup> En particulier, l'étude effectuée par Smith et coll.<sup>120</sup> a révélé que les paramètres d'ordre calculés d'une trajectoire du lysozyme du blanc d'oeuf de poule affichait une grande variation d'amplitude en fonction du numéro d'acide aminé, certains ayant un paramètre d'ordre considérablement plus bas que leurs voisins directs. Ceci étant aussi notre cas, il est intéressant de relever leurs explications qui trouvent leur origine dans les fluctuations trop grandes des angles dièdres  $\phi$  et  $\psi$ , soit des transitions concertées pouvant être la conséquence des limites de fiabilité du champ de forces. Il est aussi possible que des positions atomiques déterminées avec moins de certitude dans la structure expérimentale de départ puissent mener à des instabilités locales transitoires lors de la période d'équilibration, se stabilisant enfin dans un minimum local avec moins d'éléments de structure secondaires stabilisateurs comme des ponts hydrogènes. Une conséquence serait une flexibilité locale accrue.

### 4.3.3 Contributions structurales et énergétiques à la coopérativité

#### 4.3.3.1 Contribution électrostatique défavorable à la coopérativité associée à la liaison du $\text{Ca}^{2+}$

La différence d'énergie libre associée à la coopérativité chez la calbindine  $\text{D}_{9k}$  pour le cas où le  $\text{Ca}^{2+}$  se lie d'abord au site II, soit (voir figure 6)

$$\Delta\Delta G^{\text{coop.}} = \Delta G_{\text{I,II}} - \Delta G_{\text{II}} < 0 \quad (4.11)$$

peut être divisée en deux contributions:<sup>46</sup> une première provenant des modifications de structure et de dynamique de la protéine suite à la liaison de  $\text{Ca}^{2+}$ ; une deuxième associée à la liaison du  $\text{Ca}^{2+}$  dans le site préformé, et qui est de nature électrostatique, soit

$$\Delta\Delta G^{\text{coop.}} = \Delta\Delta G^{\text{prot.}} + \Delta\Delta G^{\text{ion}} \quad (4.12)$$

où

$$\Delta\Delta G^{\text{prot.}} = [G_{\text{I,II}}^{\text{prot.}} - G_{\text{II}}^{\text{prot.}}] - [G_{\text{II}}^{\text{prot.}} - G_{\text{A}}^{\text{prot.}}] \quad (4.13)$$

et

$$\Delta\Delta G^{\text{ion}} = [G_{\text{I,II}}^{\text{ion}} - G_{\text{II}}^{\text{ion}}] - [G_{\text{I}}^{\text{ion}} - G_{\text{A}}^{\text{ion}}]. \quad (4.14)$$

La contribution  $\Delta\Delta G^{\text{ion}}$  correspond à la différence d'énergie libre associée à l'interaction (répulsion) des deux  $\text{Ca}^{2+}$  lorsque chacun se trouve dans son site respectif. On présume ici que l'interaction entre le  $\text{Ca}^{2+}$  et la protéine est identique dans les états  $S_{\text{I(D)}}$ ,  $S_{\text{II}}$  et  $D$ .

Dans l'état  $D$ , les  $\text{Ca}^{2+}$  sont séparés en moyenne par  $12 \text{ \AA}$ , et la calbindine  $D_{9k}$  sans  $\text{Ca}^{2+}$  possède une charge totale de  $-7e$  dans sa forme A mineure (sans le groupement N-terminal Ac-Ser-Ala-Lys), répartie de façon asymétrique entre le site I ( $-2e$ ), le site II ( $-4e$ ) et le reste de la protéine. Lorsqu'un premier  $\text{Ca}^{2+}$  se lie au site II de la calbindine  $D_{9k}$ , la charge résultante de la calbindine  $D_{9k}$  dans son état  $S_{\text{II}}$  est diminuée; la contribution électrostatique à la différence d'énergie libre de la coopérativité apparaît donc défavorable au processus global.

L'énergie d'interaction entre les deux  $\text{Ca}^{2+}$  dans les sites de la calbindine  $D_{9k}$  a été évaluée en plaçant la protéine dans un milieu diélectrique continu homogène afin de modéliser l'eau, et en résolvant numériquement l'équation de Poisson pour un diélectrique inhomogène, c'est-à-dire

$$\nabla \cdot (\epsilon \nabla V) = -\frac{\rho}{\epsilon_0} \quad (4.15)$$

où  $\epsilon_0$  est la permittivité du vide et  $\epsilon$  est la permittivité relative du milieu continu, égale ici à 80;  $V$  et  $\rho$  sont le potentiel électrostatique et la densité de charge en un point donné. A l'aide d'une grille de points auxquels sont calculées les densités de charge, l'énergie d'interaction est ensuite calculée *via* l'équation de Poisson, en évaluant tous les termes de l'équation 4.14. Le résultat de cette méthode

suggère que l'interaction des deux  $\text{Ca}^{2+}$  dans les sites de la protéine est d'environ  $+3.0$  kcal/mol, donc défavorable.

Cette contribution électrostatique défavorable à la coopérativité doit être contrebalancée jusqu'au point de rendre le processus global favorable. Ainsi, sachant que l'énergie libre de coopérativité est de  $-1.84$  kcal/mol à une concentration de 50 mM/l de KCl,<sup>17</sup> la contribution associée à la modification structurale et dynamique de la protéine ( $\Delta\Delta G_{\text{prot}}$ ) se doit d'être au moins de  $-4.84$  kcal/mol.

#### **4.3.3.2 Contribution des interactions hydrophobes: interactions de van der Waals entre la chaîne latérale de la Val61 et le reste de la protéine; distributions des énergies et valeurs moyennes**

Puisque l'on sait que la protéine ne subit pas de changement conformationnel majeur suite à la liaison du  $\text{Ca}^{2+}$ , la contribution structurale à la coopérativité doit résulter de réarrangements subtils de la structure. Afin d'illustrer cela, les trajectoires des cinq simulations ont été utilisées pour calculer les distributions des énergies d'interactions de van der Waals entre la chaîne latérale de la Val61 avec le reste de la protéine. Cet acide aminé a été sélectionné pour son plus grand nombre de contacts avec d'autres atomes de la protéine et aussi parce qu'il existe des données expérimentales de changement de coopérativité associée à cette mutation.<sup>122</sup> La figure 32 montre qu'il y a un chevauchement important entre les différentes distributions, et que l'évaluation de ces énergies dans seulement une ou quelques structures n'aurait pas suffi à faire ressortir ce résultat. Les distributions montrent que cette énergie d'interaction est de plus en plus favorable dans l'ordre suivant des structures simulées: A, A(S<sub>II</sub>), S<sub>I</sub>(D), S<sub>II</sub> et D, les moyennes des distributions étant respectivement de  $-7.34$ ,  $-7.68$ ,  $-8.09$ ,  $-9.36$  et  $-9.72$  kcal/mol.

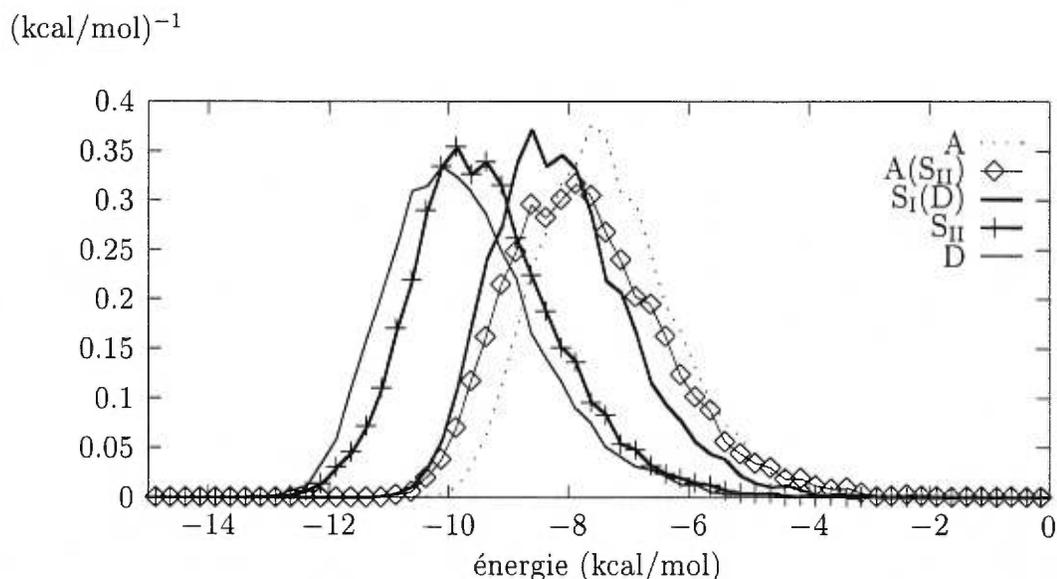


FIGURE 32. Distributions des énergies d'interaction de van der Waals entre la chaîne latérale de la Val61 et le reste de la protéine.

#### 4.3.3.3 Distributions des distances des deux ponts hydrogènes du feuillet- $\beta$ reliant les deux sites

Les deux ponts hydrogènes du petit feuillet- $\beta$  qui relie la Leu23 (site I) à la Val61 (site II) ont déjà été reconnus comme de bons candidats pour jouer un rôle dans l'interaction coopérative qui a lieu entre les deux sites.<sup>25</sup> Aussi, ces deux ponts hydrogènes sont présents dans toutes les structures des différents états d'occupation déterminées expérimentalement. La distribution radiale de la longueur de chacun de ces deux ponts hydrogènes, soit Leu23-O...HN-Val61 et Leu23-HN...O-Val61 a été calculée à partir des trajectoires pour les cinq simulations.

La figure 33 montre que les deux ponts hydrogènes sont présents dans toutes les structures simulées, sauf la structure simulée A. Dans cette dernière, le pont hydrogène Leu23-HN...O-Val61 est plus relâché que dans les autres structures, tandis que le pont Leu23-O...HN-Val61 n'est pas présent, chacun des deux parties

étant lié à une même molécule d'eau. Il semble aussi que le pont Leu23-HN...O-Val61 soit plus stable d'une structure à l'autre que le pont Leu23-O...HN-Val61. Pour ce dernier, les données de simulations indiquent un resserrement en fonction de l'état d'occupation du  $\text{Ca}^{2+}$ . Les structures simulées A( $S_{\text{II}}$ ) et  $S_{\text{I}}$ (D) ont une distribution très similaire, tandis que celle de la structure simulée  $S_{\text{II}}$  s'apparente plutôt à la D.

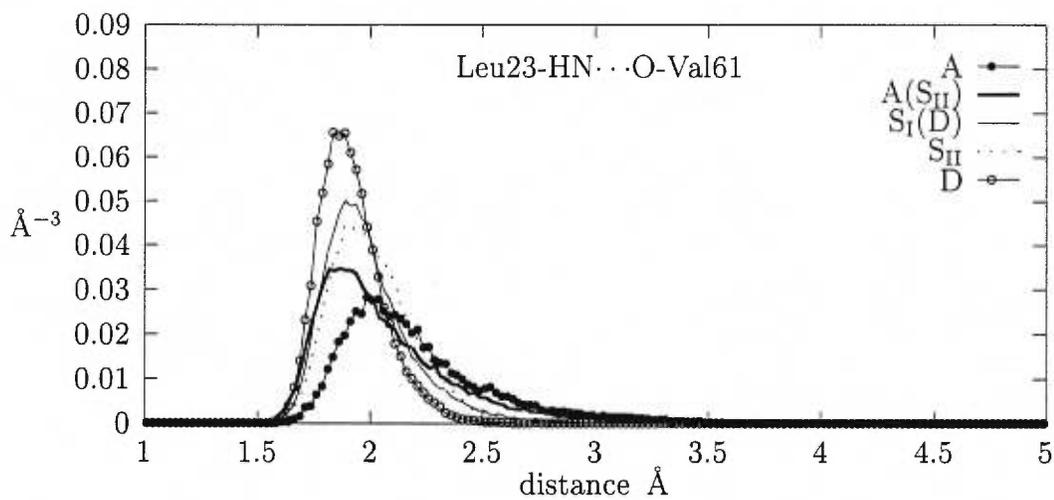
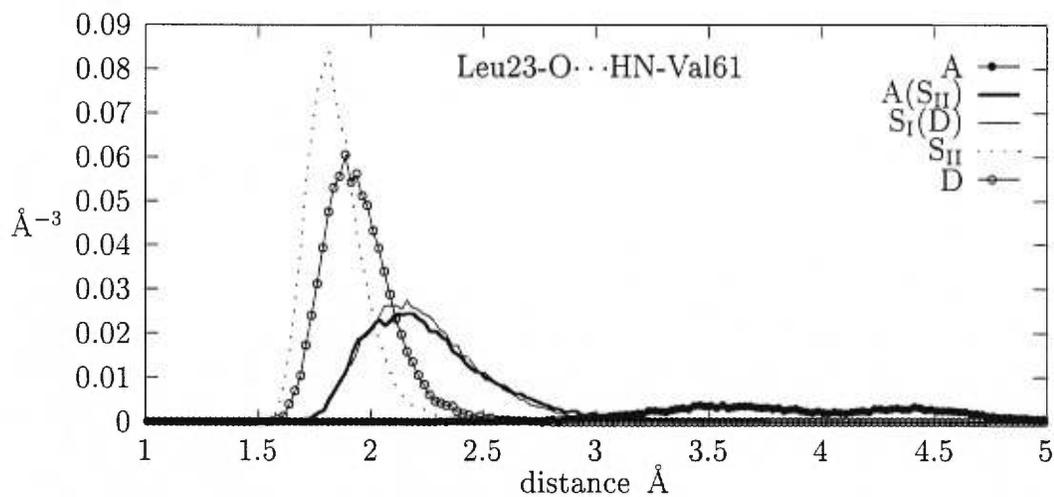


FIGURE 33. Distributions radiales des longueurs des ponts hydrogènes entre les acides aminés Leu23 et Val61 qui font partie respectivement du site I et II.

#### 4.4 Synthèse et conclusion

Une simulation de DM de 200 ps a été produite pour chacun des quatre états d'occupation de la calbindine  $D_{9k}$ , avec une simulation supplémentaire de l'état A pour un total de cinq trajectoires produites (A, A( $S_{II}$ ),  $S_I(D)$ ,  $S_{II}$  et D).

Lors des simulations, les structures simulées  $S_{II}$  et D sont demeurées structurellement proches de leur structure expérimentale de départ, tandis que la structure simulée A s'en est éloignée plus considérablement. La structure simulée A( $S_{II}$ ), construite à partir de la structure expérimentale  $S_{II}$  d'où on a enlevé le  $Ca^{2+}$  demeure davantage proche de la structure expérimentale A.

Les différences structurales entre structures simulées et expérimentales se concentrent surtout dans les bouts terminaux, les sites ainsi que dans le segment 36-45 qui relie les deux mains EF. Les hélices conservent en général un très bon accord avec leur structure expérimentale, sauf pour l'hélice C de la structure simulée A qui dévie le plus, mais dont la position s'avère mal définie dans les données expérimentales RMN.<sup>32</sup>

Suite à ces résultats, les structures simulées A( $S_{II}$ ),  $S_I(D)$ ,  $S_{II}$  et D ont été comparées entre elles. Les différences de chaîne principale entre les structures simulées  $S_I(D)$  et  $S_{II}$  sont les plus importantes, suivie par la structure simulée A( $S_{II}$ ) qui diffère considérablement de toutes les autres. Les deux structures simulées les plus similaires sont les  $S_I(D)$  et D, suivies de près par les  $S_{II}$  et D. La structure simulée  $S_I(D)$  ayant été construite à partir de la structure expérimentale D, il n'est peut-être pas très significatif que ces deux structures soient les plus similaires. Une simulation de DM plus longue serait nécessaire afin d'observer l'évolution de la structure simulée  $S_I(D)$ . Mais des résultats expérimentaux appuient pourtant cette donnée puisqu'ils ont montré que la structure de l'état  $S_I(D)$  était presque identique à celle de l'état D.<sup>114</sup> Les différences de structure entre états simulés se concentrent surtout dans les sites et le segment 36-45. Il ne faut pas

non plus perdre de vue que ces régions de la protéine sont aussi celles qui sont déterminées expérimentalement avec le moins de certitude.<sup>32, 33</sup> Néanmoins, les résultats des simulations s'accordent qualitativement avec les résultats expérimentaux, c'est-à-dire que les états liés au  $\text{Ca}^{2+}$  sont structurellement similaires entre eux et différents de l'état apo, signifiant que la majorité des changements structuraux se produisent lors de la liaison du premier  $\text{Ca}^{2+}$ . La différence considérable entre les structures simulées  $S_I(D)$  et  $S_{II}$  représente peut-être l'asymétrie du mécanisme de la coopérativité selon que l'ion se lie d'abord au site I ou au site II.

La coordinence du  $\text{Ca}^{2+}$  est altérée dans deux cas, soit dans le site I de la structure simulée D où il y a une molécule d'eau supplémentaire, ainsi que dans le site II de la structure simulée  $S_{II}$  où l'acide aminé Asp54 passe de *mono* à *bidenté*, tout comme lors de la simulation de Ahlström et coll.<sup>108</sup> Les distances  $\text{Ca}^{2+}$ -ligands sont en général trop courtes par rapport aux valeurs expérimentales, sauf pour le cas de l'Asp19 et le Gln22 de la structure simulée D dont les atomes d'oxygène se retrouvent plus distants que dans les expériences, ainsi que celui du Gln22 du site I de la structure simulée  $S_I(D)$ .

Le réseau de ponts hydrogènes HN-O de la chaîne principale de la structure simulée A est celui qui a été le plus modifié lors de la simulation, contrairement à la structure simulée  $A(S_{II})$  qui retient tous les ponts hydrogènes présents dans la structure expérimentale A. Les autres structures simulées perdent au plus chacune trois ponts hydrogènes de la chaîne principale qui sont pour la plupart remplacés par des ponts hydrogènes avec des chaînes latérales ou des molécules d'eau. Les ponts hydrogènes perdus faisaient tous partie d'un des deux sites.

Il s'avère aussi que les structures simulées  $A(S_{II})$ ,  $S_I(D)$  et  $S_{II}$  se sont décontractées lors de la simulation par rapport aux structures expérimentales. La structure simulée D demeure en excellent accord avec la structure expérimentale qui provient d'expériences de diffractions des rayons X à haute résolution. Si les structures simulées semblent indiquer qu'il y a contraction de la protéine avec le

niveau d'occupation du  $\text{Ca}^{2+}$ , les données expérimentales ne sont pas suffisantes pour confirmer ou infirmer ce résultat, ces premières n'existant que pour quelques structures.

Les structures simulées ont un profil de solvation très similaire pour les chaînes principale et latérales, bien que quelques petites différences notables existent. La liaison du  $\text{Ca}^{2+}$  dans le site II des structures simulées  $S_{II}$  et D déshydrate les chaînes latérales du site II et leur chaîne principale dans une certaine mesure, mais il semble qu'un certain effet soit ressenti dans le site I et consiste à déshydrater d'environ une molécule  $\text{H}_2\text{O}$  la chaîne principale du site I encore inoccupé. L'inverse n'est cependant pas observé, c'est-à-dire que la liaison du  $\text{Ca}^{2+}$  au site I ne semble pas affecter l'hydratation des chaînes principales et latérales du site II inoccupé.

L'examen des fluctuations des positions des acides aminés a permis de vérifier que les bouts terminaux, les sites et le segment 36-45 sont les éléments qui fluctuaient le plus, et que ces fluctuations diminuent avec le niveau d'occupation du  $\text{Ca}^{2+}$  dans la protéine. Pour la structure simulée D, les fluctuations des atomes lourds ont été comparées aux valeurs expérimentales homologues obtenues à partir des facteurs  $B$  de température (ou de Debye-Waller) et un bon accord général est ressorti. Les fluctuations des angles dièdres  $\phi$  et  $\psi$  ont mené aux mêmes conclusions que les résultats des fluctuations des acides aminés, et ont aussi révélé que certains acides aminés continuent de fluctuer considérablement même lorsqu'un ion est présent dans le site.

Les paramètres d'ordre expérimentaux des vecteurs N-H ont été comparés aux valeurs calculées à partir des trajectoires. La tendance qualitative est raisonnablement bien respectée, c'est-à-dire que la liaison successive du  $\text{Ca}^{2+}$  rend la structure plus rigide, avec la majeure partie des effets concentrés dans la liaison d'un premier ion au site I ou au site II. Mais il ressort que certains acides aminés dans les sites ont un paramètre d'ordre très faible même en présence d'un  $\text{Ca}^{2+}$ .

De plus, ces acides aminés ayant un paramètre d'ordre trop faible par rapport à la valeur expérimentale sont aussi ceux dont les angles dièdres  $\phi$  et  $\psi$  fluctuent le plus en toutes circonstances.

Une mesure de la contribution électrostatique à la coopérativité indique que cette dernière est bien défavorable par  $+3.0$  kcal/mol au processus global de coopérativité ( $-1.84$  kcal/mol), suggérant que la contribution du réarrangement de structure et de dynamique de la protéine devrait être d'environ  $-4.84$  kcal/mol.

Les interactions hydrophobes pourraient jouer un rôle important dans la coopérativité, tel que suggéré par les distributions des énergies d'interactions entre la chaîne latérale de la Val61 et le reste de la protéine pour laquelle on retrouve une interaction de plus en plus favorable avec la liaison de  $\text{Ca}^{2+}$ . Les distributions des ponts hydrogènes du feuillet- $\beta$  entre les deux sites ont révélé que les deux ponts sont présents dans toutes les structures simulées sauf A, et qu'il y avait resserrement du pont hydrogène Leu23-HN $\cdots$ O-Val61 suite à la liaison du  $\text{Ca}^{2+}$ .

Tous ces résultats indiquent que le comportement des simulations diffère quantitativement des résultats expérimentaux presque exclusivement pour les sites. Le segment 36-45 étant très flexible et mal déterminé, les différences structurales sont moins surprenantes. Que ce soit pour les différences structurales ou dynamiques, ce sont toujours les mêmes acides aminés qui éprouvent des difficultés à reproduire les résultats expérimentaux. L'origine de ces déviations peut provenir des imperfections de la fonction de potentiel de CHARMM,<sup>56</sup> ou d'un échantillonnage insuffisant de l'espace configurationnel. Par exemple, il est possible que le manque de polarisation induite ait comme conséquence que les atomes d'oxygène ligands du  $\text{Ca}^{2+}$  se positionnent trop près de ce dernier, et que ces repositionnements engendrent le bris de ponts hydrogènes dans la structure des sites. En second lieu, il se peut que cette modification de structure influe aussi sur la dynamique de certains ligands immédiats du  $\text{Ca}^{2+}$ . Mais cet argument n'a pas de sens dans le

cas de la structure simulée A(S<sub>II</sub>). Il a aussi été proposé que les trop grandes fluctuations des angles dièdres soient permises par les imperfections du champ de force (barrière d'énergie trop basse).<sup>120</sup> Il est également possible que des incertitudes dans les structures de départ contribuent à déformer les structures et modifier la dynamique.

Malgré tout, un bon accord qualitatif et dans une certaine mesure quantitatif existe entre les résultats des simulations et des expériences. Les modèles testés dans ce chapitre ont donc été retenus pour les études subséquentes. Dans le chapitre 5 qui suit, la correspondance directe entre les fluctuations des angles dièdres  $\phi$  et  $\psi$  et les paramètres d'ordre des vecteurs N-H est relevée. Une analyse exhaustive des paramètres d'ordre est effectuée en termes des fluctuations des angles dièdres  $\phi$  et  $\psi$ , puisqu'ils représentent aussi une mesure des fluctuations internes. On y développera un modèle analytique afin d'interpréter les paramètres d'ordre en termes de mouvements particuliers accomplis par les vecteurs N-H.

## CHAPITRE 5

### Développement d'une représentation analytique de la partie librationnelle des paramètres d'ordre des vecteurs N-H de la chaîne principale de la calbindine $D_{9k}$

#### 5.1 Introduction

Une brève introduction aux mesures de relaxation RMN a été exposée au chapitre 4. En particulier, les paramètres d'ordre mesurés expérimentalement ont été comparés aux valeurs calculées à partir des trajectoires de DM. Bien que le paramètre d'ordre donne une mesure de l'étendue des mouvements du vecteur N-H ainsi qu'un intervalle de temps dans lequel se produit le phénomène de relaxation, il ne fournit pas d'information quant à la nature des mouvements effectués. Ceci d'abord parce que les données expérimentales sont des moyennes et aussi à cause du couplage entre les différents types de mouvements qui existent dans la protéine et qui déterminent les différentes orientations adoptées par le vecteur N-H.<sup>123</sup>

Lors du phénomène de relaxation d'un vecteur N-H, la fonction de corrélation totale de la réorientation spin-spin  $C(t)$ , peut être factorisée en une partie reliée à la rotation globale de la protéine et une partie reliée aux mouvements internes:

$$C(t) = C_r(t) \cdot C_i(t) \quad (5.1)$$

où

$$C_r(t) = \frac{1}{5} \exp(-t/\tau_r) \quad (5.2)$$

est la partie reliée à la rotation globale avec le temps caractéristique de relaxation rotationnelle  $\tau_r$ ; l'expression exacte de la partie reliée aux mouvements internes  $C_i(t)$  a été donnée à l'équation 4.8.

