



HAL
open science

Étude du comportement de Calciprotéines à mains EF en présence de Ca^{2+} et Mg^{2+} à partir de trajectoires de dynamique moléculaire et de calculs de différences d'énergie libre

David Allouche

► **To cite this version:**

David Allouche. Étude du comportement de Calciprotéines à mains EF en présence de Ca^{2+} et Mg^{2+} à partir de trajectoires de dynamique moléculaire et de calculs de différences d'énergie libre. Sciences du Vivant [q-bio]. Université Paul Sabatier (Toulouse 3), 1997. Français. NNT : 1997TOU30301 . tel-04548873

HAL Id: tel-04548873

<https://hal.inrae.fr/tel-04548873>

Submitted on 16 Apr 2024

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

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

N° d'ordre : 2885

THÈSE

présentée au laboratoire de Physique Quantique
(UMR 5626 du CNRS). I. R. S. A. M. C.
118, route de Narbonne - 31 062 Toulouse Cedex

devant l'UNIVERSITÉ PAUL SABATIER - TOULOUSE III

en vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ PAUL SABATIER

Spécialité : Physico-chimie des biomolécules

par

David ALLOUCHE

Étude du comportement de calciprotéines à mains EF en présence de Ca^{2+} et Mg^{2+} à partir de trajectoires de dynamique moléculaire et de calculs de différences d'énergie libre.

Soutenue le 17 décembre 1997 devant la