Un premier modèle d'interprétation des données de relaxation ( $T_1$ ,  $T_2$  et NOE) a été proposé par Lipari et Szabo.<sup>117</sup> Cette approche est fondée sur l'approximation que l'on peut découpler la rotation globale de la protéine de ses mouvements internes. Dans ce contexte du modèle de Lipari et Szabo, les mouvements internes sont décrits par deux paramètres: le paramètre d'ordre  $S^2$  et le temps caractéristique de relaxation  $\tau_i$  qui ne rend compte que d'un seul type de mouvements internes dont la nature n'est pas précisée. Ainsi, on fait l'hypothèse que la fonction de corrélation  $C_i(t)$  a la forme:

$$C_i(t) = S^2 + (1 - S^2) \exp(-t/\tau_i). \quad (5.3)$$

Suite à des études sur l'interleukine-1 $\beta$  par Clore et coll.<sup>124</sup> il s'est avéré que les paramètres de relaxation  $T_1$ ,  $T_2$  et NOE des noyaux  $^{15}\text{N}$  de la chaîne principale ne pouvaient être bien ajustés au modèle de Lipari-Szabo ne comportant que deux paramètres et ne rendant compte que d'un seul type de relaxation. Clore et coll. ont donc développé l'approche de Lipari et Szabo de sorte à inclure deux types de mouvements internes dans le processus de relaxation afin de s'ajuster aux données RMN de relaxation de la chaîne principale de l'interleukine-1 $\beta$ .<sup>125</sup> Ne considérant toujours que les mouvements plus lents que le temps caractéristique de rotation de la protéine, la fonction de corrélation pour les mouvements internes devient biexponentielle:

$$C_i(t) = S^2 + (1 - S_v^2) \exp(-t/\tau_v) + (S_v^2 - S^2) \exp(-t/\tau_l) \quad (5.4)$$

où  $\tau_v$  et  $\tau_l$  sont les temps caractéristiques de corrélation associés respectivement à des mouvements rapides et lents. Si l'on pose que les deux types de mouvements sont approximativement indépendants, le paramètre d'ordre généralisé  $S^2$  devient un produit de deux paramètres d'ordre décorrelés, soit  $S^2 = S_v^2 S_l^2$ , où  $S_v^2$  est le

paramètre d'ordre associé au plus rapide des deux mouvements, tandis que  $S_1^2$  est celui associé au plus lent. La densité spectrale recalculée selon l'équation 4.5 à partir de cette fonction de corrélation donne une équation pour laquelle les échelles des temps caractéristiques de corrélation des deux types de mouvements internes doivent différer par au moins un ordre de grandeur, sinon la densité spectrale se simplifie à celle pour un seul mouvement de relaxation.

Les phénomènes de relaxation impliquent en réalité toute une gamme de mouvements dont les caractéristiques temporelles et spatiales sont très variées et qui peuvent être couplés entre eux. Des expériences de RMN et des simulations de DM effectuées sur différents peptides ou protéines solvatés tels l'antamanide,<sup>126</sup> l'interleukine-1 $\beta$ ,<sup>119, 124</sup> le lysozyme,<sup>120</sup> le xfn31,<sup>127, 128</sup> le GRDBD (*Glucocorticoid Receptor DNA-Binding Domain*)<sup>129</sup> et le ribonucléase H<sub>1</sub><sup>130, 131</sup> ou ribonucléase T<sub>1</sub><sup>132</sup> ont permis d'observer plusieurs types de mouvements internes de la chaîne principale:

- des mouvements rapides de librations de petite amplitude des plans amides responsables de la réorientation des vecteurs N-H et communs à tous les acides aminés et qui sont très bien échantillonnés par les simulations de DM.
- des mouvements beaucoup plus lents et de plus grandes amplitudes tels que des transitions entre deux états stabilisés par un pont hydrogène, ou une réorientation du vecteur N-H due à des mouvements qui impliquent des groupes d'acides aminés. Ces mouvements ne sont pas ou que peu échantillonnés par les simulations de DM puisqu'ils sont moins fréquents.

Ces différents types de mouvements propres à la chaîne principale ont donc suscité l'élaboration de plusieurs modèles dynamiques afin de caractériser plus formellement les mouvements qui définissent l'espace exploré par le vecteur N-H et d'en interpréter le paramètre d'ordre. Entre autres, Woessner<sup>133</sup> propose un

modèle pour des paires de spins identiques où les vecteurs spin-spin se réorientent autour d'un axe qui peut à son tour bouger de façon stochastique à une échelle de temps différente. Kinoshita et coll.<sup>134</sup> et Lipari et Szabo<sup>135</sup> proposent un modèle où l'extrémité du vecteur spin-spin a une probabilité égale de se trouver n'importe où sur la surface d'une section conique de demi-angle donné d'une sphère. Puis, comme mentionné précédemment, Clore et coll.<sup>124, 125</sup> ont proposé un modèle qui tient compte de deux types de mouvements différents par au moins un ordre de grandeur dans leur échelle de temps: les mouvements les plus rapides vont selon le modèle de diffusion dans un cône,<sup>134, 135</sup> tandis que les mouvements plus lents impliquent des sauts entre deux cônes d'orientations différentes dont la différence d'angle entre les cônes se situe entre 15° et 69°. Enfin, Brüshweiler et Wright<sup>136</sup> ont proposé le modèle GAE (*Gaussian Axial Fluctuations*) des fluctuations gaussiennes autour d'un axe pour les vecteurs N-H. Ce modèle restreint d'autant plus les mouvements du vecteur N-H de la chaîne principale en ne lui permettant de se déplacer que sur une portion de la calotte du cône qui ressemble à un croissant. Ce modèle sera décrit plus en détail à la section 5.3. Plus récemment, des simulations de DM<sup>130, 132</sup> ont permis d'établir que pour les mouvements rapides, la distribution des vecteurs N-H de la chaîne principale s'accordait bien avec le modèle de diffusion dans un cône, bien que la surface de la section conique soit elliptique plutôt que sphérique.

Tous ces auteurs proposent des modèles dynamiques où sont considérés différents types de mouvements couplés dans différents éléments de structure secondaire qui peuvent se produire à des échelles de temps différentes et à partir desquels on tente d'interpréter la valeur du paramètre d'ordre. Le but de ce chapitre est de proposer un modèle analytique qui caractérise exclusivement les mouvements rapides de librations des vecteurs N-H se produisant selon la normale du plan amide par rapport à sa position d'équilibre. Ces mouvements communs à tous les acides aminés seront caractérisés afin d'évaluer leur contribution dans le paramètre d'ordre, ainsi que la sensibilité de ce dernier à l'amplitude des fluctu-

ations (librations) qui le déterminent. De plus, ces librations sont interprétées à la section 5.2 en termes de mouvements anticorrélés des angles dièdres  $\phi$  et  $\psi$  qui déterminent principalement leur amplitude. Le formalisme du modèle analytique est développé à la section 5.2, et le modèle est évalué à la section 5.3 à l'aide de la trajectoire de 400 ps de l'état D de la calbindine  $D_{9k}$  et de comparaisons avec d'autres modèles.

## 5.2 Théorie

Une première approximation du modèle consiste à poser que le plan amide, qui comprend les atomes O, C, N et H de la chaîne principale, forme une unité planaire et rigide en tout temps, c'est-à-dire qu'il n'y a aucun degré de liberté de rotation dû au caractère de liaison double entre les atomes O-C et C-N. Par contre, les liaisons simples de part et d'autre du plan amide, entre le C et le  $C_\alpha$  de même qu'entre le N et le  $C_\alpha$ , représentent deux degrés de liberté de rotation selon lesquels le plan amide effectue un mouvement de libration autour du lien C-N (axe de rotation).

Notre modèle consiste donc à représenter le vecteur N-H de chaque acide aminé comme effectuant un mouvement de libration, soit un balancement de part et d'autre d'un axe fixe. Ce mouvement a déjà été observé par plusieurs auteurs.<sup>117, 119, 127, 128, 129, 130, 132</sup>

La figure 34 illustre le modèle pour lequel un système d'axes orthogonaux est utilisé, et la description suivante est faite en termes des coordonnées sphériques  $r$ ,  $\Theta$  et  $\Phi$ . Il faut mentionner qu'on suppose que tout mouvement de translation ou de rotation de la protéine a été éliminé, et que seuls les mouvements internes sont concernés. Dans cette représentation, le plan amide à sa position d'équilibre est superposé au plan  $x-z$  de sorte que le vecteur N-H est colinéaire et superposé à l'axe des  $x$  ( $\Phi = 0^\circ$ ,  $\Theta = 90^\circ$ ), et les atomes C et N demeurent en tout temps

dans le plan  $x-z$ . Le mouvement de libration du vecteur N-H, découplé des autres types de mouvements, ne s'effectue que dans le plan  $x-y$  avec en tout temps un angle de  $\Theta = 90^\circ$ . Aussi, les librations du vecteur N-H le répartissent de part et d'autre de l'axe des  $x$  selon une distribution gaussienne avec un écart type  $\sigma_\Phi$  de libration. La façon dont les axes du système de coordonnées ont été placés suggère que l'amplitude de l'angle de libration du vecteur N-H<sub>*i*+1</sub> est largement déterminée par les torsions effectuées par les angles dièdres  $\phi_{i+1}$  et  $\psi_i$  que l'on propose anticorrélées. Cette idée est formellement développée un peu plus loin dans cette section.

La prochaine étape du développement consiste à calculer la partie librationnelle du paramètre d'ordre tel que déjà définie au chapitre 4 par l'équation 4.10, soit:

$$\begin{aligned} \langle S^2 \rangle &= \frac{4\pi}{5} \sum_{m=-2}^2 |\langle Y_{2m}(\Theta, \Phi) \rangle|^2 \\ &= \frac{3}{4} |\langle Y_2(\Theta, \Phi) \rangle|^2 + 3 |\langle Y_1(\Theta, \Phi) \rangle|^2 + \frac{1}{4} |\langle Y_0(\Theta, \Phi) \rangle|^2 \end{aligned} \quad (5.5)$$

La partie librationnelle est obtenue en recalculant les intégrales des harmoniques sphériques à l'intérieur des limites imposées par le mouvement de libration, soit que  $\Theta=90^\circ$ , et si l'on considère que le vecteur N-H est réparti de part et d'autre de l'axe des  $x$  selon une distribution gaussienne, on peut alors obtenir  $\langle S^2 \rangle_{\text{lib.}}$  en fonction de l'écart type de la distribution de l'angle de libration,  $\sigma_\Phi$ . Dans ce cas, les intégrales à résoudre sont les suivantes (à évaluer pour  $\Theta = 90^\circ$ ):

$$\begin{aligned} \langle S^2 \rangle_{\text{lib.}} &= \frac{3}{4} \left| \int_{-\infty}^{+\infty} \sin^2 \Theta \exp(i2\Phi) (2\pi\sigma_\Phi^2)^{-1/2} \exp\left(\frac{-\Phi^2}{2\sigma_\Phi^2}\right) \sin \Theta d\Theta d\Phi \right|^2 \\ &+ 3 \left| \int_{-\infty}^{+\infty} \sin \Theta \cos \Theta \exp(i\Phi) (2\pi\sigma_\Phi^2)^{-1/2} \exp\left(\frac{-\Phi^2}{2\sigma_\Phi^2}\right) \sin \Theta d\Theta d\Phi \right|^2 \\ &+ \frac{1}{4} \left| \int_{-\infty}^{+\infty} (3 \cos^2 \Theta - 1) (2\pi\sigma_\Phi^2)^{-1/2} \exp\left(\frac{-\Phi^2}{2\sigma_\Phi^2}\right) \sin \Theta d\Theta d\Phi \right|^2 \end{aligned} \quad (5.6)$$

où  $\sigma_\Phi$  est l'écart type de la distribution des  $\Phi$  que l'on notera  $\sigma_{\text{lib.}}$  à partir de ce

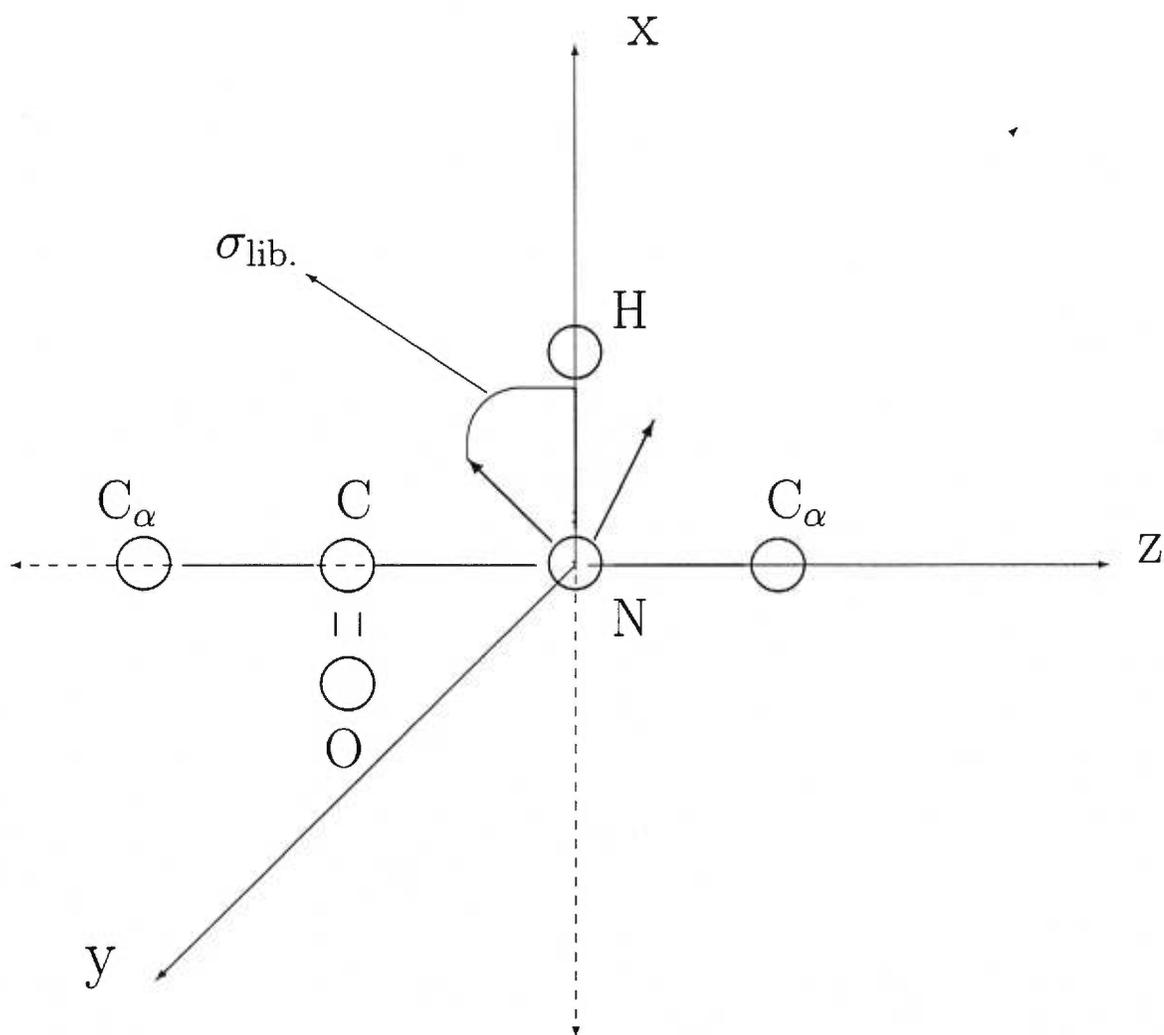


FIGURE 34. Modèle des librations d'un vecteur N-H: à la position d'équilibre du vecteur N-H ( $\Phi=0$ ), les atomes O, C, N et H se trouvent dans le plan  $x-z$ ; les librations entraînent l'atome H exclusivement dans le plan  $x-y$ , tandis que les atomes C et N demeurent dans le plan  $x-z$ .

point. Le résultat final après simplification est:

$$\begin{aligned}\langle S^2 \rangle_{\text{lib.}} &= \frac{3}{4} \left| \exp(-2\sigma_{\text{lib.}}^2) \right|^2 + 0 + \frac{1}{4} |-1|^2 \\ &= \frac{1}{4} \left[ 3 \exp(-4\sigma_{\text{lib.}}^2) + 1 \right]\end{aligned}\quad (5.7)$$

Notons que si l'on pose que  $\sigma_{\text{lib.}}$  est très petit, le  $\langle S^2 \rangle_{\text{lib.}}$  tend vers 1, soit une grande rigidité dans ce type de mouvements, tandis que pour de grandes valeurs de  $\sigma_{\text{lib.}}$ ,  $\langle S^2 \rangle_{\text{lib.}}$  tend vers 0.25, ce qui signifie que les librations ne représentent pas la totalité des orientations permises du vecteur N-H.

Maintenant, la prochaine étape du développement consiste à poser que les fluctuations des angles dièdres  $\phi_{i+1}$  et  $\psi_i$  sont anticorrélées et que ces fluctuations déterminent principalement les fluctuations de l'angle de libration  $\sigma_{\text{lib.}}$  du vecteur N-H<sub>*i+1*</sub>. Exprimé autrement, ceci devient:

$$\begin{aligned}\phi_{i+1} &= \langle \phi_{i+1} \rangle + \Delta\phi_{i+1} \\ &= \langle \phi_{i+1} \rangle + \Delta\alpha_{\text{lib.}}\end{aligned}\quad (5.8)$$

et

$$\begin{aligned}\psi_i &= \langle \psi_i \rangle + \Delta\psi_i \\ &= \langle \psi_i \rangle - \Delta\alpha_{\text{lib.}}\end{aligned}\quad (5.9)$$

où *i* est le numéro de l'acide aminé auquel appartient l'angle dièdre,  $\langle \phi_{i+1} \rangle$  et  $\langle \psi_i \rangle$  sont les valeurs moyennes des angles dièdres  $\phi_{i+1}$  et  $\psi_i$ ,  $\Delta\phi_{i+1}$  et  $\Delta\psi_i$  sont les fluctuations instantanées autour de la valeur moyenne et  $\Delta\alpha_{\text{lib.}}$  est la fluctuation instantanée de l'angle de libration. Rappelons encore que l'on suppose qu'il n'y a aucune corrélation entre la libration et tout autre type de mouvement interne. Si on réarrange ces équations de sorte à isoler  $\Delta\alpha_{\text{lib.}}$  on obtient:

$$\begin{aligned}\phi_{i+1} - \langle \phi_{i+1} \rangle = \Delta\phi_{i+1} &= \Delta\alpha_{\text{lib.}}, \\ \psi_i - \langle \psi_i \rangle = \Delta\psi_i &= -\Delta\alpha_{\text{lib.}}.\end{aligned}\quad (5.10)$$

En soustrayant les deux parties de droite des deux équations on isole  $\Delta\alpha_{\text{lib.}}$ :

$$\Delta\phi_{i+1} - \Delta\psi_i = 2\Delta\alpha_{\text{lib.}} \quad (5.11)$$

que l'on transforme de sorte à obtenir une expression en termes de la variance de la distribution de l'angle de libration  $\langle\Delta\alpha_{\text{lib.}}^2\rangle$ , soit

$$\langle(\Delta\phi_{i+1} - \Delta\psi_i)^2\rangle = 4\langle\Delta\alpha_{\text{lib.}}^2\rangle \quad (5.12)$$

ou encore en fonction de l'écart type de la distribution de l'angle de libration, c'est-à-dire  $\sigma_{\text{lib.}} = \sqrt{\langle\Delta\alpha_{\text{lib.}}^2\rangle}$ :

$$\sigma_{\text{lib.}} = \frac{1}{2}\sqrt{\langle\Delta\phi_{i+1}^2\rangle + \langle\Delta\psi_i^2\rangle - 2\langle\Delta\phi_{i+1}\Delta\psi_i\rangle} \quad (5.13)$$

où le  $\sigma_{\text{lib.}}$  est celui retrouvé aussi à l'équation 5.7,  $\langle\Delta\phi_{i+1}^2\rangle$  et  $\langle\Delta\psi_i^2\rangle$  sont respectivement les variances des distributions de l'angle dièdre  $\phi_{i+1}$  et  $\psi_i$ .

Afin de vérifier ces hypothèses, le modèle ici développé sera testé à la prochaine section à partir des données de la simulation de 400 ps de l'état D de la calbindine  $D_{9k}$  (voir chapitre 4). Entre autres, le  $\sigma_{\text{lib.}}$  du vecteur N-H de chaque acide aminé sera calculé, de même que la partie librationnelle du paramètre d'ordre du vecteur N-H. Des comparaisons avec d'autres modèles sont aussi effectuées.

### 5.3 Résultats

Afin de tester le modèle représenté par l'équation 5.7, les paramètres d'ordre des vecteurs N-H de la trajectoire totalisant 400 ps de l'état D de la calbindine  $D_{9k}$  sont tout d'abord recalculés. Les résultats sont présentés à la figure 35, de même que les paramètres d'ordre qui avaient été obtenus à partir des premiers 200 ps (voir section 4.3.2.3).

Il est intéressant de noter que les paramètres d'ordre des acides aminés 22, 44 et 60 (et d'autres moins notables) ressortent légèrement plus élevés lorsque

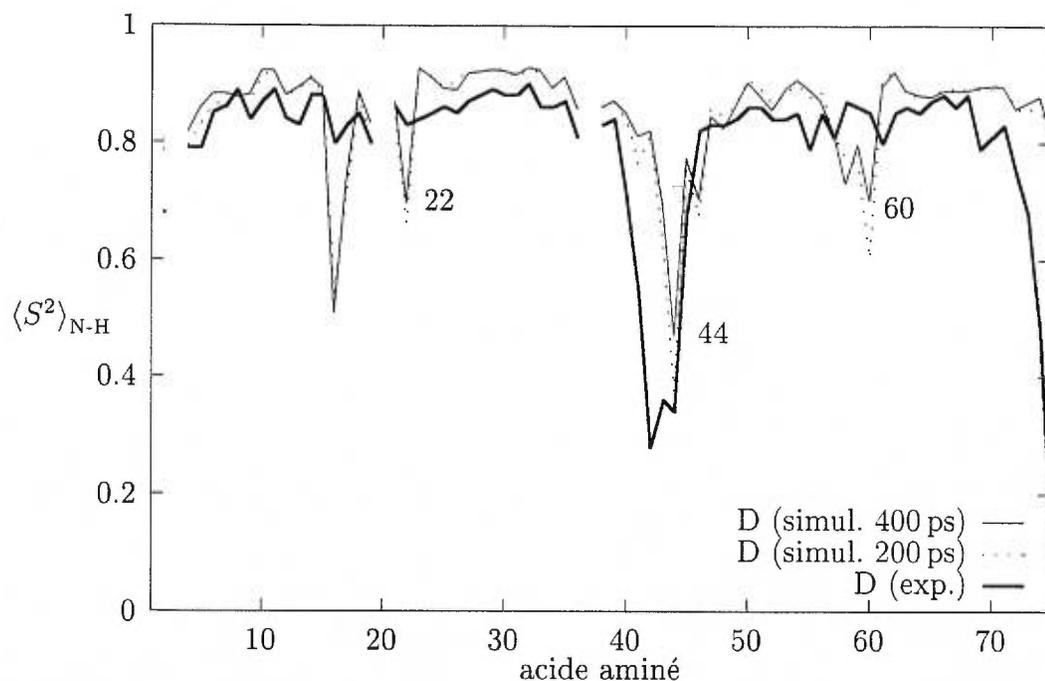


FIGURE 35. Comparaisons des paramètres d'ordre  $\langle S^2 \rangle$  des vecteurs N-H de l'état D de la calbindine  $D_{9k}$  provenant de données expérimentales avec ceux calculés à partir des trajectoires de 200 et 400 ps.

calculés à partir de la trajectoire de 400 ps. Le fait d'augmenter la durée de la simulation devrait faire baisser la valeur du paramètre d'ordre puisqu'à des durées plus longues on échantillonnerait des types de mouvements plus lents, et cela entraînerait une plus grande décorrélation et donc un paramètre d'ordre d'autant plus faible. La trop courte durée de la simulation ne permet probablement pas d'obtenir des valeurs convergées et ce qui est observé est probablement le fruit d'oscillations dans la fonction de corrélation de 400 ps. Mais à part quelques acides aminés dans les sites (16, 22, 58, 60), le segment 36-45 (43-46) et le bout C-terminal (69-75), les résultats de la simulation de 400 ps s'accordent bien avec les données expérimentales.

Afin d'isoler la partie librationnelle des paramètres d'ordre présentés à la figure 35, on suppose que les mouvements du vecteur N-H peuvent être découplés en deux types de mouvements indépendants dont le premier est le mouvement de libration tel qu'exposé à la figure 34, tandis que le second type de mouvements est associé aux fluctuations du vecteur N-H qui se produiraient dans le plan  $x-z$  et pour lesquels une mesure est prise *via* le paramètre d'ordre  $\langle S^2 \rangle_{C_\alpha-C_\alpha}$ . Le paramètre d'ordre total devient donc

$$\langle S^2 \rangle_{N-H} = \langle S^2 \rangle_{\text{lib.}} \langle S^2 \rangle_{C_\alpha-C_\alpha} \quad (5.14)$$

d'où

$$\langle S^2 \rangle_{\text{lib.}} = \frac{\langle S^2 \rangle_{N-H}}{\langle S^2 \rangle_{C_\alpha-C_\alpha}} \quad (5.15)$$

Les paramètres d'ordre  $\langle S^2 \rangle_{C_\alpha-C_\alpha}$  ont donc été calculés à partir de la trajectoire de 400 ps de l'état D, de la même façon que pour les vecteurs N-H, et sont présentés à la figure 36 où l'on voit que les mouvements sont très restreints sauf pour la région du segment 36-45, avec un paramètre d'ordre autour de 0.96. On note néanmoins que les sites sont très légèrement plus flexibles, de même que l'hélice C.

Les  $\langle S^2 \rangle_{\text{lib.}}$  calculés selon l'équation 5.15 sont présentés à la figure 37. Sauf pour les acides aminés 42 et 74, tous les rapports  $\langle S^2 \rangle_{N-H} / \langle S^2 \rangle_{C_\alpha-C_\alpha}$  sont inférieurs à un, c'est-à-dire que l'amplitude du mouvement effectué par le vecteur N-H est supérieure à l'amplitude du mouvement du vecteur  $C_\alpha-C_\alpha$ . Pour les acides aminés 42 et 74, qui se trouvent respectivement en plein coeur du segment 36-45 et dans le bout C-terminal, il faut supposer que le modèle ne s'applique tout simplement pas, étant donné leur très grande mobilité impliquant d'autres types de mouvements dont on ne tient pas compte dans notre modèle. Pour tous les autres acides aminés on note un niveau de librations donnant un  $\langle S^2 \rangle_{\text{lib.}}$  autour de 0.93, sauf pour les acides aminés 16, 22, 43, 44, 46, 58 et 60 dont la contribution librationnelle au paramètre d'ordre est nettement plus considérable.

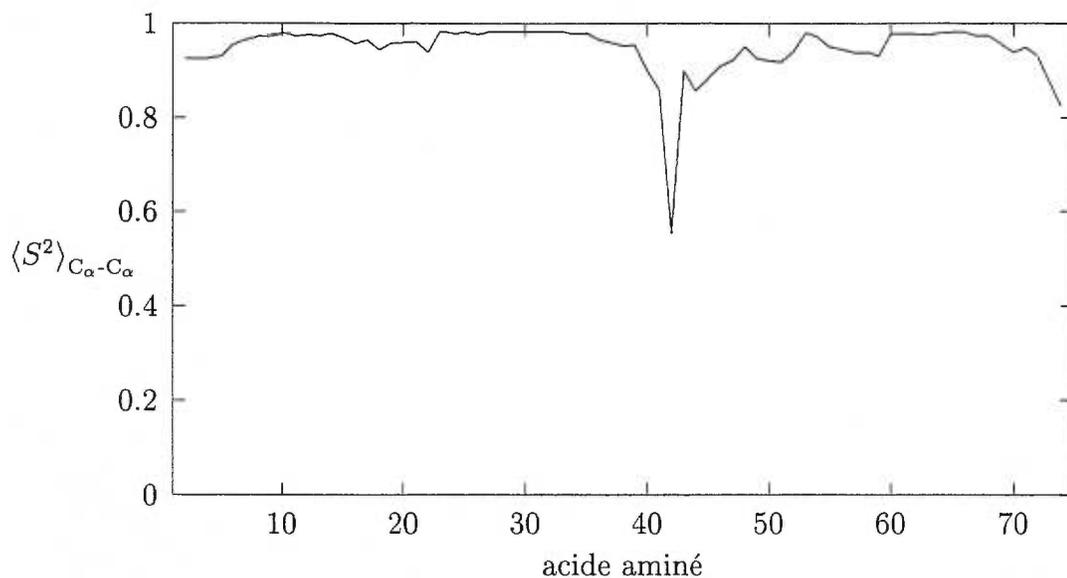


FIGURE 36. Paramètres d'ordre des vecteurs  $C_\alpha-C_\alpha$  de l'état D de la calbindine  $D_{9k}$  calculés à partir de la trajectoire de 400 ps.

Dans le but de calculer les  $\sigma_{\text{lib}}$ , selon l'équation 5.13, les fluctuations moyennes autour des valeurs moyennes des angles dièdres  $\phi$  et  $\psi$  ont aussi été recalculées à partir de la trajectoire de 400 ps. Elles sont, tout comme les  $\langle S^2 \rangle_{\text{N-H}}$ , très similaires aux résultats du chapitre 4 et ne sont donc pas présentés. Puisqu'il n'y a pas d'angle dièdre  $\phi$  pour l'acide aminé 1 ni non plus d'angle dièdre  $\psi$  pour l'acide aminé 75, les  $\sigma_{\text{lib}}$  ont été calculés pour les acides aminés 2 à 74 selon l'équation 5.13. Le terme  $\langle \Delta\phi_{i+1} \cdot \Delta\psi_i \rangle$  de cette équation doit être développé avant d'être calculé. Si l'on se réfère aux équations 5.8 et 5.9, ce terme de corrélation  $\langle \Delta\phi_{i+1} \cdot \Delta\psi_i \rangle$  devient

$$\begin{aligned} \langle \Delta\phi_{i+1} \cdot \Delta\psi_i \rangle &= \langle (\phi_{i+1} - \langle \phi_{i+1} \rangle) \cdot (\psi_i - \langle \psi_i \rangle) \rangle & (5.16) \\ &= \langle \phi_{i+1} \cdot \psi_i \rangle - \langle \phi_{i+1} \rangle \cdot \langle \psi_i \rangle \end{aligned}$$

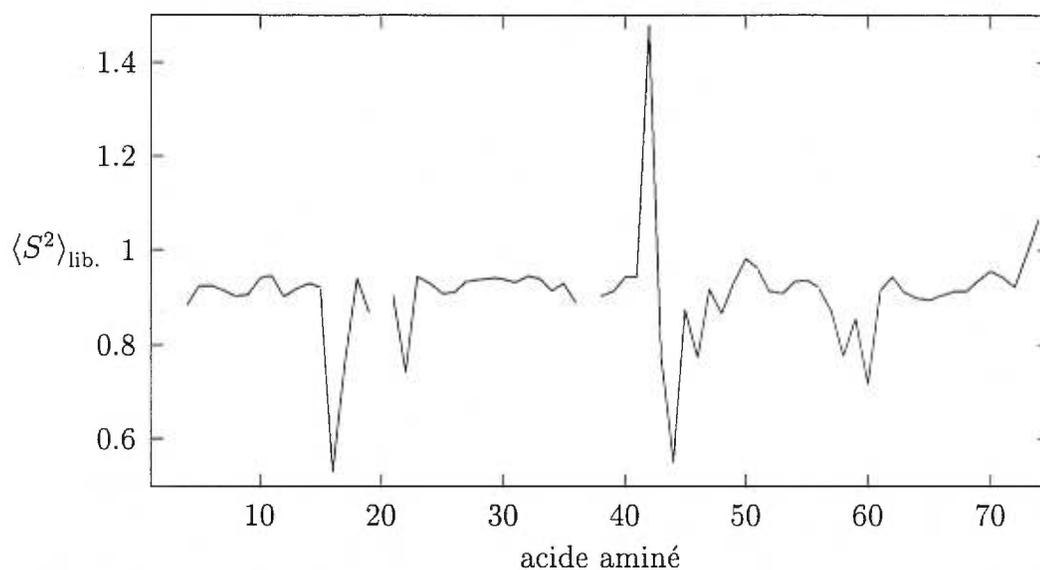


FIGURE 37. Contribution des mouvements de librations au paramètre d'ordre total calculée par  $\langle S^2 \rangle_{\text{N-H}} / \langle S^2 \rangle_{\text{C}_\alpha\text{-C}_\alpha}$ .

$$= \frac{\sum (\phi_{i+1} \cdot \psi_i)}{N} - \frac{\sum \phi_{i+1}}{N} \cdot \frac{\sum \psi_i}{N}$$

où  $N$  est le nombre de configurations dans la trajectoire.

La figure 38 compare la courbe théorique compilée à partir de l'équation 5.7 avec les résultats tirés des simulations, soit les calculs, pour chaque vecteur N-H, du  $\sigma_{\text{lib.}}$  selon les équations 5.13 et 5.16 de même que le  $\langle S^2 \rangle_{\text{lib.}}$  (équation 5.15). On voit que les résultats des simulations s'accordent très bien avec le modèle développé, sauf pour trois acides aminés (42-44) qui font partie du segment 36-45 et dans une moindre mesure l'acide aminé 74 du bout C-terminal. Pour les vecteurs N-H des acides aminés 42-44 et 74 il faut se rappeler que leurs  $\langle S^2 \rangle_{\text{N-H}}$  calculés sont trop élevés par rapport aux données expérimentales. En prenant ces valeurs expérimentales et en effectuant le recalcul du  $\langle S^2 \rangle_{\text{lib.}}$ , on trouve des valeurs qui s'accordent mieux avec le modèle. Mais cette analyse doit être regardée avec prudence, puisque ce sont présumément des mouvements lents

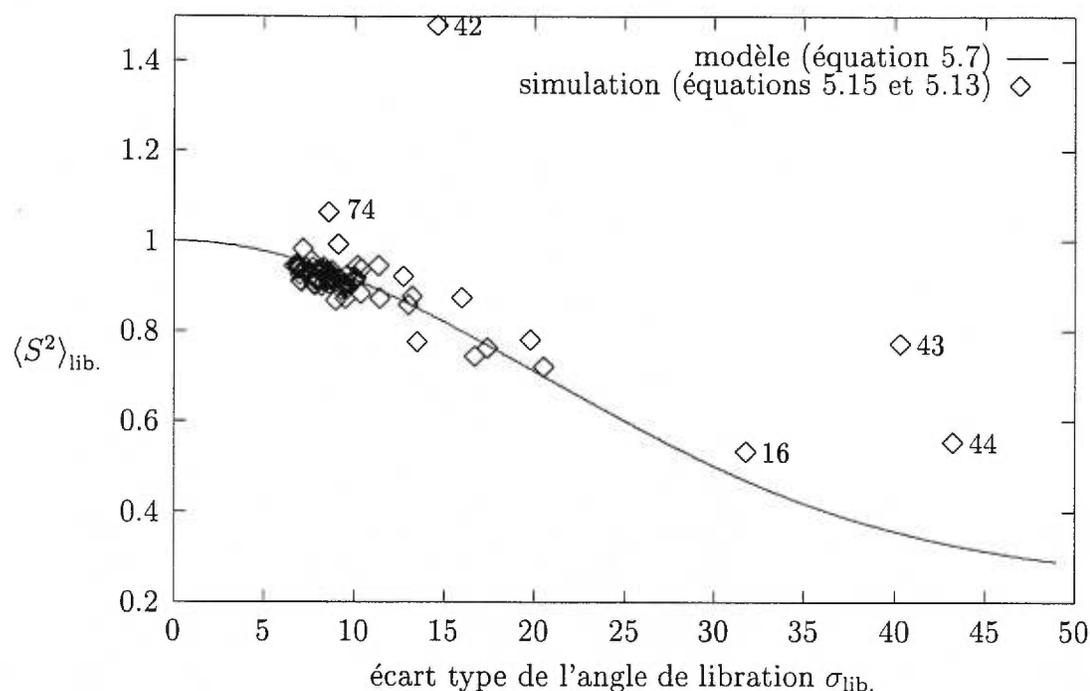


FIGURE 38. Comparaison du modèle des librations (équation 5.7) avec les valeurs tirées de la trajectoire de 400 ps de l'état D

qui ne sont pas échantillonnés et qui font que le paramètre d'ordre est trop élevé par rapport à l'expérience. Pour ces acides aminés, l'erreur provient probablement de la décomposition inadéquate proposée par l'équation 5.13, c'est-à-dire la supposition d'anticorrélation des angles  $\phi$  et  $\psi$ . Néanmoins, pour la majorité des vecteurs N-H, qui sont en accord avec l'expérience et le modèle, on observe des librations autour de  $9^\circ$ , ce qui est en accord avec les observations de Yamasaki<sup>130</sup> et Fushman.<sup>132</sup> Bien que cela dépende du modèle, on remarque que les acides aminés pour lesquels on avait obtenu des valeurs calculées de  $\langle S^2 \rangle_{\text{N-H}}$  trop faibles par rapport à l'expérience s'accordent aussi avec le modèle, et correspondent à de plus grandes fluctuations des angles dièdres  $\phi$  et  $\psi$ .

Il ressort de l'analyse que de petites variations de  $\sigma_{\text{lib.}}$  induisent des changements

considérables du  $\langle S^2 \rangle_{\text{lib.}}$ , comme le montre le tableau XVII: par exemple, pour passer de 1.0 à 0.9, le vecteur N-H doit fluctuer de  $10.7^\circ$ , et pour passer de 0.8 à 0.7, seulement  $4.2^\circ$  de fluctuations supplémentaires suffisent. Alors lorsque de trop grandes fluctuations des angles dièdres  $\phi$  et  $\psi$  sont invoquées pour expliquer les paramètres d'ordres trop faibles, il faut se rappeler que seulement quelques quatre degrés suffisent à faire baisser  $\langle S^2 \rangle_{\text{lib.}}$  de 0.1. Il ne faut pas oublier que le paramètre

$\langle S^2 \rangle_{\text{lib.}}$	$\sigma_{\text{lib.}} (^\circ)$	$\Delta^\circ$
1.0 - 0.9	0 - 10.7	10.7
0.9 - 0.8	10.7 - 15.7	5.0
0.8 - 0.7	15.7 - 19.9	4.2
0.7 - 0.6	19.9 - 24.0	4.1
0.6 - 0.5	24.0 - 28.3	4.3
0.5 - 0.4	28.3 - 33.3	5.0
0.4 - 0.3	33.3 - 40.6	7.3

TABLEAU XVII. Dépendance du  $\langle S^2 \rangle_{\text{lib.}}$  en fonction du  $\sigma_{\text{lib.}}$ .

d'ordre mesuré expérimentalement ( $\langle S^2 \rangle_{\text{N-H}}$ ) dépend, selon notre modèle, des fluctuations du vecteur  $C_\alpha-C_\alpha$ . Bien qu'il ne soit pas possible d'évaluer directement le  $\langle S^2 \rangle_{\text{lib.}}$  et de le multiplier par le  $\langle S^2 \rangle_{C_\alpha-C_\alpha}$  pour obtenir le  $\langle S^2 \rangle_{\text{N-H}}$ , on peut néanmoins savoir que le résultat est influencé par l'erreur sur le  $\langle S^2 \rangle_{\text{N-H}}$  et le  $\langle S^2 \rangle_{C_\alpha-C_\alpha}$ , et que d'autres facteurs que les fluctuations des angles dièdres sont déterminants.

D'autre part, Brühweiler et Wright<sup>136</sup> représentent les  $\langle S^2 \rangle_{\text{N-H}}$  par un modèle similaire dans lequel les mouvements du vecteur N-H correspondent à des fluctuations réparties autour d'un axe selon une distribution gaussienne. De plus, ces fluctuations sont déterminées par les fluctuations de l'angle dièdre  $\phi$  dans lequel est impliqué l'atome d'azote.

L'équation dérivée pour leur modèle est:<sup>136</sup>

$$\langle S^2 \rangle^{\text{Brüsh.}} = 1 - 3 \sin^2 \Theta \left[ \cos^2 \Theta \left( 1 - \exp \left( -\sigma_{\Phi}^2 \right) \right) + \frac{1}{4} \sin^2 \Theta \left( 1 - \exp \left( -4\sigma_{\Phi}^2 \right) \right) \right] \quad (5.17)$$

où les définitions des angles  $\Theta$  et  $\Phi$  sont les mêmes que dans notre modèle, bien que le plan amide  $y$  soit orienté de façon légèrement différente: dans leur modèle, c'est le vecteur N-C $_{\alpha}$  qui est superposé à l'axe des  $x$ , impliquant que le vecteur N-H forme un angle  $\Theta \approx 62^\circ$  plutôt que  $90^\circ$ . Dans leur modèle, le vecteur N-H se déplace dans le plan  $x-y$  et ces fluctuations sont déterminées exclusivement par les fluctuations de l'angle dièdre  $\phi$ , tandis que les fluctuations dans le plan  $x-z$  sont déterminées par les mouvements du vecteur N-C $_{\alpha}$ . De plus, l'écart type de la distribution du vecteur N-H dans le plan  $x-y$  ( $\sigma_{\Phi}$ ) correspond à l'écart type ( $\sigma_{\phi}$ ) de la distribution de l'angle dièdre  $\phi$ , puisque ce sont les fluctuations de ce dernier qui déterminent les fluctuations du vecteur N-H. Si l'on pose l'angle  $\Theta = 90^\circ$ , on retrouve l'équation 5.7. On note aussi que c'est à cette valeur de  $\Theta$  que l'influence de  $\sigma_{\Phi}$  est la plus grande, pour de petites valeurs de  $\sigma_{\Phi}$ . Pour de grandes valeurs de  $\sigma_{\Phi}$  à des valeurs de  $\Theta$  pouvant varier, on retrouve le modèle de Woessner<sup>133</sup> où toutes les positions deviennent permises.

Selon le modèle de Brühweiler et Wright,<sup>136</sup> les paramètres d'ordre  $\langle S^2 \rangle_{\text{N-H}}$  peuvent être décomposés en deux contributions: la première contribution correspond aux fluctuations du vecteur N-H dans le le plan  $x-y$ , tandis que la seconde contribution rend compte des fluctuations s'effectuant dans le plan  $x-z$  et est directement liée aux fluctuations du vecteur N-C $_{\alpha}$ . En effectuant l'opération suivante:

$$\langle S^2 \rangle_{\text{lib.}}^{\text{Brüsh.}} = \frac{\langle S^2 \rangle_{\text{N-H}}}{\langle S^2 \rangle_{\text{N-C}_{\alpha}}} \quad (5.18)$$

on isole donc la partie librationnelle correspondante à celle dans notre modèle à partir de laquelle une comparaison des deux modèles peut être effectuée. Les paramètres d'ordre des vecteurs N-C $_{\alpha}$  ont donc été calculés à partir de la trajectoire

de 400 ps afin d'évaluer les  $\langle S^2 \rangle_{\text{lib.}}^{\text{Brüsh.}}$  selon l'équation 5.18. La figure 39 montre la courbe théorique compilée à partir de l'équation 5.17 superposée aux valeurs de  $\langle S^2 \rangle_{\text{lib.}}^{\text{Brüsh.}}$  que l'on fait correspondre aux valeurs de l'écart type de l'angle dièdre  $\phi$  ( $\sigma_\phi$ ) elles aussi tirées de la trajectoire de 400 ps.

L'accord est très bon pour les points dont les  $\sigma_\phi$  sont inférieurs à  $15^\circ$ , sauf pour l'acide aminé 74 qui se comporte de la même façon que dans notre modèle. A des fluctuations de l'angle dièdre  $\phi$  supérieures à  $15^\circ$ , la distribution des points est plus dispersée que dans notre modèle, et les points les plus dissidents communs aux deux modèles représentent les acides aminés 42 et 43 du segment 36-45. Aussi, les points qui correspondent aux acides aminés 15 et 16 présentent un moins bon accord que dans notre modèle. Il ressort donc que le modèle de Brüshweiler et Wright fait défaut pour certains des acides aminés dont les valeurs calculées des  $\langle S^2 \rangle_{\text{N-H}}$  sont inférieures aux valeurs expérimentales. Yamasaki et coll.<sup>130</sup> avaient pour leur part observé que le modèle de Brüshweiler et Wright semblait bien représenter une certaine proportion des points sans pouvoir cependant prédire l'ensemble des paramètres d'ordre.

Le modèle de Brüshweiler et Wright diffère de notre modèle en ce qu'il néglige la contribution de mouvements à plus grande échelle en postulant que les fluctuations du vecteur N-H ne sont entièrement déterminées que par un type de mouvements, soit les fluctuations de l'angle dièdre  $\phi$ . Si ces fluctuations correspondent à la contribution principale dans le modèle, l'erreur associée occupe aussi une plus grande proportion. De fait, le calcul du  $\sigma_{\text{lib.}}$  dans notre modèle a pour effet de mieux représenter les fluctuations du vecteur N-H parce qu'il moyenne d'une certaine façon la valeur du  $\sigma_{\text{lib.}}$  et diminue ainsi son erreur.

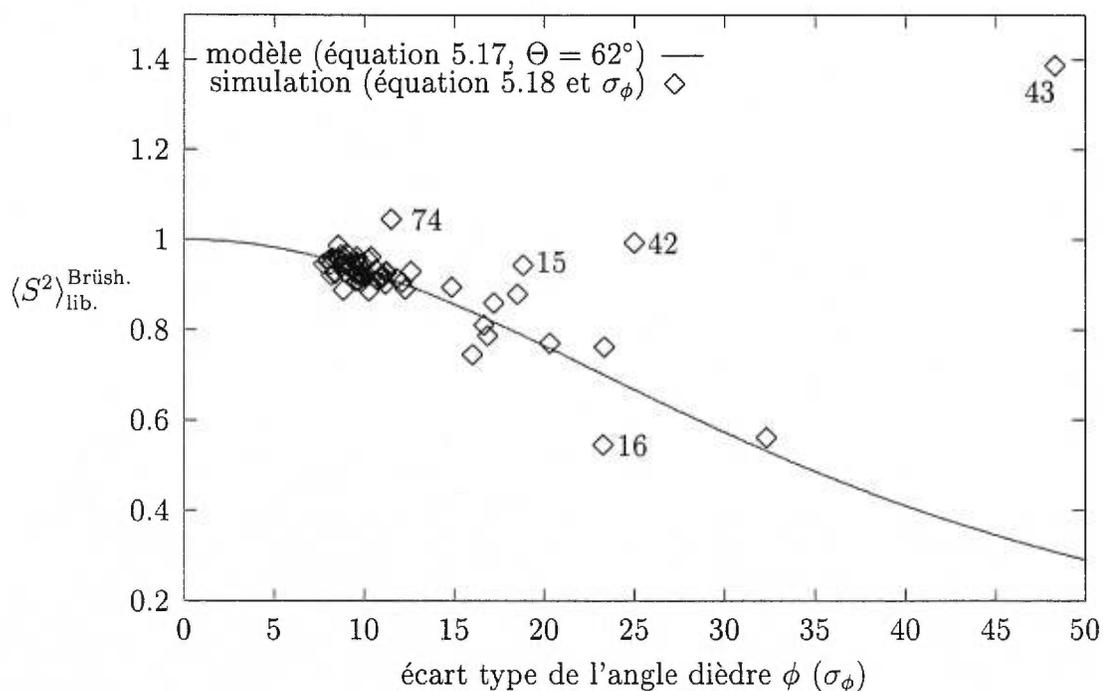


FIGURE 39. Comparaison du modèle de Brühweiler et Wright (équation 5.17,  $\Theta = 62^\circ$ ) avec les valeurs des  $\langle S^2 \rangle_{\text{lib.}}^{\text{Brüh.}}$  tirées de la trajectoire de 400 ps de l'état D et calculées selon l'équation 5.18, et les valeurs de  $\sigma_\phi$ .

#### 5.4 Conclusion

Un modèle analytique a été développé afin d'exprimer la partie librationnelle des paramètres d'ordre des vecteurs N-H en fonction de l'amplitude des librations effectuées par le vecteur N-H. De plus, ces librations sont exprimées en termes des fluctuations des angles dièdres  $\phi$  et  $\psi$  (bornant le plan amide) que l'on propose anticorrélés.

Le modèle a été testé à l'aide de la trajectoire de 400 ps de l'état D de la calbindine  $D_{9k}$ , d'où les valeurs de la partie librationnelle du paramètre d'ordre ainsi que des librations du vecteur N-H ont été extraites. Les résultats de la

trajectoire sont en très bon accord avec le modèle, sauf pour les acides aminés 42-44 (segment 36-45) et 74 (bout C-terminal). Les paramètres d'ordre calculés de ces acides aminés sont les seuls (sauf l'acide aminé 1 qui n'apparaît pas dans les tests du modèle) pour lesquels les paramètres d'ordre calculés sont supérieurs aux données expérimentales. L'analyse de ces résultats ne semble donc pas pouvoir s'expliquer en termes de fluctuations des angles dièdres  $\phi$  et  $\psi$ .

Pour les points qui s'accordent bien avec les données expérimentales (paramètres d'ordre) et avec le modèle, on trouve des librations moyennes autour de  $9^\circ$ , ce qui est raisonnablement en accord avec d'autres auteurs<sup>130, 132</sup> qui ont trouvé des librations de  $10-12^\circ$  communes à tous les acides aminés. Les résultats suggèrent aussi que des fluctuations trop grandes des angles dièdres  $\phi$  et  $\psi$  représentent en partie la source des trop faibles valeurs calculées pour certains paramètres d'ordre. Ainsi, une surestimation des librations par seulement quelques degrés est suffisante pour expliquer le désaccord avec les données expérimentales.

Notre modèle a été comparé à un modèle analytique similaire (Brüshweiler et Wright) qui exprime les paramètres d'ordre des vecteurs N-H exclusivement en termes des fluctuations des angles dièdres  $\phi$ , et auquel il s'accorde raisonnablement, sauf pour quelques acides aminés dont les valeurs calculées des paramètres d'ordre sont inférieures aux valeurs expérimentales. Il est proposé que ce modèle donne une représentation trop restreinte des paramètres d'ordre en négligeant la contribution des angles dièdres  $\psi$ .

Il serait intéressant de caractériser la partie des librations qui s'effectuent dans le plan  $x-z$  afin de confirmer que le vecteur N-H diffuse dans un cône elliptique. Des valeurs moyennes de librations dans le plan  $x-z$  autour de  $3-5^\circ$  ont été observés par d'autres auteurs.<sup>130, 132</sup> A cette fin, les harmoniques sphériques du paramètre d'ordre devraient être recalculées avec les nouvelles contraintes spatiales, et cette contribution librationnelle dans le plan devrait être comparée au paramètre d'ordre  $\langle S^2 \rangle_{C_\alpha-C_\alpha}$  correspondant au bon acide aminé.

## CHAPITRE 6

### Calculs de la différence d'énergie libre associée à la mutation V61A et contribution à la coopérativité

#### 6.1 Introduction

Au chapitre 4, une analyse comparative détaillée de la structure de la calbindine  $D_{9k}$  dans quatre de ses états d'occupation a été effectuée et l'importance de l'acide aminé en position 61 (Val61) a été examinée. La chaîne latérale de la Val61 est impliquée dans un grand nombre de contacts de van der Waals à l'intérieur de la protéine, et la chaîne principale forme deux ponts hydrogènes reliant les deux sites (Leu23-Val61). L'importance des interactions hydrophobes impliquant cet acide aminé a aussi été vérifiée et il a été montré que la chaîne latérale de la Val61 avait des énergies d'interaction de plus en plus favorables avec le niveau d'occupation du  $Ca^{2+}$  dans la protéine ( $D > S_{II} \geq S_I(D) > A(S_{II})$ ).

Dans ce chapitre nous étudierons les effets de la mutation V61A (Val61→Ala61) sur la coopérativité de la calbindine  $D_{9k}$  par des calculs de différences d'énergies libres afin de saisir une partie du rôle structural et fonctionnel joué par la Val61. Comme il a déjà été mentionné, la calbindine  $D_{9k}$  étudiée dans cette thèse est un mutant (P43G) qui a été caractérisé afin de valider son utilisation pour représenter la protéine sauvage. Les différences structurales du mutant P43G sont mineures et localisées dans la région de la mutation.<sup>14</sup> La disparité des constantes macroscopiques de liaison de même que l'affinité totale pour le  $Ca^{2+}$  ( $K_1K_2$ ) entre la calbindine  $D_{9k}$  sauvage et son mutant P43G sont très mineures.<sup>137</sup>

Pour le mutant P43M-V61A, le changement de coopérativité a été évalué à  $\Delta(\Delta\Delta G^{\text{coop.}}) = +2.13 \text{ kcal/mol}$  à une concentration de KCl  $\leq 2 \text{ mM}$ ,<sup>122</sup> c'est-à-dire que la coopérativité devient négative puisqu'elle est de  $-1.96 \text{ kcal/mol}$  pour le mutant P43M à la même concentration. La mutation V61A sur la calbindine D<sub>9k</sub> P43G est ici effectuée par des méthodes de calculs d'énergies libres dont les bases ont déjà été exposées aux sections 2.3 et 3.2.1. Bien qu'il ne s'agisse pas exactement du même double mutant (expériences: P43M-V61A; simulations: P43G-V61A), la similitude entre les comportements des mutants P43M et P43G (constantes de liaison macroscopiques et énergie libre de liaison totale),<sup>138</sup> suggère qu'une comparaison avec le mutant V61A est valide. Ces méthodes sont développées davantage à la section 6.2 afin de s'ajuster aux contraintes du système beaucoup plus imposant que représente la calbindine D<sub>9k</sub> hydratée. Bien que des différences d'énergies libres associées à des mutations aient déjà été effectuées par des méthodes similaires, ce calcul effectué sur la calbindine D<sub>9k</sub> dans le but d'obtenir des éclaircissements sur la coopérativité est une première. Ces calculs permettront d'aborder le thème central de cette thèse, soit l'étude de la coopérativité dans la liaison des Ca<sup>2+</sup> aux sites de la calbindine D<sub>9k</sub> *via* l'étude du rôle joué par la Val61 dans le processus de coopérativité en comparant la structure des sites de la calbindine D<sub>9k</sub> P43G étudiée tout au long de cette thèse avec son mutant P43G-V61A.

Les effets d'un grand nombre de mutations sur la calbindine D<sub>9k</sub> ont déjà été étudiés<sup>14, 16, 17, 31, 43, 114, 137, 139</sup> et il existe toute une liste de mutations effectuées sur la calbindine D<sub>9k</sub> P43M<sup>122</sup> pour lesquelles le changement de coopérativité a été mesuré et dont fait partie le mutant V61A. Les mutations P20G, P20G avec retrait de l'Asn21 et Y13F avec Pro20 retiré ont des effets structuraux très localisés au site I mais affectent cependant les deux constantes de liaison macroscopiques.<sup>16, 43, 139</sup> Les mutations impliquant la neutralisation d'acides aminés chargés à la surface de la calbindine D<sub>9k</sub> (site I: E17Q, D19N et E26Q)<sup>17, 23, 43</sup> ont aussi des conséquences structurales mineures, voire indétectables, qui se limitent

au site I, mais qui peuvent avoir des conséquences considérables sur l'affinité du  $\text{Ca}^{2+}$  pour les deux sites ainsi que sur la coopérativité. De plus, les effets de ces mutations sont additifs. Ces résultats ont donc révélé l'importance des interactions électrostatiques dans la liaison des ions  $\text{Ca}^{2+}$  aux sites, de même que dans le processus de la coopérativité.

La neutralisation de la charge sur l'Asp60 (E60Q) n'a eu que des effets mineurs sur la liaison des ions.<sup>17</sup> Cependant, la mutation E60D<sup>31</sup> où la charge est conservée mais où la longueur de la chaîne latérale est diminuée, a eu un effet dramatique: l'affinité totale pour le  $\text{Ca}^{2+}$  ( $K_1K_2$ ) est diminuée par un facteur de 38 par rapport à la protéine sauvage et, contrastant avec ce résultat, il y a une augmentation de la coopérativité. Une étude structurale par diffraction des rayons X<sup>31</sup> a révélé que le fait de diminuer la longueur de la chaîne latérale de l'Asp60 empêchait cette dernière d'établir un pont hydrogène avec la molécule d'eau qui est directement liée au  $\text{Ca}^{2+}$  dans le site I. Il a été suggéré que l'absence de ce pont hydrogène diminuait la stabilité de plusieurs ligands aux deux sites et par conséquent leur affinité pour le  $\text{Ca}^{2+}$ .<sup>31</sup>

Les travaux de Johansson et coll.<sup>140, 141, 142, 143</sup> ont démontré, entre autres, que la double mutation A15D-P20G avec retrait des Ala14 et Asn21 dans le site I transformaient le domaine N-terminal en une main EF normale. Enfin, le mutant N56A a servi à caractériser l'état  $S_1(D)$  puisqu'il permet la liaison séquentielle du  $\text{Ca}^{2+}$  avec très peu de perturbations de la structure des états A et D. Ceci étant donné que la chaîne latérale de l'Asn56 n'a que très peu de contacts avec le reste de la protéine.<sup>114</sup>

Dans ce chapitre, la section 6.2 rappelle les bases théoriques et explique la méthode utilisée pour effectuer la mutation d'une alanine en valine. A la section 6.3, un exemple test est présenté afin d'illustrer cette méthode sur un petit peptide et de la valider. Les résultats et la discussion de la mutation effectuée sur la calbindine  $D_{9k}$  sont présentés à la section 6.4. Les conclusions et suggestions sont

données à la dernière section (6.5).

## 6.2 Méthode

### 6.2.1 Formalisme

Se référant à la figure 40 (et à l'équation 1.8), les changements d'énergies libres associés à chaque étape de liaison du  $\text{Ca}^{2+}$  à la calbindine  $\text{D}_{9k}$ , de même que le changement total associé à la coopérativité sont donnés par:

$$\begin{aligned}\Delta\Delta G_{\text{Val}}^{\text{coop.}} &= \Delta G_{\text{Val}}^{\text{S}_{\text{II}} \rightarrow \text{D}} - \Delta G_{\text{Val}}^{\text{A} \rightarrow \text{S}_{\text{I}}(\text{D})} \\ &= \left(G_{\text{Val}}^{\text{D}} - G_{\text{Val}}^{\text{S}_{\text{II}}}\right) - \left(G_{\text{Val}}^{\text{S}_{\text{I}}(\text{D})} - G_{\text{Val}}^{\text{A}}\right)\end{aligned}\quad (6.1)$$

où l'indice Val signifie qu'il s'agit de la calbindine  $\text{D}_{9k}$  avec une valine en position 61. De la même façon, pour le mutant V61A avec une alanine en position 61, les mêmes changements d'énergies libres sont notés par:

$$\begin{aligned}\Delta\Delta G_{\text{Ala}}^{\text{coop.}} &= \Delta G_{\text{Ala}}^{\text{S}_{\text{II}} \rightarrow \text{D}} - \Delta G_{\text{Ala}}^{\text{A} \rightarrow \text{S}_{\text{I}}(\text{D})} \\ &= \left(G_{\text{Ala}}^{\text{D}} - G_{\text{Ala}}^{\text{S}_{\text{II}}}\right) - \left(G_{\text{Ala}}^{\text{S}_{\text{I}}(\text{D})} - G_{\text{Ala}}^{\text{A}}\right)\end{aligned}\quad (6.2)$$

de sorte que la *différence* de coopérativité entre la protéine sauvage et son mutant est obtenue par:

$$\Delta(\Delta\Delta G_{\text{Val} \rightarrow \text{Ala}}^{\text{coop.}}) = \left(\Delta G_{\text{Val} \rightarrow \text{Ala}}^{\text{D}} - \Delta G_{\text{Val} \rightarrow \text{Ala}}^{\text{S}_{\text{II}}}\right) - \left(\Delta G_{\text{Val} \rightarrow \text{Ala}}^{\text{S}_{\text{I}}(\text{D})} - \Delta G_{\text{Val} \rightarrow \text{Ala}}^{\text{A}}\right) \quad (6.3)$$

où les  $\Delta G_{\text{Val} \rightarrow \text{Ala}}^{\text{état } i} = G_{\text{Ala}}^{\text{état } i} - G_{\text{Val}}^{\text{état } i}$  sont les changements d'énergies libres induits par la mutation dans chacun des quatre états et qui sont les quatre observables calculées dans ce chapitre, afin de pouvoir évaluer le changement de coopérativité total induit par cette mutation,  $\Delta(\Delta\Delta G_{\text{Val} \rightarrow \text{Ala}}^{\text{coop.}})$ .

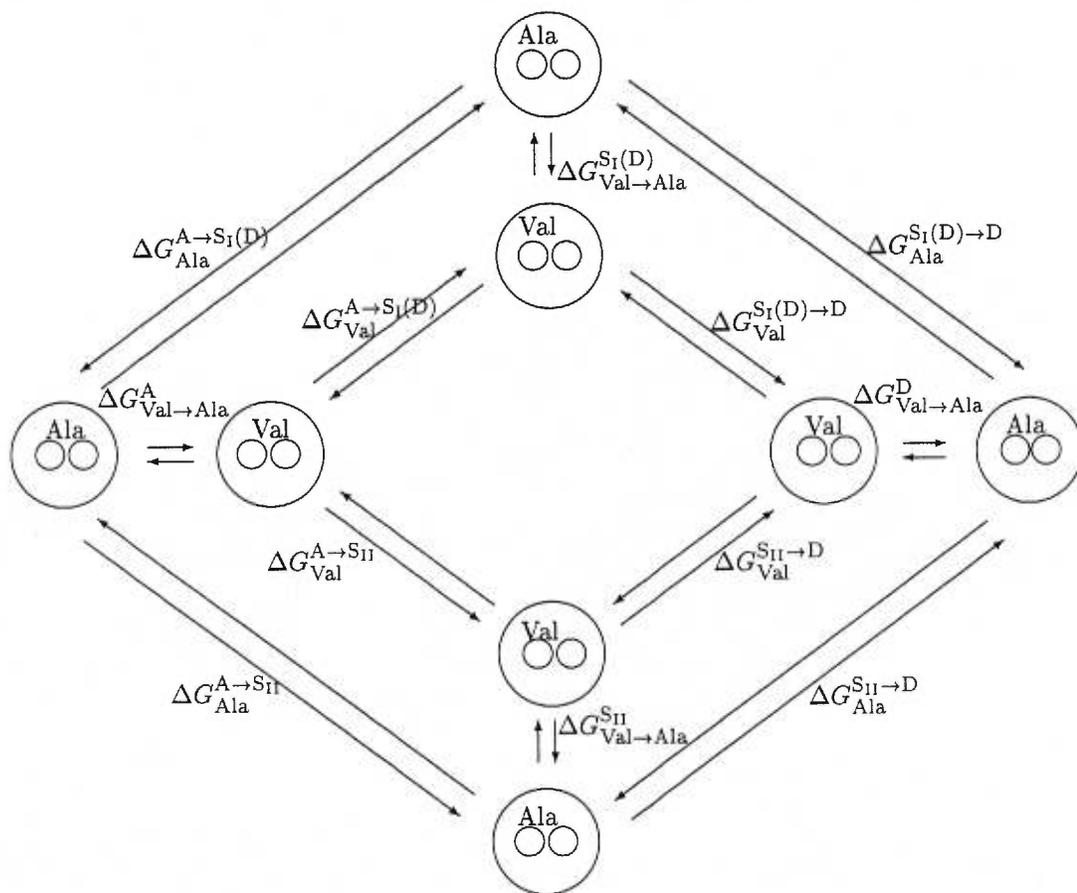


FIGURE 40. Cycles thermodynamiques des changements d'énergies libres associés aux liaisons successives du  $\text{Ca}^{2+}$  dans la calbindine  $D_{9k}$  (coopérativité) dans les cas où l'acide aminé en position 61 est une valine (cycle interne) ou une alanine (cycle externe). Les étapes représentées par des flèches horizontales ou verticales correspondent aux changements d'énergies libres associées à la mutation V61A pour chaque état d'occupation de la calbindine  $D_{9k}$ .

Les changements des constantes macroscopiques  $K_1$  et  $K_2$  associés à la mutation ayant été mesurés expérimentalement, il est donc d'un intérêt particulier de trouver une expression reliant ces mesures expérimentales à nos calculs de différences d'énergies libres. Si l'on pose que  $k_I = k_{II}$  et que  $k_{I,II} = k_{II,I}$  pour les constantes microscopiques, alors les constantes macroscopiques sont données par  $K_1 = 2k_I = 2k_{II}$  et  $K_2 = k_{I,II}/2 = k_{II,I}/2$  (selon les équations 1.6 et 1.7), le rapport d'une constante macroscopique ( $K_1$  ou  $K_2$ ) entre la protéine mutée et la protéine sauvage est donné par:

$$\begin{aligned} \left( \frac{K_1^{\text{Ala}}}{K_1^{\text{Val}}} \right) &= \exp \left[ -\beta \left( \Delta G_{\text{Val} \rightarrow \text{Ala}}^{\text{SI}(\text{D})} - \Delta G_{\text{Val} \rightarrow \text{Ala}}^{\text{A}} \right) \right] \quad (a) \\ &= \exp \left[ -\beta \left( \Delta G_{\text{Val} \rightarrow \text{Ala}}^{\text{SII}} - \Delta G_{\text{Val} \rightarrow \text{Ala}}^{\text{A}} \right) \right] \quad (b) \end{aligned} \quad (6.4)$$

ou

$$\begin{aligned} \left( \frac{K_2^{\text{Ala}}}{K_2^{\text{Val}}} \right) &= \exp \left[ -\beta \left( \Delta G_{\text{Val} \rightarrow \text{Ala}}^{\text{D}} - \Delta G_{\text{Val} \rightarrow \text{Ala}}^{\text{SI}(\text{D})} \right) \right] \quad (a) \\ &= \exp \left[ -\beta \left( \Delta G_{\text{Val} \rightarrow \text{Ala}}^{\text{D}} - \Delta G_{\text{Val} \rightarrow \text{Ala}}^{\text{SII}} \right) \right] \quad (b) \end{aligned} \quad (6.5)$$

La figure 41 illustre les chaînes latérales d'une valine et d'une alanine et établit la nomenclature des atomes concernés par la mutation et qui sera utilisée dans le reste du chapitre. Les spécifications topologiques des deux acides aminés existent simultanément lors des calculs. Selon la procédure que nous avons suivie, cela implique que les atomes de type HB, CG et HG de la valine, ainsi que les HB de l'alanine contribuent au nombre total de particules. Cependant, il faut s'assurer qu'il n'y ait pas d'interactions entre les chaînes latérales des deux espèces, c'est-à-dire que l'une soit invisible pour l'autre (c'est-à-dire que les HB, CG1, CG2 et les six HG★ de la valine n'interagissent pas avec les HB1, HB2 et HB3 de l'alanine). Les interactions entre chaque chaîne latérale et son environnement sont calculées par une fonction de potentiel  $H(\lambda)$  hybride:

$$\begin{aligned} H(\lambda) &= (1 - \lambda) H_{\text{Val}(\text{Ala})} + \lambda H_{\text{Ala}(\text{Val})} \\ &= H_{\text{Val}(\text{Ala})} + \lambda \left( H_{\text{Ala}(\text{Val})} - H_{\text{Val}(\text{Ala})} \right) \end{aligned} \quad (6.6)$$

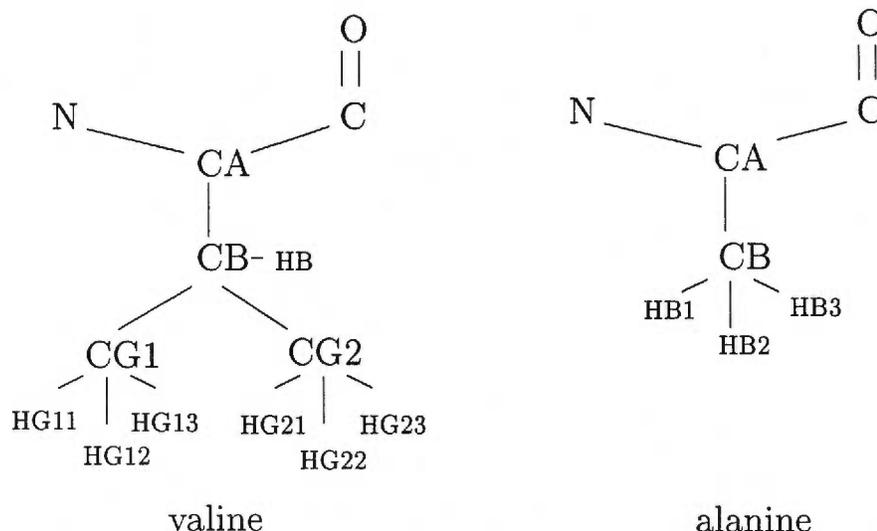


FIGURE 41. Représentation des acides aminés faisant partie de la mutation: les atomes faisant partie de la mutation sont les atomes qui ne sont pas communs aux deux acides aminés, c'est-à-dire que les deux groupements méthyles de la valine sont échangés pour deux HB dans l'alanine.

où le Val(Ala) ou Ala(Val) signifie que les deux chaînes latérales sont présentes, mais que celle entre parenthèses est invisible pour tous les autres atomes.

La mutation d'un acide aminé en un autre nécessite la perturbation à la fois du paramètre  $\sigma$  de Lennard-Jones de chaque atome impliqué dans la mutation en plus de sa charge. La perturbation du paramètre de Lennard-Jones, associée à la formation de l'espace physique occupé par l'atome, nécessite plus de soins que lorsque seule la charge est perturbée, probablement parce que cela implique d'abord des ajustements translationnels lents des atomes environnants, tandis que la croissance d'une charge implique principalement la réorientation de dipôles.<sup>144</sup> Les calculs de perturbations impliquant l'apparition et la disparition d'atomes nécessitent donc des périodes d'équilibration plus longues avant de commencer l'échantillonnage. De plus, il est nécessaire de découper plus finement le paramètre de couplage  $\lambda$  là où le changement d'énergie libre associé est le plus grand, soit aux extrémités  $\lambda = 0$  et  $\lambda = 1$ . Dans chaque cas, le paramètre de couplage  $\lambda$

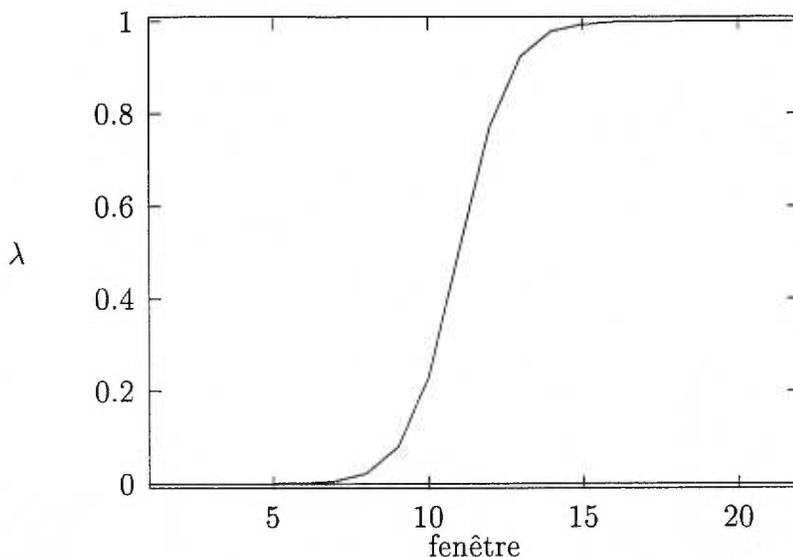


FIGURE 42. Variation du paramètre de couplage  $\lambda$  en fonction de la fenêtre de simulation. Une fonction  $\tanh$  a servi à générer les valeurs de  $\lambda$ .

a été divisé en 22 fenêtres avec une augmentation non linéaire, tel qu'illustré à la figure 42. Pour les méthodes de perturbations thermodynamiques (P.T.) et d'intégrations thermodynamiques (I.T.), chaque perturbation est effectuée à  $\lambda \pm \Delta\lambda$  et les valeurs de  $\lambda$  sont distantes par  $2\Delta\lambda$ , sauf pour les cas  $\lambda = 0$  et  $\lambda = 1$  où les perturbations sont effectuées à  $\lambda$  et respectivement à  $\lambda + \Delta\lambda$  et  $\lambda - \Delta\lambda$ .

La qualité du calcul de perturbation est entièrement tributaire du fait que l'ensemble des configurations généré durant la simulation à une valeur de  $\lambda$  donnée doit être représentatif d'un système à l'équilibre thermodynamique pour l'Hamiltonien  $H(\lambda)$ . Autrement dit, il faut que les changements de  $\lambda$  soient lents et que l'échantillonnage soit suffisant. Un système dont l'espace des configurations est beaucoup plus restreint par rapport à un autre nécessitera un temps de simulation (échantillonnage) moins grand à chacune des fenêtres. Un symptôme que le calcul de perturbations n'est pas convergé est la disparité entre les résultats des calculs à l'aller ( $\lambda = 0 \rightarrow \lambda = 1$ ) et au retour ( $\lambda = 1 \rightarrow \lambda = 0$ ).<sup>145</sup> Cependant, le fait que

les résultats des calculs à l'aller et au retour soient identiques est une condition nécessaire, mais non suffisante, pour juger de la qualité du résultat final.

Trois méthodes ont été utilisées afin d'effectuer ces calculs, soit la méthode des perturbations thermodynamiques (P.T.) qui a été exposée à la section 2.3, de même que la méthode d'intégration thermodynamique (I.T.).

La troisième méthode consiste en l'algorithme WHAM (*Weighted Histogram Analysis Method*)<sup>146, 147</sup> qui a été utilisé pour combiner les valeurs des énergies perturbées à chaque fenêtre et ainsi obtenir une meilleure valeur de la perturbation totale. Ceci consiste en une méthode itérative auto-cohérente dont le but est d'obtenir les constantes d'énergies libres  $G_k$  optimales correspondant à l'Hamiltonien hybride donné à l'équation 6.6. Ces constantes d'énergies libres  $G_k$  sont données par:

$$\exp[-\beta G_k] = \sum_{i=1}^{N_f} \sum_{t=1}^{T_i} \exp[-\beta \lambda_k \Delta H_i(t)] \times \left[ \sum_{i=j}^{N_f} T_j \exp[\beta (G_j - \lambda_j \Delta H_i(t))] \right]^{-1} \quad (6.7)$$

où  $N_f$  est le nombre de fenêtres de  $\lambda$ ,  $T_i$  est la durée de la simulation à la fenêtre  $i$ ,  $\lambda_k$  est la valeur du paramètre de couplage à la  $k^e$  fenêtre et  $\Delta H_i(t)$  est la différence d'énergie potentielle ( $H_{\text{Ala}(\text{Val})} - H_{\text{Val}(\text{Ala})}$ ) obtenue au temps  $t$  à la  $i^{\text{ème}}$  fenêtre. C'est la différence des valeurs convergées des  $G_k$  à  $\lambda_{k=1} = 0$  et  $\lambda_{k=N_f} = 1$  qui donne la meilleure valeur de la différence d'énergie libre totale.

### 6.3 Illustration et validation de la méthode

Avant de procéder à la mutation V61A sur la calbindine D<sub>9k</sub>, le protocole a d'abord été testé sur un petit peptide composée de trois alanines, où l'alanine du centre a été mutée en valine. La différence d'énergie libre d'hydratation entre une alanine et une valine dans un tripeptide a été calculée selon le cycle thermo-

dynamique illustré à la figure 43.

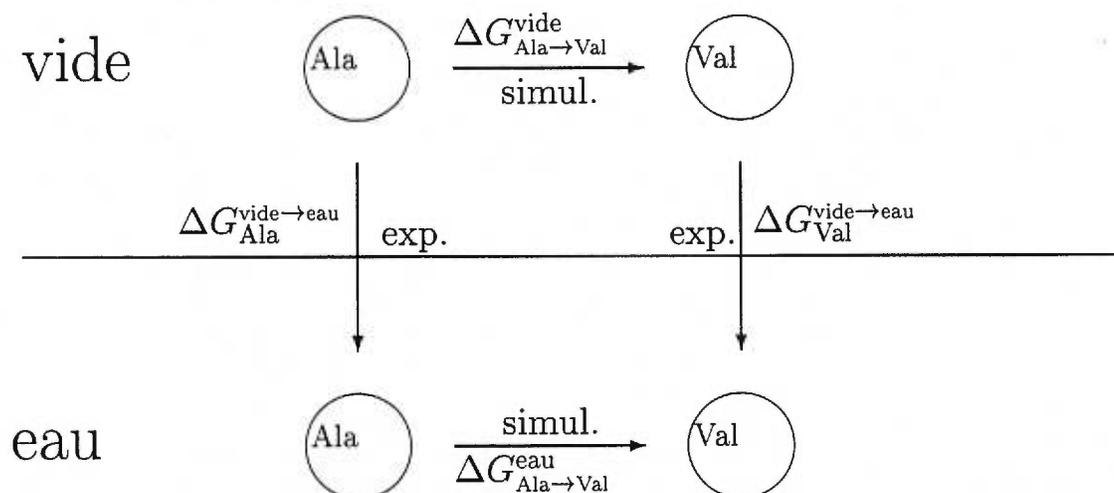


FIGURE 43. Cycle thermodynamique des changements d'énergies libres associés à la mutation Ala→Val.

Le terme à calculer est:

$$\begin{aligned}
 \Delta\Delta G^{\text{hydra.}} &= \Delta G_{\text{Val}}^{\text{hydra}} - \Delta G_{\text{Ala}}^{\text{hydra}} & (6.8) \\
 &= (G_{\text{Val}}^{\text{eau}} - G_{\text{Val}}^{\text{vide}}) - (G_{\text{Ala}}^{\text{eau}} - G_{\text{Ala}}^{\text{vide}}) \\
 &= (G_{\text{Val}}^{\text{eau}} - G_{\text{Ala}}^{\text{eau}}) - (G_{\text{Val}}^{\text{vide}} - G_{\text{Ala}}^{\text{vide}}) \\
 &= \Delta G_{\text{Ala} \rightarrow \text{Val}}^{\text{eau}} - \Delta G_{\text{Ala} \rightarrow \text{Val}}^{\text{vide}}
 \end{aligned}$$

où les  $\Delta G_{\text{Ala} \rightarrow \text{Val}}^{\text{eau}}$  et  $\Delta G_{\text{Ala} \rightarrow \text{Val}}^{\text{vide}}$  sont les différences d'énergies libres associées à la transformation d'une Ala en Val dans l'eau ou dans le vide et qui sont calculés par les trois méthodes *via* la DM.

Le même protocole de simulation a été suivi pour le cas du tripeptide d'alanine dans le vide ainsi que dans une sphère de 73 molécules d'eau, soit le nombre minimum afin que les acides aminés demeurent hydratés tout au long des simulations. Les conditions à la frontière sphérique du solvant<sup>61</sup> sont celles exposées à la section 2.2.2. Tout d'abord, les énergies des structures qui

correspondent aux  $\lambda = 0$  et  $\lambda = 1$  ont été minimisées sur 1000 pas de minimisation *Steepest-Descent*. S'ensuit une période d'équilibration de 2 ps avec un pas d'intégration de 0.001 ps. L'algorithme SHAKE<sup>65</sup> a été utilisé afin de maintenir fixes tous les liens chimiques impliquant des atomes d'hydrogène; aucun rayon de coupure des interactions n'a été appliqué. De la dynamique de Langevin a été simulée à 300 K avec un coefficient de friction correspondant à un temps caractéristique de relaxation de  $25 \text{ ps}^{-1}$  et appliqué à tous les atomes sauf à ceux d'hydrogène. Une constante de force de  $10 \text{ kcal}/(\text{mol} \cdot \text{Å}^2)$  a été appliquée au tripeptide de sorte à ce qu'il demeure au centre de masse du système.

Puis, dans les mêmes conditions, on simule une période d'équilibration de 2 ps à chaque valeur de  $\lambda$  avant que ne débute la période de production de la trajectoire dont les structures échantillonnées serviront au calcul de perturbations. Toujours pour chaque valeur de  $\lambda$ , 13 650 configurations (ou 13.65 ps) ont été récoltées pour un total de 300 ps simulées dans chaque direction de  $\lambda$ . Cette durée particulière de simulation a été choisie parce qu'il a été suggéré<sup>144, 148</sup> que c'est la durée minimale requise pour obtenir une convergence raisonnable pour des systèmes de taille similaire.

Les résultats des calculs de perturbations (mutations Ala→Val) effectués dans le vide et dans la sphère de 73 molécules d'eau sont donnés aux figures 44 et 45 et montrent respectivement les changements des  $\Delta G_{\text{Ala} \rightarrow \text{Val}}^{\text{vide}}$  et  $\Delta G_{\text{Ala} \rightarrow \text{Val}}^{\text{eau}}$  en fonction de  $\lambda$ . Ces figures font ressortir l'importance d'un découpage plus fin dans les régions de  $\lambda = 0$  et  $\lambda = 1$  car elles donnent lieu à des changements d'énergies libres plus importants. Le potentiel hybride (équation 6.6) de paire avec la grandeur des variations de l'énergie du terme de Lennard-Jones (équation 2.22) en sont les raisons. On voit aussi que c'est pour le résultat obtenu de l'algorithme WHAM que les calculs à l'aller et au retour convergent le mieux dans l'eau. La courbe WHAM aller et retour (figure 44 du bas) correspond au meilleur résultat que l'on peut obtenir à partir de cette distribution des  $\Delta H_j$ .

On notera cependant une légère disparité entre la forme de la courbe WHAM "aller" combinée avec seulement les résultats à l'aller, et la courbe WHAM "retour" combinée avec seulement les résultats au retour. Les structures finales recueillies après toute la simulation aux  $\lambda = 0$  et  $\lambda = 1$  ne révèlent pas de différences majeures pouvant expliquer cette différence. Mais il se peut, étant donnée la grande liberté de mouvements que confère le vide, que la prise de deux échantillonnages dans différentes parties d'une surface d'énergie potentielle très complexe puisse expliquer cette différence.

Le tableau XVIII résume les résultats des figures 44 et 45 en ne donnant que les points finaux des calculs de perturbations, c'est-à-dire les valeurs cumulées de toutes les perturbations à chaque fenêtre.

		$\Delta G_{\text{Ala} \rightarrow \text{Val}}^{\text{vide}}$	$\Delta G_{\text{Ala} \rightarrow \text{Val}}^{\text{eau}}$
aller ( $\lambda_{0 \rightarrow 1}$ )	P.T.	-0.91	2.26
	I.T.	-0.87	1.87
	WHAM	-0.60	2.24
retour ( $\lambda_{1 \rightarrow 0}$ )	P.T.	0.67	-2.39
	I.T.	0.49	-2.05
	WHAM	0.71	-2.45
aller/retour	WHAM	-0.66	2.41

TABLEAU XVIII. Résultats des calculs de différences d'énergies libres (kcal/mol) à l'aller et au retour effectués par les méthodes des perturbations (P.T.), d'intégration thermodynamique (I.T.) et de l'algorithme WHAM. Les résultats aller/retour de WHAM signifient que toutes les données pour les 44 fenêtres ont été analysées simultanément.

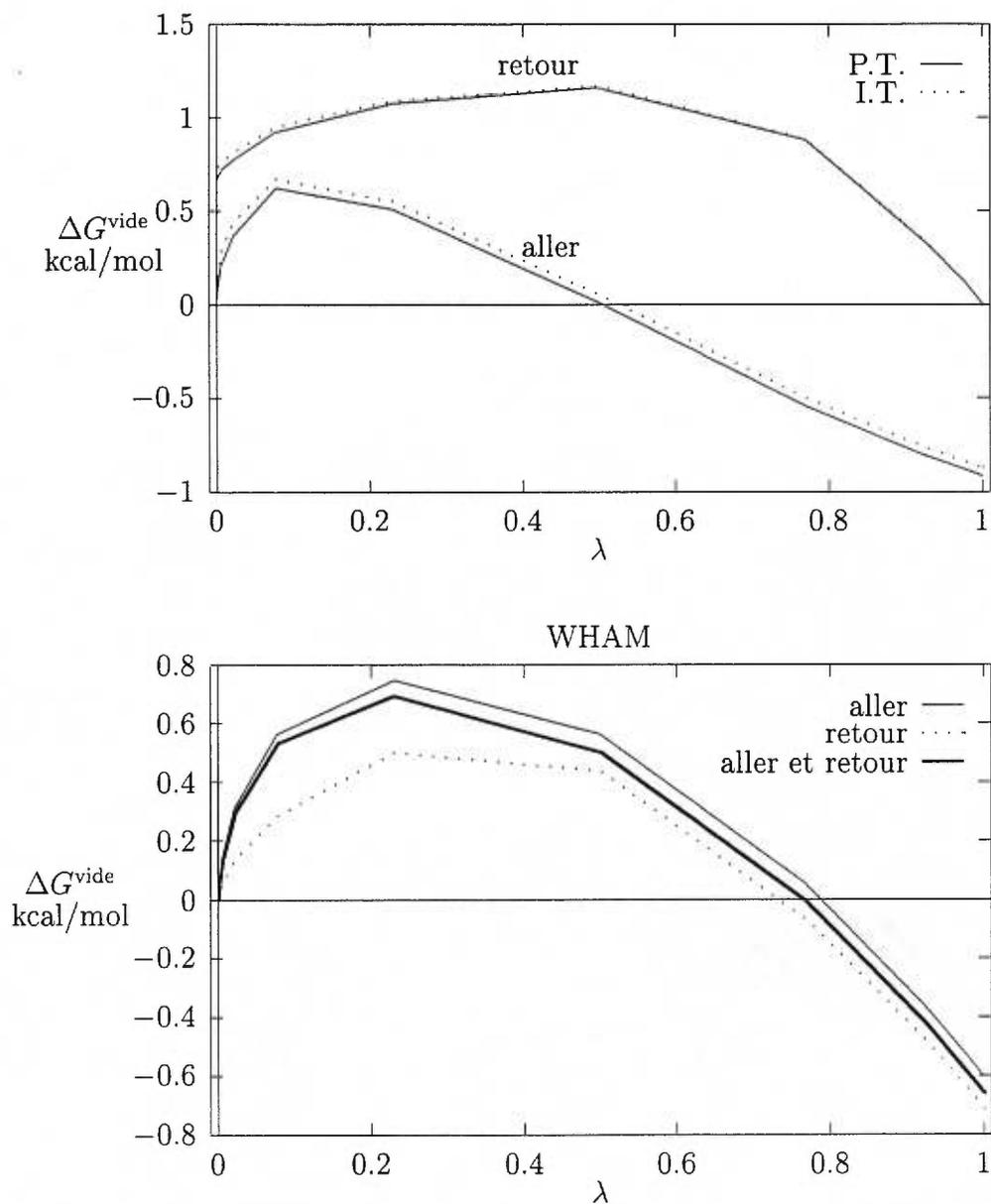


FIGURE 44. Variation de l'énergie libre  $\Delta G^{\text{vide}}$  lors des calculs de perturbations du tripeptide dans le vide en fonction de la valeur de  $\lambda$ . La courbe du haut correspond aux résultats des méthodes P.T. et I.T., et celle du bas aux résultats obtenus de l'algorithme WHAM en utilisant soit les données pour l'aller seulement, le retour seulement, ou les données aller-retour combinées.

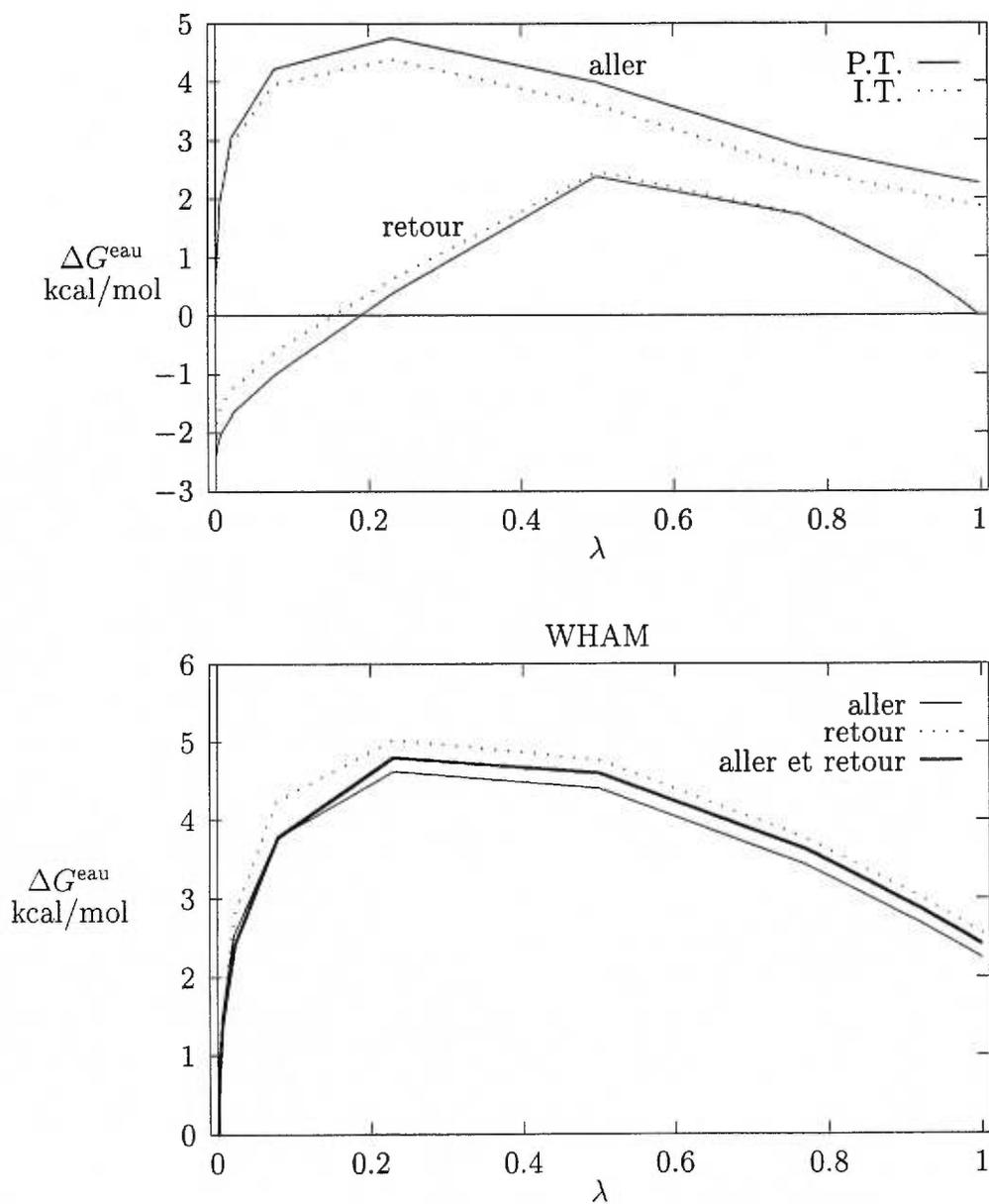


FIGURE 45. Voir figure 44. Lors de ces calculs, le tripeptide a été simulé dans 73 molécules d'eau.

Les différences d'énergies libres d'hydratation obtenues à partir de ces résultats sont données au tableau XIX et correspondent à la différence  $\Delta G_{\text{Ala} \rightarrow \text{Val}}^{\text{eau}} - \Delta G_{\text{Ala} \rightarrow \text{Val}}^{\text{vide}}$ . Le tableau XIX donne les valeurs calculées par les méthodes P.T., I.T. et WHAM. Il ressort que l'hydratation de la chaîne latérale de la valine dans un tripeptide

	$\Delta\Delta G_{\text{Ala} \rightarrow \text{Val}}^{\text{hydra.}}$ (kcal/mol)
P.T.	3.13
I.T.	2.64
WHAM	3.07

TABLEAU XIX. Différences d'énergies libres d'hydratation entre l'alanine et la valine obtenues par les méthodes des perturbations (P.T.), d'intégration thermodynamique (I.T.) et WHAM.

demanderait 3 kcal/mol supplémentaires d'énergie libre par rapport à une chaîne latérale d'alanine. Ceci est dû à la présence de deux groupements méthyles hydrophobes dans la valine. La déformation de l'environnement aqueux pour faire place à ces deux méthyles ne serait donc pas compensée par assez d'interactions favorables avec l'eau environnante. La valeur expérimentale la plus pertinente avec laquelle comparer est la différence d'énergie libre d'hydratation entre la chaîne latérale d'une alanine et d'une valine, obtenue par des expériences de transfert de méthane (pour la chaîne latérale de l'alanine) et du propane (pour la chaîne latérale de la valine) de la phase gaseuse à la phase aqueuse.<sup>149</sup> Cette valeur expérimentale est de 0.05 kcal/mol,<sup>149</sup> et indique un accord qualitatif de notre calcul, bien qu'il ne s'agisse pas de la même molécule, c'est-à-dire que l'on néglige les interactions entre la chaîne latérale et le reste du polypeptide. Un calcul de différence d'énergie libre d'hydratation par notre méthode entre le méthane et le propane pourrait établir la pertinence d'utiliser cette donnée expérimentale comme valeur de référence pour des calculs de mutations sur des acides aminés et des polypeptides.

Yamaotsu et coll.,<sup>150</sup> qui ont effectué cette mutation dans un seul acide aminé et qui procèdent par la méthode de croissance lente du paramètre de couplage  $\lambda$  (*Slow growth*) avec le logiciel AMBER 4.0,<sup>151</sup> obtiennent une valeur plus près de cette valeur expérimentale. Cette méthode consiste à faire varier le paramètre de couplage de façon continue (en fonction du temps) plutôt qu'à des valeurs discrètes comme il a été fait lors de nos calculs. Bien que cette méthode s'avère plus efficace dans certains cas, lorsqu'il s'agit de perturbations où des changements configurationnels considérables surviennent, il se peut que le système n'ait pas tout à fait le temps de s'équilibrer tout au long du changement du  $\lambda$  si le temps de simulation n'est pas assez long. On se retrouve avec un système qui est dans une configuration non représentative de  $H(\lambda)$ . Dans ces cas les méthodes de variations discrètes de  $\lambda$  s'avèrent plus avantageuses.<sup>67</sup> Néanmoins, il semble que ce ne soit pas le cas de Yamaotsu et coll.,<sup>150</sup> puisqu'ils obtiennent une valeur raisonnable de 1.57 kcal/mol. Une comparaison plus détaillée n'est pas possible puisque le calcul du  $\Delta\Delta G^{\text{hydra}}$  n'est fait qu'à partir des contributions d'énergies électrostatique et de van der Waals, ce qui semble incomplet.

Les résultats des calculs du tripeptide d'alanines donnent une valeur plus grande que la valeur expérimentale de la différence d'énergie libre d'hydratation entre la chaîne latérale d'une alanine et d'une valine. Mais la pertinence d'utiliser la différence d'énergie d'hydratation entre le méthane et le propane comme valeur de référence demeure à vérifier. De plus, cette valeur expérimentale n'est pas appuyée par un grand nombre d'autres données. Rien n'indique dans la procédure suivie qu'un mauvais protocole a été utilisé. Suite aux simulations sur le tripeptide d'alanine, un protocole très similaire a été utilisé pour effectuer la mutation V61A dans la calbindine D<sub>9k</sub>. Les résultats présentés dans la prochaine sous-section rendent compte de la première étape du lancement de ce calcul, mais ne constituent pas les résultats finaux qui nécessiteront une analyse exhaustive. Le bon déroulement des simulations (maintient des structures, convergence des calculs, etc.) est examiné, afin que ce projet puisse être poursuivi en reposant sur

une bonne base.

#### 6.4 Résultats et discussion de la mutation V61A dans la calbindine $D_{9k}$ P43G

Pour obtenir une estimation de l'effet de la mutation V61A sur la coopérativité, il est nécessaire d'effectuer un calcul de perturbations pour chacun des quatre états de la calbindine  $D_{9k}$ . Les structures de départ correspondent aux dernières structures générées à la fin des 200 ps de simulation dont les résultats ont été présentés au chapitre 4. Cependant, c'est la dernière structure simulée A(S<sub>II</sub>) qui a été utilisée pour ce calcul de perturbation sur l'état apo puisque cette dernière est en meilleur accord avec la structure expérimentale (voir entre autres le tableau IX et aussi la figure 30 pour l'aspect dynamique). Rappelons que chaque structure simulée est incluse dans une sphère de 2 285 molécules d'eau auxquelles est appliqué le potentiel à la frontière sphérique exposé à la section 2.2.2.<sup>61</sup>

Le protocole des calculs de perturbations pour le cas de la calbindine  $D_{9k}$  est le même que de celui du tripeptide d'alanine (voir section 6.3), sauf pour les quelques différences suivantes: tout d'abord, pour les structures qui correspondent aux  $\lambda = 0$  et  $\lambda = 1$ , les énergies ont été minimisées sur 100 pas de minimisation *steepest-descent*. Puis, s'ensuit une période d'équilibration de 10 ps avec un pas d'intégration de 0.002 ps. Rappelons qu'aucun rayon de coupure des interactions non-liées n'a été appliqué. Un coefficient de friction correspondant à un temps caractéristique de relaxation de  $5 \text{ ps}^{-1}$  a été appliqué à tous les atomes sauf à ceux d'hydrogène. Une constante de force de  $10 \text{ kcal}/(\text{mol} \cdot \text{Å}^2)$  a été appliquée au centre de masse de la protéine. De plus, une contrainte harmonique de force  $30 \text{ kcal}/(\text{mol} \cdot \text{rad}^2)$  a été appliquée à l'angle dièdre  $\chi$  (N, CA, CB, CG2) de la chaîne latérale de la Val61 dans l'état  $\lambda = 1$  (alanine invisible aux autres atomes) afin d'éviter toute réorientation de cette chaîne latérale que l'on voulait maintenir autour de  $180^\circ$  (valeur initiale). Comme aucune transition de la sorte n'était

survenue lors du calcul sur le tripeptide d'alanine, il ne s'est pas avéré nécessaire d'appliquer une telle contrainte.

Puis, dans les mêmes conditions, on simule une période d'équilibration de 1 ps pour chaque valeur de  $\lambda$  avant que ne débute la période de production de la trajectoire dont les structures échantillonnées serviront au calcul de perturbations. Chaque nouvelle valeur de  $\lambda$  est appliquée à la dernière configuration générée à la précédente valeur de  $\lambda$ . A chaque valeur de  $\lambda$ , la période d'équilibration a été suivie par une période de production de 3 ps pour un total de 1500 configurations échantillonnées pour le calcul de perturbation. Un total de 66 ps a été simulé dans chaque direction de  $\lambda$ .

Seuls les résultats obtenus par l'algorithme WHAM sont ici présentés, mais il a été vérifié que les résultats des méthodes P.T. et I.T. se comportaient normalement, c'est-à-dire que les changements d'énergie libre en fonction de  $\lambda$  étaient en accord raisonnable avec ceux obtenus de l'algorithme WHAM. Il faut mentionner que, contrairement au cas du tripeptide qui jouit d'une grande liberté de mouvements et peut explorer une surface d'énergie potentielle complexe, un acide aminé dans une protéine se voit restreint par l'empêchement stérique dû à la compacité des acides aminés, et devrait à ce point de vue être plus facile à échantillonner. Mais en contrepartie, la rigidité de l'environnement peut demander plus de temps d'équilibration afin que la cavité des atomes naissants se forme totalement. Les résultats de l'algorithme WHAM des calculs de perturbations (mutations Val61→Ala) effectués dans les quatre états d'occupation de la calbindine  $D_{9k}$  sont donnés aux figures 46 et 47 et montrent les changements des  $\Delta G_{\text{Val} \rightarrow \text{Ala}}^{\text{état } i}$  en fonction de  $\lambda$ . Alors que les structures simulées A( $S_{II}$ ), et particulièrement  $S_I(D)$  démontrent une disparité des résultats à l'aller et au retour (non convergence), les états  $S_{II}$  et D apparaissent pour leur part mieux convergés.

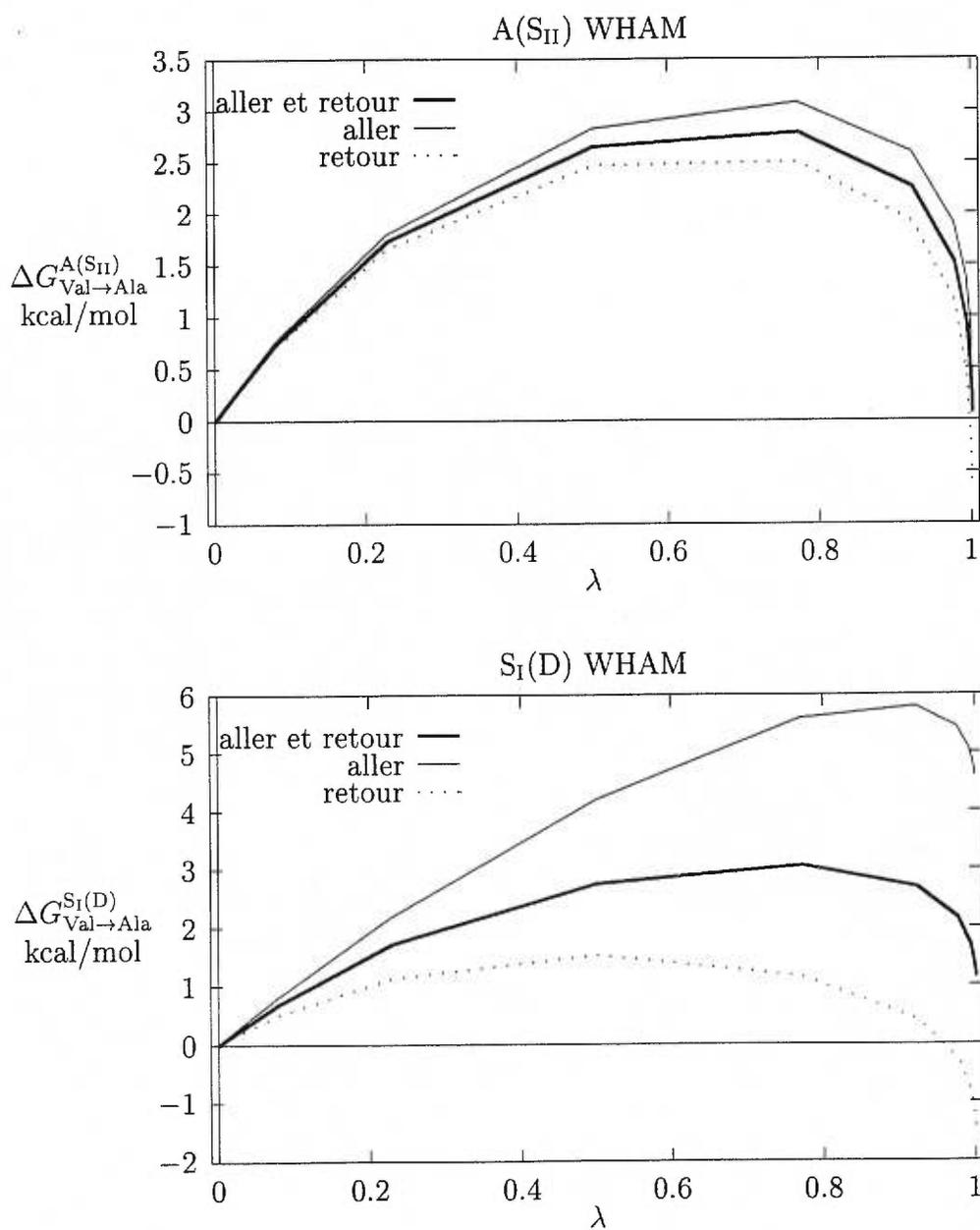


FIGURE 46. Variations de l'énergie libre en fonction de la valeur de  $\lambda$  pour la mutation effectuée dans la structure simulée A(S<sub>II</sub>) (haut) et S<sub>I</sub>(D) (bas).

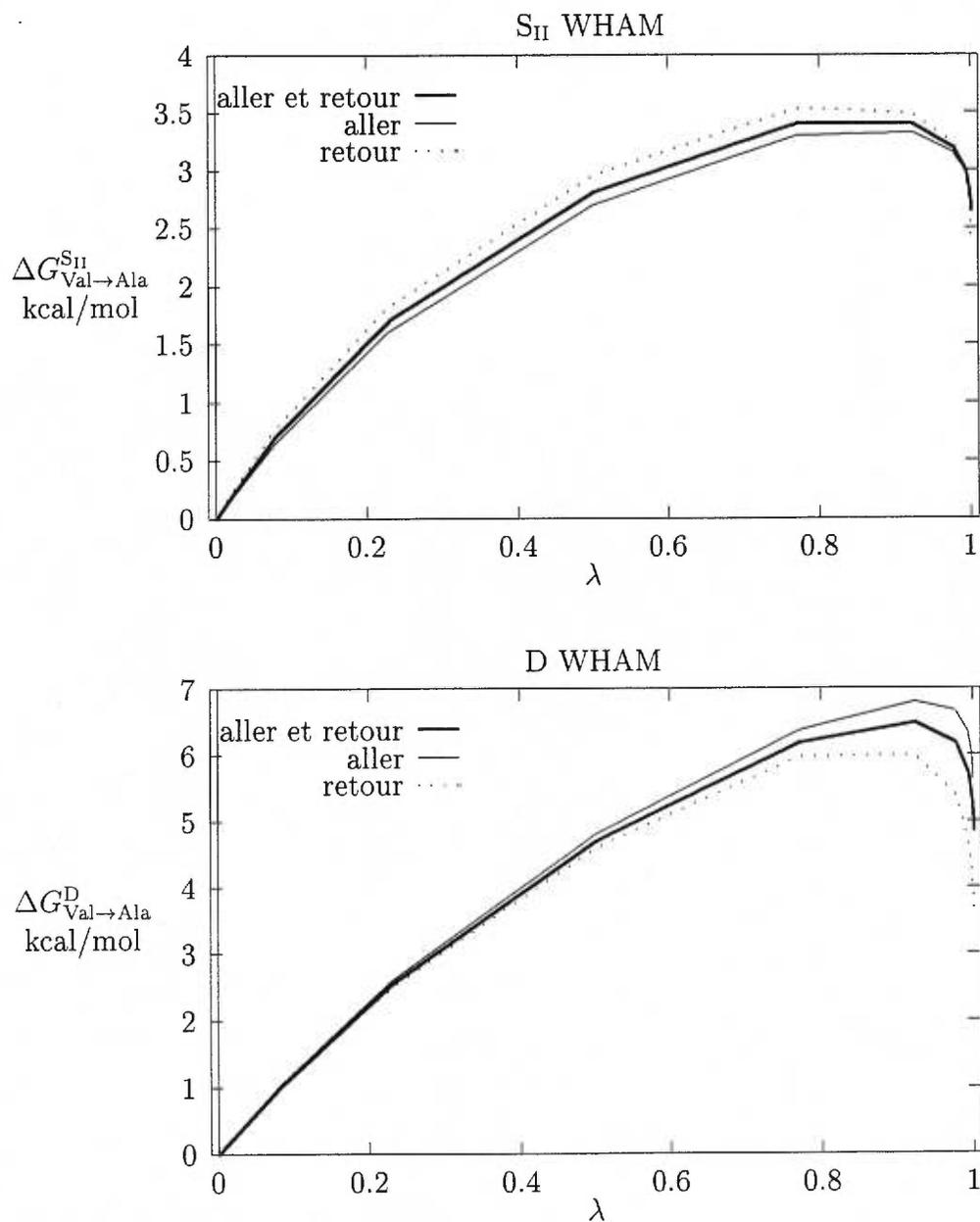


FIGURE 47. Variations de l'énergie libre en fonction de la valeur de  $\lambda$  pour la mutation effectuée dans l'état S<sub>II</sub> (haut) et D (bas).

Rappelons que la valine en position 61 fait partie du site II et que ce dernier est occupé par un ion  $\text{Ca}^{2+}$  dans les états  $S_{\text{II}}$  et D et que sa flexibilité s'en trouve grandement réduite. En d'autres mots, l'espace configurationnel du site II dans les états  $S_{\text{II}}$  et D est plus restreint dans ces deux cas et donc plus propice à une meilleure convergence à échantillonnage égal avec les états A et  $S_{\text{I}}(\text{D})$ .

Une comparaison visuelle des structures simulées  $S_{\text{I}}(\text{D})$  à la valeur de  $\lambda = 0.9$  entre les calculs à l'aller et au retour n'a révélé aucune différence de conformation majeure. Cette valeur de  $\lambda$  a été choisie dans la région où une grande différence entre les calculs à l'aller et au retour est ressortie. Comme une seule configuration a été comparée à une seule autre, il est difficile de préciser davantage la raison à l'origine de la moins bonne convergence pour cette structure. Le tableau XX donne la différences entre les structures simulées moyennes à  $\lambda = 0.9$  lors des calculs à l'aller et au retour. Ces différences de structure ont été calculées de la même façon qu'à la section 4.3.1.2 (équation 4.1). Ce tableau ne révèle pas non plus de différences pouvant expliquer la non convergence des calculs à l'aller et au retour. Le bon maintien de la valeur de l'angle dièdre  $\chi_{61}$  par la contrainte

		acides aminés							
		1-75	3-15	16-24	25-35	36-45	46-53	54-62	63-73
			hélice A	site I	hélice B	segment	hélice C	site II	hélice D
$S_{\text{I}}(\text{D})$	tout	1.29	0.85	0.83	0.95	0.90	0.84	0.76	0.91
	c.p.	1.01	0.52	0.56	0.37	0.73	0.51	0.45	0.45

TABLEAU XX. Différences structurales ( $\text{\AA}$ ) (équation 4.1) entre la structure simulée moyenne à la valeur de  $\lambda = 0.9$  à l'aller et avec la structure moyenne à la même valeur de  $\lambda$  au retour. Ces différences sont données pour chaque élément de structure secondaire. L'abréviation c.p. signifie chaîne principale; tous les résultats excluent les atomes d'hydrogène.

harmonique a aussi été vérifiée et s'est révélé efficace. La valeur de  $\chi_{61}$  la plus différente de  $180^\circ$  a été  $155^\circ$  et elle n'est survenue qu'une fois. Dans tous les autres

cas, la valeur s'est maintenue à  $180 \pm 8^\circ$ . Une explication possible demeure, tout comme dans le cas du tripeptide d'alanine, que les configurations échantillonnées diffèrent légèrement et que certaines transitions rares aient pu survenir dans une direction mais pas dans l'autre. Il a d'autre part été vérifié que les résultats après 3 ps de simulation de DM par fenêtre étaient mieux convergés qu'après la première picoseconde de simulation par fenêtre. De plus longues simulations devraient donc améliorer davantage ces résultats.

Le tableau XXI donne les résultats des changements d'énergies libres associés à la mutation V61A pour les quatre états d'occupation de la calbindine  $D_{9k}$ . Ces calculs suggèrent que la mutation V61A est défavorable dans tous les états

		$\Delta G_{\text{Val} \rightarrow \text{Ala}}^{\text{A}(S_{\text{II}})}$	$\Delta G_{\text{Val} \rightarrow \text{Ala}}^{\text{S}_I(\text{D})}$	$\Delta G_{\text{Val} \rightarrow \text{Ala}}^{\text{S}_{\text{II}}}$	$\Delta G_{\text{Val} \rightarrow \text{Ala}}^{\text{D}}$
WHAM	aller	0.66	4.60	2.80	5.63
	retour	-0.58	-1.40	2.45	3.68
	aller et retour	0.07	1.16	2.65	4.85

TABLEAU XXI. Résultats des calculs de différences d'énergies libres (kcal/mol) obtenus par l'algorithme WHAM. Les résultats totaux de WHAM signifient que toutes les données pour les 44 fenêtres ont été analysées simultanément.

d'occupation. De plus, la valeur absolue de cette énergie libre augmente en allant de l'état A vers les états  $S_I(\text{D})$ ,  $S_{\text{II}}$  et D. Selon l'équation 6.3, le changement d'énergie libre associé à la mutation et contribuant à la coopérativité ( $\Delta(\Delta\Delta G_{\text{Val} \rightarrow \text{Ala}}^{\text{coop}})$ ) est de 1.11 kcal/mol. La valeur expérimentale étant de 2.13 kcal/mol,<sup>122</sup> un accord qualitatif a donc été obtenu par la méthode de calcul que nous avons utilisée.

En substituant les résultats du tableau XXI dans les équations 6.4 et 6.5 on obtient les rapports de constantes macroscopiques  $K_1^{\text{Ala}}/K_1^{\text{Val}}$  et  $K_2^{\text{Ala}}/K_2^{\text{Val}}$ . Ces résultats sont donnés au tableau XXII. Ces rapports indiquent l'effet de la mutation V61A sur les deux constantes macroscopiques de liaison du  $\text{Ca}^{2+}$ , en

équation	$K_1^{\text{Ala}}/K_1^{\text{Val}}$	$K_2^{\text{Ala}}/K_2^{\text{Val}}$
6.4(a) (S <sub>I</sub> -A)	0.162	—
6.4(b) (S <sub>II</sub> -A)	0.013	—
6.5(a) (D-S <sub>I</sub> )	—	0.002
6.5(b) (D-S <sub>II</sub> )	—	0.025
exp. <sup>122</sup>	1.580	0.030

TABLEAU XXII. Rapports des constantes macroscopiques entre les calbindines D<sub>9k</sub> P43G (Val) et P43G-V61A (Ala).

ayant fait l'hypothèse que  $k_I = k_{II}$  et  $k_{I,II} = k_{II,I}$ . Cela implique aussi que des résultats identiques devraient être obtenus pour les équations 6.4 (a) et (b), de même que pour les équations 6.5 (a) et (b). Nos calculs ne donnent cependant pas des résultats identiques pour les deux voies de liaisons et suggèrent que la mutation V61A diminue considérablement les deux constantes macroscopiques  $K_1$  et  $K_2$ . Les données expérimentales<sup>122</sup> (exp. dans le tableau) indiquent que la constante  $K_1$  augmente par un facteur 1.58 et que la constante  $K_2$  diminue par un facteur 0.03. Il est intéressant de remarquer que les résultats des calculs de perturbations ayant les mieux convergés (S<sub>II</sub> et D) sont ceux qui s'accordent bien avec les données expérimentales et que le moins bon accord est obtenu pour les rapports faits à partir des calculs de perturbations des structures A(S<sub>II</sub>) et S<sub>I</sub>(D) qui ne sont pas complètement convergés. Mais puisque les mutants finaux ne sont pas les mêmes dans les simulations (P43G-V61A) et l'expérience (P43M-V61A), effectuer davantage de comparaisons ne serait pas approprié.

Bien qu'une analyse comparative finale ne puisse être effectuée étant donnée la durée encore insuffisante de la simulation, il est intéressant d'examiner quels sont les atomes lourds qui se trouvent dans l'environnement de la chaîne latérale affectée par la mutation. Les tableaux XXIII et XXIV donnent les listes des distances moyennes de quelques atomes lourds les plus près des atomes lourds

de la chaîne latérale de la valine ( $\lambda = 0$ ) ou de l'alanine ( $\lambda = 1$ ). Les distances moyennes dans le cas de la valine ont été calculées à partir des trajectoires de 200 ps analysées au chapitre 4. Pour le cas de l'alanine, les configurations utilisées sont les 20 récoltées lors de la simulation de 3 ps de la structure à la valeur de  $\lambda = 1$  préalablement équilibrée pendant 10 ps.

	A(S <sub>II</sub> )	S <sub>I</sub> (D)	S <sub>II</sub>	D
CB(Val61)-O(Leu23)	3.78±0.42	3.62±0.23	3.78±0.23	3.86±0.24
CB(Val61)-CE1(Phe50)	5.38±0.41	4.99±0.51	5.73±0.76	4.68±0.49
CB(Val61)-CE2(Phe50)	6.42±0.90	6.55±0.72	5.27±0.48	5.89±0.86
CB(Val61)-CZ(Phe50)	5.46±0.53	5.56±0.57	5.03±0.40	5.07±0.55
CG2(Val61)-CE1(Phe50)	4.10±0.44	3.83±0.45	4.45±0.67	3.80±0.42
CG2(Val61)-CE2(Phe50)	5.17±0.94	5.20±0.72	4.07±0.56	4.79±0.82
CG2(Val61)-CZ(Phe50)	4.23±0.55	4.35±0.52	3.84±0.35	4.17±0.50

TABLEAU XXIII. Distances moyennes ( $\text{\AA}$ ) entre les atomes lourds de la chaîne latérale de la valine et les atomes lourds les plus proches appartenant à d'autres acides aminés non adjacents. ( $n = 20000$  configurations utilisées pour ces calculs.)

	A(S <sub>II</sub> )	S <sub>I</sub> (D)	S <sub>II</sub>	D
CB(Ala61)-O(Leu23)	3.58±0.19	3.79±0.34	3.69±0.19	4.03±0.23
CB(Ala61)-CE1(Phe50)	4.47±0.26	4.65±0.38	5.93±0.56	5.03±0.43
CB(Ala61)-CE2(Phe50)	6.20±0.37	6.49±0.41	4.57±0.23	3.74±0.17
CB(Ala61)-CZ(Phe50)	4.93±0.33	5.33±0.43	4.84±0.35	4.01±0.35

TABLEAU XXIV. Distances moyennes ( $\text{\AA}$ ), dans les mutants V61A entre les atomes lourds de la chaîne latérale de l'alanine (Ala61) et les atomes lourds les plus proches appartenant à d'autres acides aminés non adjacents. ( $n = 20$  configurations utilisées pour ces calculs.)

Les deux acides aminés les plus proches de la valine ou de l'Ala61 sont dans tous les cas la Leu23 qui fait partie du site I, ainsi que le Phe50 situé dans l'hélice C. La figure 48 illustre ces deux acides aminés et spécifie la nomenclature utilisée aux tableaux XXIII et XXIV. En détail, l'atome d'oxygène de la chaîne principale

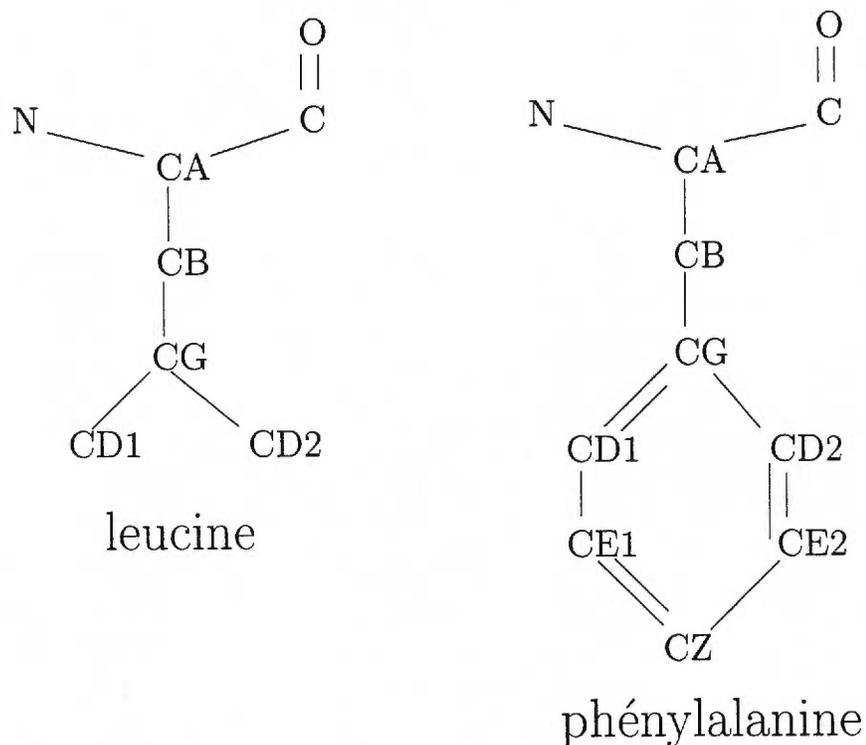


FIGURE 48. Représentation des acides aminés les plus proches des atomes soit de la valine ou de l'alanine et nomenclature à utiliser pour lire les tableaux XXIII et XXIV.

de la Leu23 (qui est impliqué dans un pont hydrogène avec l'atome d'azote de la Val61) demeure celui qui est le plus proche du CB de la Val61 ou de l'Ala61 dans tous les cas sauf dans la structure D où les atomes CG2(Val61)-CE1(Phe50) sont les plus rapprochés. Ces deux atomes sont aussi les suivants les plus rapprochés dans les structures A(S<sub>II</sub>), S<sub>I</sub>(D) et S<sub>II</sub>. Comme dans le mutant V61A il n'y a pas d'atomes CG1 et CG2, il faut réévaluer ces distances en vérifiant la compacité des atomes lourds autour du CB de l'Ala61. Bien que nos résultats ne soient que préliminaires, le tableau XXIV montre que l'assemblage des atomes n'est pas

aussi serré autour du CB dans la structure du mutant V61A qu'autour du CG2 dans la structure non mutée. Mais bien sûr, ces résultats restent à confirmer avec des simulations de DM plus longues à chaque fenêtre. Les résultats du tableau XXIII montrent aussi que les distances entre les atomes lourds demeurent assez stables, mais il semble que les distances impliquant l'atome CE2 du Phe50 aient de plus grandes fluctuations. Encore une fois, il faudra plus de données pour le mutant V61A pour effectuer des comparaisons valables.

La figure 49 donne un exemple de l'évolution temporelle de la distance moyenne entre les atomes CG2(Val61) et CE1(Phe50) tout au long de la simulation de 200 ps (A, S<sub>I</sub>(D) et S<sub>II</sub>) ou de 400 ps (D). Cette distance moyenne dans l'état D demeure stable pendant les premières 150 ps de simulation mais augmente par la suite et fluctue autour d'une valeur moyenne plus élevée (autre conformation). Pour la structure simulée S<sub>II</sub>, qui a l'écart type de la distance le plus élevé, la figure 49 montre qu'il y aurait échantillonnage alternatif de plusieurs conformations plutôt que d'une seule, comme c'est le cas pour la structure simulée S<sub>I</sub>(D). Pour la structure simulée A(S<sub>II</sub>), on remarque que suite aux premières 10 ps de simulations, il y a un changement de distance qui dure environ 60 ps, cette dernière revenant par la suite autour de sa valeur initiale. Les calculs d'énergies libres ayant le mieux convergés pour l'état S<sub>II</sub>, un argument plausible serait qu'une bonne flexibilité non associée à des changements conformationnels peu fréquents est plus favorable à ce genre de calculs. Par contre, une plus grande rigidité de l'environnement demandera plus de temps avant d'échantillonner toutes les configurations. Il serait intéressant de comparer ces résultats avec les homologues dans le mutant V61A de la calbindine D<sub>9k</sub> pour le ou les quelques atomes lourds les plus près du CB. L'évolution temporelle de ces distances pourraient révéler une partie de l'origine des résultats mais cela suppose un temps de simulation au moins aussi long pour les quatre structures à la valeur de  $\lambda = 1$ .

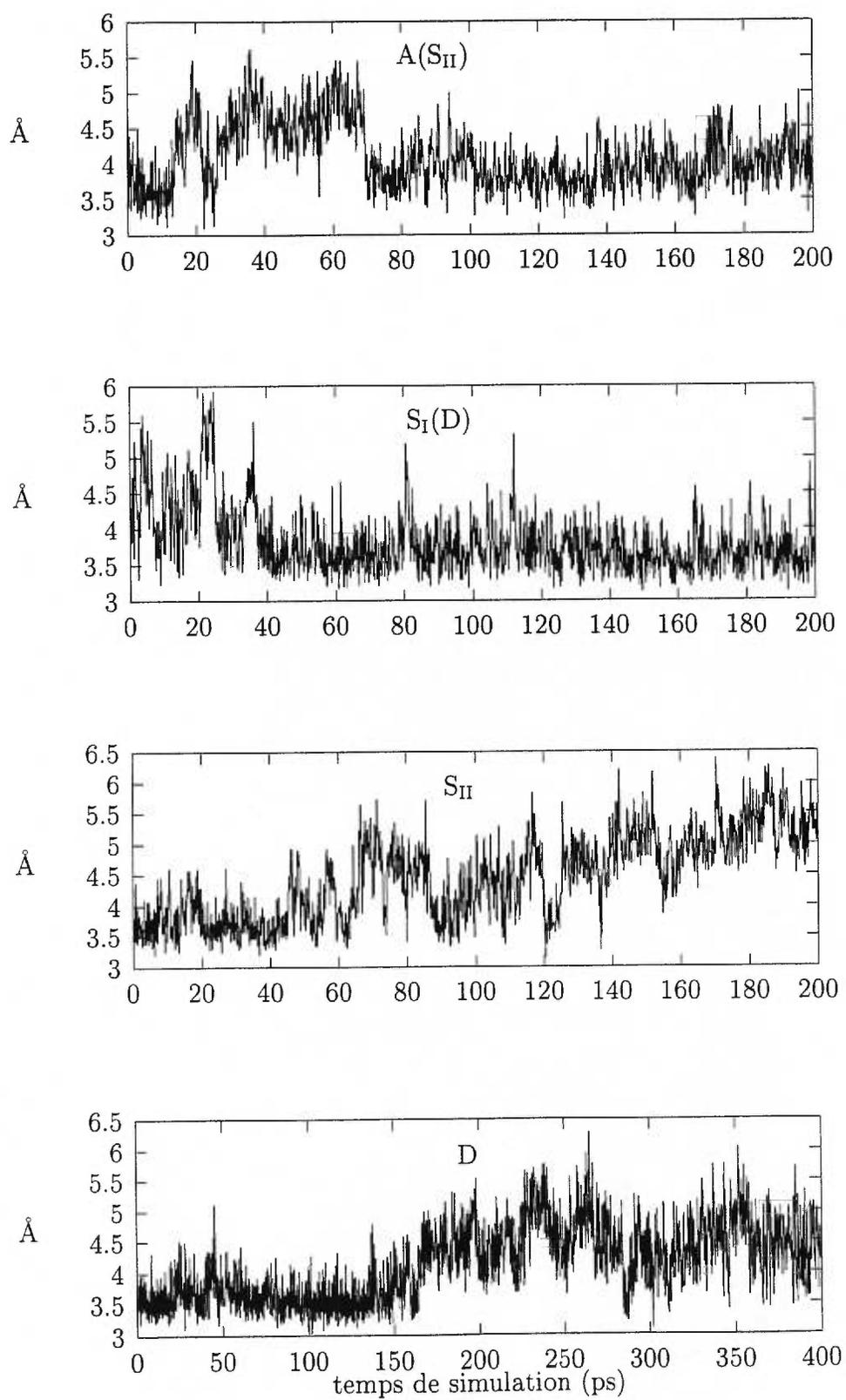


FIGURE 49. Évolution de la distance moyenne CG2(Val61)-CE1(Phe50) en fonction du temps de simulation (ps) (L'état D compte 400 ps).

Afin d'avoir une idée de l'emplacement des molécules d'eau dans l'environnement de la mutation, la figure 50 montre les fonctions de distributions radiales des atomes d'oxygène des molécules d'eau entourant la chaîne latérale de la Val61 dans tous les états d'occupation de la calbindine  $D_{9k}$ . Comme on peut le constater,

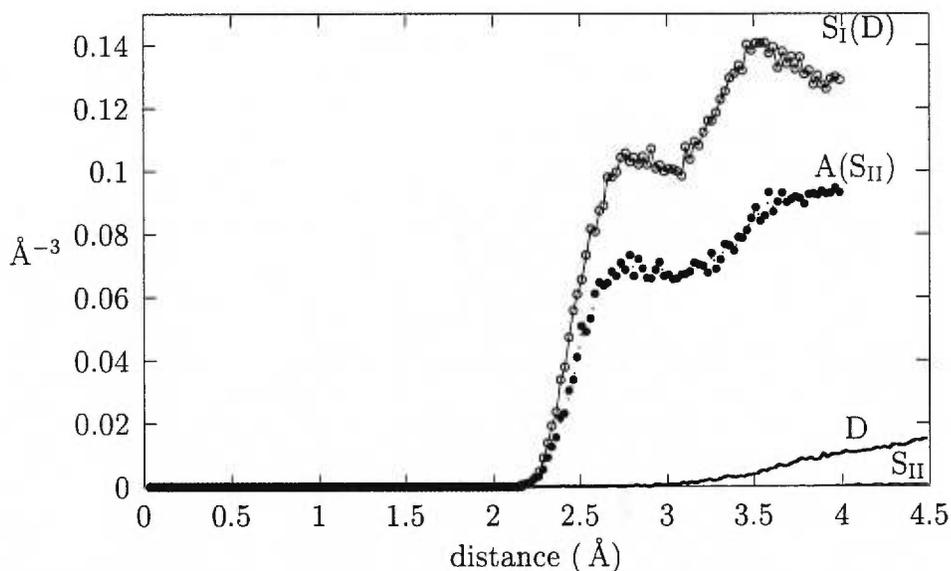


FIGURE 50. Distributions radiales  $\text{\AA}^{-3}$  des atomes O (de  $\text{H}_2\text{O}$ ) en fonction de la distance à partir du centre de masse de la chaîne latérale de la Val61, calculées à partir des trajectoires de DM analysées au chapitre 4.

il y a peu d'eau aux environs (jusqu'à 4  $\text{\AA}$ ) de la chaîne latérale de la Val61 dans les structures simulées  $A(S_{II})$  et  $S_I(D)$ , et pas du tout dans les structures simulées  $S_{II}$  et  $D$  à cause de la présence du  $\text{Ca}^{2+}$  dans le site II. On peut aussi se référer à la figure 25 pour confirmer qu'il y a moins qu'une molécule d'eau dans un rayon de 4  $\text{\AA}$  autour de cette chaîne latérale. La mutation implique donc à un premier degré le réarrangement d'atomes environnants qui appartiennent à d'autres acides aminés ou peut-être leur remplacement par des molécules d'eau. Bien que les fonctions de distributions radiales ne soient pas présentées en raison du nombre insuffisant de configurations, on a pu tout de même observer qu'une présence accrue de molécules d'eau autour de la chaîne latérale de l'Ala61.

## 6.5 Conclusion

Les bases d'une méthode et d'un protocole de simulations de DM ont été développées afin de calculer la différence d'énergie libre associée à la mutation V61A dans la calbindine  $D_{9k}$ . En plus de la construction des angles dièdres rattachant la chaîne latérale de l'Ala61 sur la chaîne principale de la calbindine  $D_{9k}$ , cette mutation implique surtout la croissance et la décroissance simultanées des charges et des paramètres de Lennard-Jones des atomes des deux chaînes latérales appartenant aux deux mutants. Cela nécessite des temps de simulations de DM considérables à cause de la complexité des changements au niveau atomique. Le but de ce chapitre était de présenter des résultats préliminaires afin de rendre compte du bon déroulement des calculs selon le protocole développé et d'illustrer la faisabilité de ce type de calculs difficiles.

Les calculs de perturbations faits en variant le paramètre de couplage dans les deux directions ( $\lambda = 0 \rightarrow 1$  et  $\lambda = 1 \rightarrow 0$ ) n'apparaissent pas complètement convergés à ce point-ci. Cette difficulté semble être reliée à l'absence de  $Ca^{2+}$  dans le site II (états A et  $S_I(D)$ ) qui implique une flexibilité accrue de ce site où est effectuée la mutation (voir les figures 28, 29 et 30 à la section 4.3.2). Une plus grande flexibilité donne lieu à une surface d'énergie potentielle plus difficile à échantillonner. En effet, les calculs apparaissent moins bien convergés dans les cas des structures simulées A( $S_{II}$ ) et  $S_I(D)$  que pour les structures simulées  $S_{II}$  et D. Il a été vérifié qu'il n'y avait pas de différence conformationnelle majeure entre les configurations échantillonnées à l'aller et au retour des calculs de perturbations.

Après avoir effectué 3 ps de simulations de DM à chacune des 22 valeurs découpées du paramètre de couplage  $\lambda$ , la valeur calculée de la différence d'énergie libre associée à la mutation V61A dans la calbindine  $D_{9k}$ ,  $\Delta(\Delta\Delta G_{Val \rightarrow Ala}^{coop.}) = 1.11$  kcal/mol, se révèle en accord qualitatif avec la donnée expérimentale (2.13 kcal/mol).<sup>122</sup> De plus, les rapports de constantes macroscopiques entre la protéine mutée et la

protéine sauvage ont été évalués et un bon accord qualitatif est obtenu à partir des calculs de perturbations qui ont le mieux convergés ( $S_{II}$  et D). Ces rapports  $K_1^{Ala}/K_1^{Val}$  et  $K_2^{Ala}/K_2^{Val}$  ont été respectivement évalués à 0.013 et 0.025 dans les meilleures simulations, en comparaison à 1.58 et 0.03 obtenus expérimentalement.

Bien qu'il soit trop tôt pour conclure de façon définitive, il semble que l'organisation des atomes autour de la chaîne latérale mutée (alanine) ne soit pas aussi compacte que dans la calbindine  $D_{9k}$  P43G originale. Les acides aminés les plus proches de la Val61 à muter sont la Leu23 faisant partie du site I ainsi que le Phe50. La position de ces acides aminés sera à examiner dans l'analyse finale de ces calculs. La présence de l'eau dans l'environnement des chaînes latérales de la mutation devra aussi être examinée.

On peut donc présumer que des bases rigoureuses pour le calcul de la coopérativité dans liaison du  $Ca^{2+}$  ont été établies, mais des temps plus longs de simulations devront être accumulés à chaque fenêtre de perturbation afin de mieux comprendre l'origine des changements à la coopérativité. Une étude plus en profondeur du mutant P43G-V61A pourra être poursuivie. Entre autres, avec l'accumulation de quelques 200 ps de simulations du mutant P43G-V61A (à la valeur  $\lambda = 1$ ), un histogramme tel celui présenté à la figure 32 pourrait être compilé afin d'évaluer les différentes contributions qui amènent une diminution de la coopérativité de la calbindine  $D_{9k}$ .

# CHAPITRE 7

## Conclusion

### 7.1 Rappel du but de la thèse

Le but de cette thèse était d'étudier divers aspects des conséquences de la liaison des ions  $\text{Ca}^{2+}$  et de la coopérativité chez la calbindine  $\text{D}_{9k}$  par des méthodes de modélisation moléculaire. A cette fin, on doit disposer d'un bon modèle pour le système biomoléculaire et pour les méthodes de simulation. Si tel est le cas, ces méthodes permettent d'effectuer une analyse au niveau atomique et d'évaluer les différentes contributions qui pourraient jouer un rôle dans la coopérativité.

Le cas particulier de la calbindine  $\text{D}_{9k}$  implique la liaison de deux ions  $\text{Ca}^{2+}$  dont chacun est chargé  $+2e$ . Ceci ramène d'emblée au fait que la fonction d'énergie potentielle utilisée pour nos simulations (CHARMM<sup>56</sup>) calcule les forces par paires d'atomes, ce qui implique que les effets à plusieurs corps ou non-additifs sont négligés. C'est le cas dans la plupart des simulations effectuées jusqu'à maintenant. Par contre, nos simulations sont avantagées par le fait qu'il n'y a aucun rayon de coupure imposé aux interactions non-liées. Ceci est particulièrement important pour les interactions électrostatiques à longue portée, surtout pour notre cas où deux ions chargés  $+2e$  ne sont séparés que par 12 Å.

La première partie de cette thèse consistait à déterminer des paramètres de Lennard-Jones pour le  $\text{Ca}^{2+}$  et d'évaluer la faisabilité d'inclure un terme explicite d'énergie qui tiendrait compte des interactions à plusieurs corps entre le  $\text{Ca}^{2+}$  et ses ligands dans les sites de la calbindine  $\text{D}_{9k}$ . La deuxième partie de cette thèse

consiste à utiliser les modèles de la calbindine  $D_{9k}$  et de simulations de dynamique moléculaire afin d'étudier la structure et la dynamique de la protéine en fonction du niveau d'occupation du  $Ca^{2+}$ . Nos résultats y sont comparés aux données expérimentales homologues afin d'évaluer la qualité des modèles utilisés et des observations sont faites au sujet de la coopérativité.

## 7.2 Contributions principales

Des paramètres de Lennard-Jones pour le  $Ca^{2+}$  ont été déterminés afin de reproduire l'énergie libre d'hydratation du  $Ca^{2+}$  et des propriétés statiques telles la structure immédiate des molécules d'eau autour du  $Ca^{2+}$  et la coordinence du  $Ca^{2+}$ . Les courbes d'énergie d'interaction  $Ca^{2+}$ -ligand calculées par des méthodes *ab initio* (HF) ne sont cependant pas reproduites de façon identique, étant donné que les charges attribuées aux atomes des acides aminés du champ de force *param22*<sup>51</sup> sont fixes, ce qui ne laisse pas la possibilité aux nuages électroniques de se réorganiser en présence d'ions chargés  $+2e$ .

Les contributions non-additives dans les interactions  $Ca^{2+}$ - $(H_2O)_n$  ont été évaluées par des méthodes *ab initio* (HF) puis comparées aux calculs d'énergies d'interaction effectués avec *param22*. Cette contribution non-additive reliée aux effets à plusieurs corps compte, par exemple, pour autant que 32 kcal/mol dans l'interaction  $Ca^{2+}$ - $(H_2O)_6$ . Cette contribution non-additive jouant un rôle dans les propriétés structurales et dynamiques des liquides,<sup>74, 99</sup> et sachant qu'elle est bien représentée par l'énergie de polarisation induite au second ordre pour les interactions  $Ca^{2+}$ - $H_2O$ <sup>94</sup> et  $Ca^{2+}$ -peptide,<sup>95</sup> des polarisabilités pour les atomes et groupements chimiques de molécules représentant les ligands du  $Ca^{2+}$  dans les sites de la calbindine  $D_{9k}$  ont été déterminées. Afin de déterminer ces polarisabilités, une méthode approximative et simple a été élaborée dans laquelle seuls les atomes lourds de molécules modèles d'acides aminés peuvent être polarisés par une charge ponctuelle. Les résultats obtenus s'accordent raisonnablement bien avec les données

provenant d'expériences ou d'autres calculs.

Ensuite, une tentative d'inclure les termes d'énergie de polarisation au premier et second ordre pour le cas du  $\text{Ca}^{2+}$  hydraté a été accomplie et lors de laquelle les paramètres déterminés dans cette thèse ont été utilisés. Cette tentative démontre la grande sensibilité des résultats aux paramètres et suggère fortement l'existence de difficultés inhérentes au développement d'une fonction d'énergie potentielle qui tiendrait compte des effets de polarisation induite. Devant l'ampleur du travail qu'impliquerait la reparamétrisation d'un champ de forces incluant un terme d'énergie attribuable aux effets à plusieurs corps, il a été préféré d'utiliser tel quel le champ de force *param22* pour les protéines, d'approximer que le  $\text{Ca}^{2+}$  était une sphère de charge  $+2e$  et que ses ligands n'étaient pas polarisables.

Des simulations de DM de 200 ps ont été produites pour les quatre états d'occupation de la calbindine  $\text{D}_{9k}$ , avec une simulation supplémentaire de l'état A pour un total de cinq trajectoires produites (A, A( $\text{S}_{\text{II}}$ ),  $\text{S}_{\text{I}}$ (D),  $\text{S}_{\text{II}}$  et D). Les structures simulées  $\text{S}_{\text{II}}$  et D sont demeurées structurellement proches de leur structure expérimentale de départ, tandis que la structure simulée A s'en est éloignée plus considérablement. La structure simulée A( $\text{S}_{\text{II}}$ ), construite à partir de la structure expérimentale  $\text{S}_{\text{II}}$  de laquelle on a enlevé le  $\text{Ca}^{2+}$  demeure davantage proche de la structure expérimentale A. Les différences structurales entre structures simulées et expérimentales se concentrent surtout dans les bouts terminaux, les sites ainsi que dans le segment 36-45 qui relie les deux mains EF. Les hélices conservent en général un très bon accord avec leur structure expérimentale, sauf pour l'hélice C de la structure simulée A qui dévie le plus, mais dont la position s'avère mal définie dans les données expérimentales RMN.<sup>32</sup>

Les différences de structure entre états simulés se concentrent surtout dans les sites et le segment 36-45. Mais ces régions de la protéine sont celles qui sont déterminées expérimentalement avec le moins de certitude.<sup>32, 33</sup> Pour l'aspect structural, les résultats des simulations s'accordent qualitativement bien avec les

résultats expérimentaux.<sup>25, 33, 32</sup> Les simulations effectuées à partir de structures déterminées par RMN (A, A(S<sub>II</sub>), S<sub>I</sub>(D) et S<sub>II</sub>) semblent en moins bon accord avec les données expérimentales que la simulation de l'état D dont la structure a été déterminée par des expériences de diffraction des rayons X à haute résolution. Les états liés au Ca<sup>2+</sup> sont structurellement similaires entre eux et différents de l'état A, signifiant que la majorité des changements structuraux se produit lors de la liaison du premier Ca<sup>2+</sup>. La différence considérable entre les structures simulées S<sub>I</sub>(D) et S<sub>II</sub> représentent peut-être l'asymétrie du mécanisme de la coopérativité selon que l'ion se lie d'abord au site I ou au site II. La liaison du Ca<sup>2+</sup> dans le site II des structures simulées S<sub>II</sub> et D déshydrate les chaînes latérales du site II et leur chaîne principale dans une certaine mesure, et il semble qu'un certain effet soit aussi ressenti dans le site I et consiste à déshydrater d'environ une molécule H<sub>2</sub>O la chaîne principale du site I encore inoccupé.

Pour l'aspect dynamique, les études des fluctuations des angles dièdres  $\phi$  et  $\psi$ , ainsi que les comparaisons des paramètres d'ordre des vecteurs N-H calculés à partir des trajectoires montrent que les bouts terminaux, les sites et le segment 36-45 sont les éléments qui fluctuent le plus et que ces fluctuations diminuent avec le niveau d'occupation du Ca<sup>2+</sup> dans la protéine. Les liaisons successives du Ca<sup>2+</sup> rendent la structure plus rigide, avec la majeure partie des effets concentrée dans la liaison d'un premier ion au site I ou au site II. Ces résultats sont en bon accord qualitatif avec les données expérimentales.<sup>24, 28, 33, 39</sup>

Une mesure de la contribution électrostatique à la coopérativité indique que cette dernière est défavorable par +3.0 kcal/mol au processus global de coopérativité (-1.84 kcal/mol), suggérant que la contribution du réarrangement de structure et de dynamique de la protéine devrait être d'environ -4.84 kcal/mol. Les interactions hydrophobes pourraient jouer un rôle important dans la coopérativité, tel que suggéré par les distributions des énergies d'interactions entre la chaîne latérale de la Val61 et le reste de la protéine pour laquelle on retrouve une interaction

de plus en plus favorable avec la liaison de  $\text{Ca}^{2+}$ . Les distributions des ponts hydrogènes du feuillet- $\beta$  entre les deux sites ont révélé que les deux ponts sont présents dans toutes les structures simulées sauf A, et qu'il y avait raffermissement du pont hydrogène Leu23-HN $\cdots$ O-Val61 suite à la liaison du  $\text{Ca}^{2+}$ .

L'origine des déviations quantitatives des données expérimentales peut provenir des imperfections de la fonction d'énergie potentielle de CHARMM<sup>56</sup> (ex: absence de polarisation induite) ou d'un échantillonnage insuffisant de l'espace configurationnel. Il est également possible que des incertitudes dans les structures de départ contribuent à déformer les structures simulées et à modifier la dynamique.

Un modèle analytique a été développé afin d'exprimer la partie librationnelle des paramètres d'ordre des vecteurs N-H en fonction de l'amplitude des librations effectuées par le vecteur N-H. De plus, ces librations sont exprimées en termes des fluctuations des angles dièdres  $\phi$  et  $\psi$  (bornant le plan amide) que l'on propose anticorrélés. Pour les paramètres d'ordre qui s'accordent bien avec les données expérimentales et avec le modèle, on trouve des librations moyennes autour de  $9^\circ$ , ce qui est raisonnablement en accord avec d'autres auteurs<sup>130, 132</sup> qui ont trouvé des librations de  $10\text{-}12^\circ$  communes à tous les acides aminés. Les résultats suggèrent aussi que des fluctuations trop grandes des angles dièdres  $\phi$  et  $\psi$  représentent en partie la source des trop faibles valeurs calculées pour certains paramètres d'ordre. Ainsi, une surestimation des librations par seulement quelques degrés est suffisante pour expliquer le désaccord avec les données expérimentales.

Des bases rigoureuses d'une méthode et d'un protocole de simulations de DM ont été développées afin de calculer la différence d'énergie libre associée à la mutation V61A dans la calbindine D<sub>9k</sub>. Le but de ce chapitre était de présenter des résultats préliminaires afin de rendre compte du bon déroulement des calculs selon le protocole développé et d'illustrer la faisabilité de ce type de calculs difficiles. Les calculs de perturbations faits en variant le paramètre de couplage dans les

deux directions ( $\lambda = 0 \rightarrow 1$  et  $\lambda = 1 \rightarrow 0$ ) n'apparaissent pas complètement convergés à ce point-ci, mais après avoir effectué 3 ps de simulations de DM à chacune des 22 valeurs de  $\lambda$ , la valeur calculée de la différence d'énergie libre associée à la mutation V61A est de  $\Delta(\Delta\Delta G_{\text{Val} \rightarrow \text{Ala}}^{\text{coop}}) = 1.11$  kcal/mol, et se révèle en accord qualitatif avec la donnée expérimentale (2.13 kcal/mol).<sup>122</sup> Les rapports de constantes macroscopiques entre la protéine mutée et la protéine sauvage ont été évalués et un bon accord qualitatif est obtenu à partir des calculs de perturbations qui ont le mieux convergés ( $S_{\text{II}}$  et D). Ces rapports  $K_1^{\text{Ala}}/K_1^{\text{Val}}$  et  $K_2^{\text{Ala}}/K_2^{\text{Val}}$  ont été respectivement évalués à 0.013 et 0.025 dans les meilleurs simulations, en comparaison à 1.58 et 0.03 obtenus expérimentalement. Il semble que l'organisation des atomes autour de la chaîne latérale mutée (alanine) ne soit pas aussi compacte que dans la calbindine D<sub>9k</sub> P43G originale. Des temps plus longs de simulations devront être accumulés à chaque fenêtre de perturbation afin de mieux comprendre l'origine des changements à la coopérativité.

### 7.3 Suggestions pour de futures simulations

Parmi les améliorations qui pourraient être apportées au modèle on notera l'inclusion dans la fonction d'énergie potentielle, d'un terme d'énergie de polarisation induite afin de représenter les interactions à plusieurs corps. Cela impliquerait probablement la reparamétrisation d'une grande partie du champ de forces *param22*.<sup>51</sup> Ensuite, des simulations de DM de plusieurs centaines de ps devraient être effectuées afin d'observer la qualité des simulations par rapport aux données expérimentales. En particulier, la coordinence du  $\text{Ca}^{2+}$  devrait être examinée afin de confirmer que les ligands supplémentaires dans les simulations D et  $S_{\text{I}}(\text{D})$  peuvent être attribués au manque de polarisation induite, de même que la structure des sites par rapport aux structures expérimentales. Les paramètres d'ordre des vecteurs N-H dont les valeurs calculées étaient trop faibles par rapport aux données expérimentales devraient être analysés à l'aide du modèle des librations afin d'observer des

différences dans les fluctuations des angles  $\phi$  et  $\psi$ .

Un raffinement de la structure expérimentale A (RMN) serait aussi à envisager afin d'améliorer sa qualité et d'avoir une meilleure structure de référence pour des simulations ultérieures. Si possible, on devrait aussi obtenir une structure expérimentale de l'état  $S_I(D)$  pour lequel les données actuelles se rapportent au mutant N56A.<sup>114</sup> Des simulation devraient être refaites à partir de cette structure expérimentale.

Des calculs de différences d'énergies libres associées à la mutation V61A dans la calbindine  $D_{9k}$  devraient être poursuivis afin d'observer une convergence raisonnable des structures A( $S_{II}$ ) et  $S_I(D)$ . Une simulation de quelques centaines de ps du mutant P43G-V61A pourrait être effectuée afin d'analyser un ensemble de configurations raisonnable et de faire des comparaisons avec la calbindine  $D_{9k}$  originale (P43G). Des calculs de perturbations sur d'autres mutants pourraient aussi être envisagés puisqu'il existe un liste considérable de ces mutants pour lesquels le changement de coopérativité a été mesuré expérimentalement en fonction de la concentration saline du milieu dans lequel se trouve la calbindine  $D_{9k}$ . Des calculs basés sur l'équation de Poisson-Boltzman qui étudieraient l'effet de la salinité sur le changement d'énergie libre associé à la coopérativité seraient aussi d'un intérêt certain.

## RÉFÉRENCES

- [1] Rasmussen, H. The Cycling of Calcium as an Intracellular Messenger. *Scientific American* **261(4)**, 66–73 (1989).
- [2] Heizmann, C.W. et Hunziker, W. Intracellular calcium-binding proteins: more sites than insights. *Trends in Biochemical Sciences* **16**, 98–103 (1991).
- [3] Pelmont, J. *Enzymes*. Presses Universitaires de Grenoble, (1993).
- [4] Kretsinger, R.H. et Nockolds, C.E. Carp Muscle Calcium-binding Protein. Structure Determination and General Description. *Journal of Biological Chemistry* **248(9)**, 3313–3326 (1973).
- [5] Fullmer, C.H. et Wasserman, R.H. Bovine Intestinal Calcium-binding Proteins. Purification and some properties. *Biochimica and Biophysica Acta* **317**, 172–186 (1973).
- [6] Balmain, N. Calbindin D<sub>9k</sub>, A Vitamin D-dependent, Calcium-Binding Protein in Mineralized Tissues. *Clinical Orthopaedics and Related Research* **265**, 265–276 (1991).
- [7] L'Horset F., Perret C., Brehier A. et Thomasset M. 17 beta-estradiol stimulates the calbindin-D9k (CaBP9k) gene expression at the transcriptional and posttranscriptional levels in the rat uterus. *Endocrinology* **127(6)**, 2891–2897 (1990).
- [8] Darwish, H., Krisinger, J., Furlow, J.D., Smith, C., Murdoch, F.E. et DeLuca, H.F. An estrogen-responsive element mediates the transcriptional regulation of calbindin D-9K gene in rat uterus. *Journal of Biological Chemistry* **266(1)**, 551–558 (1991).

- [9] Zanello, S.B., Boland, R.L. et Norman, A.W. cDNA sequence identity of a vitamin D-dependent calcium-binding protein in the chick to calbindin D-9K. *Endocrinology* **136(6)**, 2784–2787 (1995).
- [10] Schroder, B., Schlumbohm, C., Kaune, R. et Breves, G. Role of Calbindine D<sub>9k</sub> in Buffering Cytosolic Free Ca<sup>2+</sup> Ions in Pig Duodenal Enterocytes. *Journal of Physiology* **492(3)**, 715–722 (1996).
- [11] Christakos, N., Gabrielides, C. et Rhoten, W. Vitamin D-Dependent Calcium Binding Proteins: Chemistry, Distribution, Functional Considerations, and Molecular Biology. *Endocrine Reviews* **10**, 3–26 (1989).
- [12] Fullmer, C.H. et Wasserman, R.H. The Amino Acid Sequence of Bovine Intestinal Calcium Binding Protein. *Journal of Biological Chemistry* **256**, 5669–5674 (1981).
- [13] Szebenyi, D.M., Obendorf, S.K. et Moffat, K. Structure of vitamin D-dependent calcium-binding protein from bovine intestine. *Nature* **294**, 327–332 (1981).
- [14] Kördel, J., Forsén, S., Drakenberg, T. et Chazin, W. J. The Rate et Structural Consequences of Proline Cis-Trans Isomeration in Calbindin D<sub>9k</sub>: NMR Studies of the Minor (*cis*-Pro43) Isoform and the Pro43Gly Mutant. *Biochemistry* **29**, 4400–4409 (1990).
- [15] Svensson, L.A., Thulin, E. et Forsén, S. Proline *Cis-Trans* Isomers in Calbindin D<sub>9k</sub> Observed by X-Ray Crystallography. *Journal of Molecular Biology* **223**, 601–606 (1992).
- [16] Linse, S., Brodin, P., Drakenberg, T., Thulin, E., Sellers, P., Elmdén, K., Grundström, T. et Forsén, S. Structure-Function Relationships in EF-hand Ca<sup>2+</sup>-binding Proteins. Proteins Engineering and Biophysical Studies of Calbindin D<sub>9k</sub>. *Biochemistry* **26**, 6723–6735 (1987).

- [17] Linse, S., Johansson, C., Brodin, P., Grundström, T., Drakenberg, T. et Forsén, S. Electrostatic Contribution to the Binding of Calcium in Calbindin D<sub>9k</sub>. *Biochemistry* **31**, 154–162 (1991).
- [18] Monod, J., Changeux, J.-P. et Jacob, F. Allosteric Proteins and Cellular Control Systems. *Journal of Molecular Biology* **6**, 306–329 (1963).
- [19] Monod, J., Wyman, J. et Changeux, J.-P. On the Nature of Allosteric Transitions: A Plausible Model. *Journal of Molecular Biology* **12**, 88–118 (1965).
- [20] Koshland, D.E., Némethy, G. et Filmer, D. Comparison of Experimental Binding Data et Theoretical Models in Proteins Containing Subunits. *Biochemistry* **5**, 365–385 (1966).
- [21] Weber, G. Ligand Binding et Internal Equilibria in Proteins. *Biochemistry* **11**, 864–878 (1972).
- [22] Weber, G. Energetics of Ligand Binding to Proteins. *Advances in Protein Chemistry* **29**, 2–84 (1975).
- [23] Linse, S. Biophysical Studies of the Cooperativity of Calcium Binding in Calmodulin et Calbindin D<sub>9k</sub>. *Thèse de doctorat, Université de Lund, Suède*. (1993).
- [24] Akke, M. et Skelton, N.J. et Kördel, J. et Palmer, A. G. et Chazin, W. J. Effects of Ion Binding on the Backbone Dynamics of Calbindin D<sub>9k</sub> Determined by <sup>15</sup>N NMR Relaxation. *Biochemistry* **32**, 9832–9844 (1993).
- [25] Akke, M., Forsén, S. et Chazin, W.J. Molecular Basis for Cooperativity in Ca<sup>2+</sup> Binding to Calbindin D<sub>9k</sub>. <sup>1</sup>H Nuclear Magnetic Resonance Studies of (Cd<sup>2+</sup>)<sub>1</sub>-Bovine Calbindin D<sub>9k</sub>. *Journal of Molecular Biology* **220**, 173–189 (1991).

- [26] Vogel, H.J., Drakenberg, T., Forsén, S., O'Neil, J.O.J. et Hofmann, T. Structural Differences in the Two Calcium Binding Sites of the Porcine Intestinal Calcium Binding Protein: A Multinuclear NMR Study. *Biochemistry* **24**, 3870–3876 (1985).
- [27] Skelton, N.J. et Kördel, J. et Forsén, S. et Chazin, W.J. Comparative Structural Analysis of the Calcium Free and Bound States of the Calcium Regulatory Protein Calbindin D<sub>9k</sub>. *Journal of Molecular Biology* **213**, 593–598 (1990).
- [28] Skelton, N.J., Kördel, J., Akke, M. et Chazin, W.J. Nuclear Magnetic Resonance Studies of the Internal Dynamics in Apo, (Cd<sup>2+</sup>)<sub>1</sub> and (Ca<sup>2+</sup>)<sub>2</sub> Calbindin D<sub>9k</sub>. The Rates of Amide Proton Exchange with Solvent. *Journal of Molecular Biology* **227**, 1100–1117 (1992).
- [29] Akke, M., Forsén, S. et Chazin, W.J. <sup>15</sup>N NMR Assignments of (Cd<sup>2+</sup>)<sub>2</sub>-Calbindin D<sub>9k</sub> and Comparison with (Ca<sup>2+</sup>)<sub>2</sub>-Calbindin D<sub>9k</sub>. Cadmium as a Substitute for Calcium in Calcium-Binding Proteins. *Magnetic Resonance in Chemistry* **31**, S128–S132 (1993).
- [30] Linse, S., Thulin, E. et Sellers, P. Disulfide bonds in homo- and heterodimers of EF-hand subdomains of calbindin D<sub>9k</sub>: stability, calcium-binding and NMR studies. *Protein Science* **2(6)**, 985–1000 (1993).
- [31] Linse, S., Bylsma, N.R., Drakenberg, T., Sellers, P., Forsén, S., Thulin, E., Svensson, L.A., Zajtzeva, I., Vjacheslav, Z. et Marek, J. A Calbindin D<sub>9k</sub> Mutant with Reduced Calcium Affinity et Enhanced Cooperativity. Metal Ion Binding, Stability, and Structural Studies. *Biochemistry* **33**, 12478–12486 (1994).
- [32] Skelton, N.J., Kördel, J., Akke, M. et Chazin, W.J. Determination of the Solution Structure of Apo Calbindin D<sub>9k</sub> by NMR Spectroscopy. *Journal of Molecular Biology* **249**, 441–462 (1995).

- [33] Akke, M. et Chazin, W.J. Solution Structure of  $(\text{Cd}^{2+})_1$ -Calbindin  $\text{D}_{9k}$  Reveals Details of the Stepwise Structural Changes along the Apo $\rightarrow(\text{Ca}^{2+})_1\rightarrow(\text{Ca}^{2+})_2$  Binding Pathway. *Journal of Molecular Biology* **252**, 102–121 (1995).
- [34] Linse, S., Jönsson, B. et Chazin, W.J. The Effect of Protein Concentration on Ion Binding. *Proceedings of the National Academy of Sciences* **92**, 4748–4752 (1995).
- [35] Linse, S. et Chazin, W.J. Quantitative measurements of the cooperativity in an EF-hand protein with sequential calcium binding. *Protein Science* **4(6)**, 1038–1044 (1995).
- [36] Wimberley, B., Thulin, E. et Chazin, W.J. Characterization of the N-terminal half-saturated state of calbindin  $\text{D}_{9k}$ : NMR Studies of the N56A mutant. *Protein Science* **4(6)**, 1045–1055 (1995).
- [37] Denisov, P.V. et Halle, B. Direct Observation of Calcium-Coordinated Water in Calbindin  $\text{D}_{9k}$  by Nuclear Magnetic Relaxation Dispersion. *Journal of the American Chemical Society* **117**, 8456–8465 (1995).
- [38] Strynadka, N.C. et James, M.N.G. Crystal Structures of the Helix-Loop-Helix Calcium-Binding Proteins. *Annual Reviews of Biochemistry* **58**, 951–998 (1989).
- [39] Linse, S., Teleman, O. et Drakenberg, T.  $\text{Ca}^{2+}$ -binding to Calbindin  $\text{D}_{9k}$  Strongly Affects the Backbone Dynamics: Measurements of Exchange Rates of Individual Amide Protons Using  $^1\text{H}$  NMR. *Biochemistry* **29**, 5925–5934 (1990).
- [40] Cooper, A. et Dryden, D.T.F. Allostery without conformational change. A plausible model. *European Biophysics Journal* **11**, 103–109 (1984).

- [41] Colombo, M.F., Rau, D.C. et Parsegian, V.A. Protein Solvation in Allosteric Regulation: A Water Effect on Hemoglobin. *Science* **256**, 655–659 (1992).
- [42] Colombo, M.F. et Bonilla-Rodriguez, G.O. The Water Effect on Allosteric Regulation of Hemoglobin Probed in Water/Glucose and Water/Glycine Solutions. *Journal of Biological Chemistry* **271(9)**, 4895–4899 (1996).
- [43] Linse, S., Johansson, C., Brodin, P., Grundström, T., Thulin, E. et Forsén, S. The Role of Protein Surface Charges in Ion Binding. *Nature* **335**, 651–652 (1988).
- [44] Kesvatera, T., Jönsson, B., Thulin, E. et Linse, S. Binding of  $\text{Ca}^{2+}$  to Calbindin  $\text{D}_{9\text{k}}$ : Structural Stability and Function at High Salt Concentration. *Biochemistry* **33**, 14170–14176 (1994).
- [45] Falke, J.J., Drake, S., Hazard, A.L. et Peersen, O. Molecular Tuning of Ion Binding to Calcium Signaling Proteins. *Quarterly Reviews of Biophysics* **27(3)**, 219–290 (1994).
- [46] Akke, M. Molecular Basis for Cooperativity in Calcium Binding by Calbindin  $\text{D}_{9\text{k}}$  as Studied by NMR. *Thèse de doctorat* (1993).
- [47] Szabo, A. et Ostlund, N.S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. McGraw-Hill, New York, (1989).
- [48] Cohen-Tannoudji, C., Diu, B. et Laloë, F. *Mécanique quantique*. Hermann, Paris, (1977).
- [49] Bunker, P.R. *Molecular Symmetry et Spectroscopy*. Academic Press, San Diego (Californie), (1979).
- [50] Frisch, M.J., Head-Gordon, M., Trucks G.W., Foresman, J.B., Schlegel, H.B., Raghavachari, K., Robb, M.A., Binkley, J.S., Gonzalez, C., Defrees, D.J., Fox, D.J., Whitesides, R.A., Seeger, R., Melius, C.F., Baker, J.,

- Martin, R.L., Khan, L.R., Steward, J.J.P., Topiol, S. et Pople, J.A. Gaussian 90. *Gaussian Inc.* (1990).
- [51] Mackerell, A. D. Jr., Bashford, D., Bellot, M., Dunbrack, R. L., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Nguyen, D. T., Ngo, T., Prod'hom, B., Roux, B., Schlenkrich, B., Smith, J., Stote, R., Straub, J., Wiorkiewicz-Kuczera, J. et Karplus, M. Self-Consistent Parametrization of Biomolecules for Molecular Modelling and Condensed Phase Simulations. *Biophysical Journal* **61**, A143 (1992).
- [52] Karplus, M. et Porter, R.N. *Atoms and Molecules*. Benjamin, Melon Park, (1970).
- [53] McCammon, J.A., Gelin, B.R. et Karplus, M. Dynamics of folded proteins. *Nature* **267**, 585–590 (1977).
- [54] Karplus, M. et McCammon, J.A. Dynamics of Proteins: Elements and Function. *Annual Review of Biochemistry* **53**, 263–300 (1983).
- [55] Brooks, C.L., Karplus, M. et Pettitt, B.M. Proteins. A theoretical perspective of dynamics, structure and thermodynamics. In I. Prigogine et S. A. Rice, editor, *Advances in Chemical Physics Vol. LXXI*. John Wiley & Sons, New York, (1988).
- [56] Brooks, B.R., Bruccoleri, R.B., Olafson, B.D., States, D.J., Swaminathan, S. et Karplus, M. CHARMM: A program for macromolecular energy minimization and dynamics calculations. *Journal of Computational Chemistry* **4**, 187–217 (1983).
- [57] Israelachvilli, J. *Intermolecular et Surface Forces*. Academic press, San Diego (Californie), (1992).
- [58] Jackson, J. D. *Classical Electrodynamics*. John Wiley & Sons, New York, (1962).

- [59] Lorrain, P. et Corson, D.R. *Champs et ondes électromagnétiques*. W.H. Freeman and Company, É.-U., (1979).
- [60] Jorgensen, W.L., Chandrasekhar, J. and Madura, J.D., Impey, R.W. et Klein, M.L. Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics* **79**, 926–935 (1983).
- [61] Beglov, D. et Roux, B. Solvent Boundary Potential for Computer Simulations. *Journal of Chemical Physics* **100**, 9050–9063 (1994).
- [62] McQuarrie, D. A. *Statistical Mechanics*. Harper and Row, New York, (1976).
- [63] Reif, R. *Fundamentals of Statistical and Thermal Physics*. McGraw-Hill, New York, (1965).
- [64] Kirkwood, J.G. Statistical Mechanics of Fluid Mixtures. *Journal of Chemical Physics* **3**, 300–313 (1935).
- [65] Ryckaert, J.P., Ciccotti, G. et Berendsen, H.J.C. Numerical Integration of the Cartesian Equation of Motions of a System with Constraints: Molecular Dynamics of *n*-Alkanes. *Journal of Computational Chemistry* **23**, 327–341 (1977).
- [66] Van Gunsteren, W.F. et Berendsen, H.J.C. Algorithms for Macromolecular Dynamics and Constraint Dynamics. *Molecular physics* **34**, 1311 (1977).
- [67] Van Gunsteren, W.F. et Weiner, P.K. *Computer Simulation of Biomolecular Systems: Theoretical and Experimental Applications*. Escom, Leiden, (1989).
- [68] R. W. Zwanzig. High temperature equation of state by a perturbation method. *J. Chem. Phys.* **22**, 1420–1426 (1954).
- [69] Åqvist, J. Ion-Water Interaction Potentials Derived from Free Energy Perturbation Simulations. *Journal of Physical Chemistry* **94**, 8021–8024 (1990).

- [70] Franks, F., et Reid, D.S. *In Water, A Comprehensive Treatise; Vol. II.* Franks, F., Ed., Plenum, New-York, (1973).
- [71] Ahlström, P., Teleman, O., Jönsson, B. et Forsén, S. Molecular Dynamics Simulation of Parvalbumin in Aqueous Solution. *Journal of American Chemical Society* **109**, 1541–1551 (1987).
- [72] Sussman, F. et Weinstein, H. On The Ion Selectivity in Ca-Binding Proteins: the Cyclo(-L-Pro-Gly-)3 Peptide as a Model. *Proceedings of the National Academy of Sciences* **86**, 7880 (1989).
- [73] Prod'hom, B. et Karplus, M. The Nature of the Ion Binding Interactions in EF-Hand Peptide Analogs: Free Energy Simulation of Asp to Asn Mutations. *Protein Engineering* **6(6)**, 585–592 (1993).
- [74] Obst, S. et Bradaczek, H. Molecular Dynamics Study of the Structure et Dynamics of the Hydration Shell of Alkaline and Alkaline-Earth Metal Cations. *Journal of Physical Chemistry* **100**, 15677–15687 (1996).
- [75] Neilson, G.W. et Enderby, J.E. *In Water, A Comprehensive Treatise; Vol.XI.* Franks, F., Ed., Plenum, New-York, (1979).
- [76] Einspahr, H. et Bugg, C.E. The Geometry of Calcium-Water Interactions in Crystalline Hydrates. *Acta Crystallographica* **B36**, 264–271 (1980).
- [77] Hewish, N.A., Neilson, G.W. et Enderby, J.E. Environment of  $\text{Ca}^{2+}$  Ions in Aqueous Solvent. *Nature* **297**, 138–139 (1982).
- [78] Neilson, G.W. et Enderby, J.E. The Coordination of Metal Aquaions. *Advances in Inorganic Chemistry* **34**, 195–218 (1989).
- [79] Shiratori, Y. et Nakagawa, S. Parametrization of Calcium Binding Sites in Proteins and Molecular Dynamics Simulation on Phospholipase  $\text{A}_2$ . *Journal of Computational Chemistry* **12(6)**, 717–730 (1991).

- [80] Kitano, M. et Kuchitsu, K. Molecular Structure of Acetamide as Studied by Gas Electron Diffraction. *Bulletin of the Chemical Society of Japan* **46**, 3048–3051 (1973).
- [81] Kitano, M., Fukuyama, T. et Kuchitsu, K. Molecular Structure of N-Methyl-Acetamide as Studied by Gas Electron Diffraction. *Bulletin of the Chemical Society of Japan* **46**, 384–387 (1973).
- [82] Binkley, J.S., Pople, J.A. et Hehre, W.J. Self-Consistent Molecular Orbital Methods. Small Split-Valence Basis Sets For First-Row Elements. *Journal of Chemical Physics* **102**, 939–947 (1980).
- [83] Hariharan, P.C. et Pople, J.A. The Influence of Polarization Functions on Molecular Orbital Hydrogenation Energies. *Theoretica Chimica Acta* **28**, 213–222 (1973).
- [84] Hehre, W.J., Ditchfield, R. et Pople J.A. Self-Consistent Molecular Orbital Methods. XII. Further Extension of Gaussian-Type Basis Sets For Use in Molecular Orbital Studies of Organic Molecules. *Journal of Chemical Physics* **56**, 2257 (1972).
- [85] Masamura, M. The structures of  $\text{HCOO}^-$ ,  $\text{CH}_3\text{COO}^-$ ,  $\text{C}_2\text{H}_5\text{COO}^-$  et  $\text{CH}_3\text{O}^-$  in Gas Phase and in Crystal Structure by *ab initio* and Resonance Theory. *Theoretica Chimica Acta* **75**, 433–446 (1989).
- [86] Nagy, P.I., Smith, D.A., Alagona, G. et Ghio, C. *Ab Initio* Studies of Free and Monohydrated Carboxylic Acids in the Gas Phase. *Journal of Physical Chemistry* **98**, 486–493 (1994).
- [87] Wachters, A.J.H. Gaussian Basis Set for Molecular Wave Functions Containing Third Row Atoms. *Journal of Chemical Physics* **52**, 1033–1036 (1970).

- [88] Gianolo, L., Pavani, R. et Clementi, E. A New Algorithm For Obtaining Contracted Basis Sets From Gaussian Type Function. *Gazzetta Chimica Italiana* **108**, 181–205 (1978).
- [89] Boys, S.F. et Bernardi, F. The Calculation of Small Molecular Interactions by the Differences of Separate Total Energies. Some Procedures with Reduced Errors. *Molecular Physics* **19**, 553–566 (1970).
- [90] Kochanski, E. et Prissette, J. Effect of Ionic Charge on Energy Contributions to the Interaction on an Ion With a Polar Molecule: *Ab Initio* Calculations For M...H<sub>2</sub>O (M= Mg, Mg<sup>+</sup>, Mg<sup>2+</sup>, Ca, Ca<sup>+</sup>, Ca<sup>2+</sup>). *Chemical Physics Letters* **80(3)**, 564 (1981).
- [91] Kaufman Katz, A., Glusker, J.P., Beebe, S.A. et Bock, C.W. Calcium Ion Coordination: A Comparison with That of Beryllium, Magnesium and Zinc. *Journal of the American Chemical Society* **118**, 5752–5763 (1996).
- [92] Garmer, D.R. et Gresh, N. A Comprehensive Energy Component Analysis of the Interaction of Hard and Soft Dications with Biological Ligands. *Journal of the American Chemical Society* **116**, 3556–3567 (1994).
- [93] Einspahr, H. et Bugg, C.E. The geometry of Calcium-Carboxylate Interactions in Crystalline Complexes. *Acta Crystallographica* **B37**, 1044–1052 (1981).
- [94] Clementi, E., Kistenmacher, H., Kolos, W. et Romano, S. Non-Additivity in Water-Ion-Water Interactions. *Theoretica Chimica Acta* **55**, 257–266 (1980).
- [95] Roux, B. Non-Additivity in Cation-Peptide Interactions. A Molecular Dynamics and *Ab Initio* Study of Na<sup>+</sup> in the Gramicidin Channel. *Chemical Physics Letters* **212(3,4)**, 231–239 (1993).
- [96] Goodfellow, J.M. Cooperative Effects in Water-Biomolecule Crystal Systems. *Proceedings of the National Academy of Sciences* **79**, 4977 (1982).

- [97] Belle, D.V., Couplet, M., Prévost, M. et Wodak, S.J. Calculations of Electrostatic Properties in Proteins. Analysis of Contributions From Induced Protein Dipoles. *Journal of Molecular Biology* **198**, 721 (1987).
- [98] Perera, L. et Berkowitz, M.L. Many-Body Effects in Molecular Dynamics Simulations of  $\text{Na}^+(\text{H}_2\text{O})_n$  and  $\text{Cl}^-(\text{H}_2\text{O})_n$  clusters. *Journal of Chemical Physics* **95(3)**, 1954–1962 (1991).
- [99] Ortega-Blake, I., Hernández, J. and Novaro, O. The Role of Nonadditive Effects in the Second Hydration Shell of  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$ . A Molecular Orbital Study of the Three-Body Potential Energy Surface. *Journal of Chemical Physics* **81(4)**, 1894–1900 (1984).
- [100] Probst, M.M. A Study of the Additivity of Interactions in Cation-Water Systems. *Chemical Physics Letters* **137(3)**, 229–233 (1987).
- [101] Applequist, J., Carl, J.R. et Fung, K.K. An Atom Dipole Interaction Model for Molecular Polarizability. Application to Polyatomic Molecules and Determination of Atom Polarizabilities. *Journal of the American Chemical Society* **94(9)**, 2952–2960 (1972).
- [102] Thole, B.T. Molecular Polarizabilities Calculated with a Modified Dipole Interaction. *Chemical Physics* **59**, 341–350 (1981).
- [103] Miller, K.J. Additivity Methods in Molecular Polarizability. *Journal of the American Chemical Society* **112**, 8533–8542 (1990).
- [104] Miller, K.J. Calculation of the Molecular Polarizability Tensor. *Journal of the American Chemical Society* **112**, 8543–8551 (1990).
- [105] Wolfram, S. *Mathematica. A System for Doing Mathematics by Computer*. Addison-Wesley Publishing Company, Inc., (1991).
- [106] Vogel, A. . *Journal of the Chemical Society* page 1833 (1948).

- [107] Karplus, M. et McCammon, J.A. . *Nature* **277**, 578 (1979).
- [108] Ahlström, P., Teleman, O., Kördel, J., Forsén, S. et Jönsson, B. A Molecular Dynamics Simulation of Bovine Calbindin D<sub>9k</sub>. Molecular Structure and Dynamics. *Biochemistry* **28**, 3205–3211 (1989).
- [109] Kördel, J. et Teleman, O. Backbone Dynamics of Calbindin D<sub>9k</sub>: Comparison of Molecular Dynamics Simulations and <sup>15</sup>N NMR Relaxation Measurements. *Journal of the American Chemical Society* **114**, 4934–4936 (1992).
- [110] Mehler, E.L., Kushick, J.N. et Weinstein, H. Consequences of sequential Ca<sup>2+</sup> occupancy for the structure et dynamics of calbindin D<sub>9k</sub>: computational simulations and comparison to experimental determinations in solution. *Molecular Simulation* **10(2-6)**, 309–334 (1993).
- [111] Teleman, O. et Jönsson, B. Vectorizing a General Purpose Molecular Dynamics Simulation Program. *Journal of Computational Chemistry* **7(1)**, 58–66 (1986).
- [112] Szebenyi, D.M.E. et Moffat, K. The Refined Structure of Vitamin D-dependent Calcium-binding Protein from Bovine Intestine. *The Journal of Biological Chemistry* **261**, 8761–8777 (1986).
- [113] Strang, G. *Linear Algebra et Its Applications (Second Edition)*. Academic Press, Orlando, Florida, (1980).
- [114] Wimberly, B., Thulin, E. et Chazin, W.J. Characterization of the N-Terminal Half-Saturated State of Calbindin D<sub>9k</sub>: NMR Studies of the N56A Mutant. *Protein Science* **4**, 1045–1055 (1995).
- [115] Lee, B. et Richards, F.M. The Interpretation of Protein Structures: Estimation of Static Accessibility. *Journal of Molecular Biology* **55(3)**, 379–400 (1971).

- [116] Abragham, A. *Principles of Nuclear Magnetism*. Oxford University Press, Oxford, (1961).
- [117] Lipari, G. et Szabo, A. Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 1. Theory et Range of Validity. *Journal of the American Chemical Society* **104**, 4546–4559 (1982).
- [118] Brüshweiler, R. Normal Modes and NMR Order Parameters in Proteins. *Journal of the American Chemical Society* **114**, 5341–5344 (1992).
- [119] Chandrasekhar, I., Clore, G.M., Szabo, A., Gronenborn, A.M. et Brooks, B.R. A 500 ps Molecular Dynamics Simulation Study of Interleukin-1 $\beta$  in Water. *Journal of Molecular Biology* **226**, 239–250 (1992).
- [120] Smith, L.J., Mark, A.E., Dobson, C.M. et Van Gunsteren, W.F. Comparison of MD Simulations and NMR Experiments for Hen Lysozyme. Analysis of Local Fluctuations, Cooperative Motions, et Global Changes. *Biochemistry* **34**, 10918–10923 (1995).
- [121] Smith, P.E., Van Schaik, R.C., Szyperski, T., Wütrich, K. et Van Gunsteren, W.F. Internal Mobility of the Basic Pancreatic Trypsin Inhibitor in Solution: A Comparison of NMR Spin Relaxation Measurements and Molecular Dynamics Simulations. *Journal of Molecular Biology* **246**, 356–365 (1995).
- [122] Linse, S. . *Communication personnelle* (1995).
- [123] Dobson, C.M. et Karplus, M. Internal Motion of Proteins: Nuclear Magnetic Resonance Measurements and Dynamic Simulations. *Methods in Enzymology* **131**, 362–389 (1986).
- [124] Clore, G.M., Driscoll, P.C., Wingfield, P.T. et Gronenborn, A.M. Analysis of the Backbone Dynamics of Interleukin-1 $\beta$  Using Two-Dimensional Inverse

- Detected Heteronuclear  $^{15}\text{N}$ - $^1\text{H}$  NMR Spectroscopy. *Biochemistry* **29**, 7387–7401 (1990).
- [125] Clore, G.M., Szabo, A., Bax, A., Kay, L., Driscoll, P.C. et Gronenborn, A.M. Deviations from the Simple Two-Parameter Model-Free Approach to the Interpretation of Nitrogen-15 Nuclear Magnetic Relaxation of Proteins. *Journal of the American Chemical Society* **112**, 4989–4991 (1990).
- [126] Brüshweiler, R., Roux, B., Blackledge, M., Griesinger, C., Karplus, M. et Ernst, R.R. Influence of Rapid Intramolecular Motion on NMR Cross-Relaxation Rates. A Molecular Dynamics Study of Antamanide in Solution. *Journal of the American Chemical Society* **114**, 2289–2302 (1992).
- [127] Palmer, A.G., Rance, M. et Wright, P.E. Intramolecular Motions of a Zinc Finger DNA-Binding Domain from Xfin Characterized by Proton-Detected Natural Abundance  $^{13}\text{C}$  Heteronuclear NMR Spectroscopy. *Journal of the American Chemical Society* **113**(12), 4371–4380 (1991).
- [128] Palmer, A.G. et Case, D.A. Molecular Dynamics Analysis of NMR Relaxation in and Zinc-Finger Peptide. *Journal of the American Chemical Society* **114**, 9059–9067 (1992).
- [129] Eriksson, M.A.L., Berglund, H., Härd, T. and Nilsson, L. A Comparison of  $^{15}\text{N}$ -NMR Relaxation Measurements With a Molecular Dynamics Simulation: Backbone Dynamics of the Glucocorticoid Receptor DNA-Binding Domain. *Proteins: Structure, Function et Genetics* **17**, 375–390 (1993).
- [130] Yamasaki, K., Minoru, S., Oobatake, M. et Kanaya, S. Characterization of the Internal Motions of Escherichia Coli Ribonuclease HI by a Combination of  $^{15}\text{N}$ -NMR Relaxation Analysis and Molecular Dynamics Simulation: Examination of Dynamic Models. *Biochemistry* **34**, 6587–6601 (1995).

- [131] Philippopoulos, M. et Lim, C. Molecular Dynamics Simulation of E. Coli Ribonuclease H<sub>1</sub> in Solution: Correlation with NMR and X-ray Data and Insights into Biological Function. *Journal of Molecular Biology* **254**, 771–792 (1995).
- [132] Fushman, D., Weisemann, R., Thüring, H. Ohlenschläger, O. et Rüterjans, H. Backbone Dynamics of Proteins Studied by Two-Dimensional Heteronuclear NMR Spectroscopy and Molecular Dynamics Simulations. *International Journal of Quantum Chemistry* **59**, 291–300 (1996).
- [133] Woessner, D.E. Spin Relaxation Processes in a Two-Proton System Undergoing Anisotropic Reorientation. *Journal of Chemical Physics* **36(1)**, 1–4 (1962).
- [134] Kinoshita, K., Kawato, S. et Ikegami, A. A Theory of Fluorescence Polarization Decay in Membrane. *Biophysical Journal* **20**, 289–305 (1977).
- [135] Lipari, G. et Szabo, A. Effect of Librational Motion on Fluorescence Depolarization and Nuclear Magnetic Resonance Relaxation in Macromolecules and Membranes. *Biophysical Journal* **30**, 489–506 (1980).
- [136] Brüshweiler, R. et Wright, P.E. NMR Order Parameters of Biomolecules: A New Analytical Representation and Application to the Gaussian Axial Fluctuation Model. *Journal of the American Chemical Society* **116**, 8426–8427 (1994).
- [137] Groves, P., Linse, S., Thulin, E. et Förse, S. A Calbindin D<sub>9k</sub> Mutant Containing a Novel Structural Extension: <sup>1</sup>H Nuclear Magnetic Resonance Studies. *Protein Science* **6**, 323–330 (1997).
- [138] Groves, P., Linse, S., Thulin, E. et Forsén, S. A calbindin D<sub>9k</sub> Mutant Containing a Novel Structural Extension: <sup>1</sup>H Nuclear Magnetic Resonance Studies. *Protein Science* **6**, 323–330 (1997).

- [139] Forsén, S., Linse, S., Thulin, E., Wendt, B., Martin, S.R., Bayley, P.M., Brodin, P. et Grundström, T. Kinetics of Calcium Binding to Calbindin Mutants. *European Journal of Biochemistry* **177**, 47–52 (1988).
- [140] Johansson, C., Brodin, P., Grundström, T., Thulin, E., Forsén, S. et Drakenberg, T. Biophysical Studies of Engineered Mutant Proteins Based on Calbindin D<sub>9k</sub> Modified in the Pseudo EF-Hand. *European Journal of Biochemistry* **187**, 455–460 (1990).
- [141] Brodin, P., Johansson, C., Forsén, S., Drakenberg, T. et Grundström, T. Functional Properties of Calbindin D<sub>9k</sub> Mutants with Exchanged Ca<sup>2+</sup> Binding Sites. *Journal of Biological Chemistry* **265**, 11125–11130 (1990).
- [142] Johansson, C., Brodin, P., Grundström, T., Forsén, S. et Drakenberg, T. Mutation of the pseudo EF-hand of calbindin D<sub>9k</sub> into a normal EF-hand. Biophysical studies. *European Journal of Biochemistry* **202**, 1283–1290 (1991).
- [143] Johansson, C., Ullner, M. et Drakenberg, T. The Solution Structures of Mutant Calbindin D<sub>9k</sub>'s, as Determined by NMR, Show that the Calcium Binding Site can Adopt Different Folds. *Biochemistry* **32(33)**, 8429–8438 (1993).
- [144] Kollman, P. Free Energy Calculations: Applications to Chemical and Biochemical Phenomena. *Chemical Reviews* **93**, 2395–2417 (1993).
- [145] Mark, A.E., Van Gunsteren, W.F. et Berendsen, H.J.C. Calculation of Relative Free Energy *via* Indirect Pathways. *Journal of Chemical Physics* **94(5)**, 3808–3816 (1991).
- [146] Roux, B. The Calculation of the Potential of Mean Force Using Computer Simulations. *Computer Physics Communications* **91**, 275–282 (1995).
- [147] Souaille, M. et Roux, B. . *Manuscrit en préparation* (1997).

- [148] Mitchell, M.J. et McCammon, J.A. Free Energy Difference Calculations by Thermodynamic Integration: Difficulties in Obtaining a Precise Value. *Journal of Computational Chemistry* **12(2)**, 271–275 (1991).
- [149] Wolfenden, R., Andersson, L., Cullis, P.M. and Southgate, C.C.B. Affinities of Amino Acid Side Chains for Solvent Water. *Biochemistry* **20**, 849–855 (1981).
- [150] Yamaotsu, N., Moriguchi, I. et Hirono, S. Solvation Free Energies of Amino Acids Calculated by Molecular Dynamics/Free Energy Perturbation Method. *Chemical Pharmaceutical Bulletin* **43(5)**, 717–721 (1995).
- [151] Pearlman, D.A., Case, D.A., Caldwell, J.C., Seibel, G.L., Singh, U.C., Weiner, P. et Kollman, P.A. *AMBER 4.0*. University of California, San Francisco, (1991).

## REMERCIEMENTS

Je remercie mon directeur de thèse, Benoît Roux, pour m'avoir permis d'étudier dans son groupe et attribué ce sujet de recherche, pour les ressources de qualité qu'il a mises à ma disposition et pour sa collaboration et son efficacité lors des dernières semaines critiques de ce travail. Je remercie également le Groupe de Recherche en Transport Membranaire et Benoît Roux pour la bourse (F.C.A.R.) qui m'a été attribuée tout au long de mes études doctorales.

J'aimerais aussi remercier les professeurs Jean-Robert Derome, Raynald Laprade et John Gunn pour leur soutien et collaboration à l'achèvement de cette thèse, de même que les docteurs Matthew Bramley et Marc Souaille.

Et finalement, il y a ceux dont la personnalité et l'amitié ont marqué les cinq années passées à ce travail: Matthew Bramley, Pierre-Nicholas Roy, John Tromp et Isabelle Gambu pour l'ambiance des premières années; ainsi que Laura Deakin, Marc et Catherine Souaille qui ont été d'une amitié précieuse lors des derniers mois de ce travail.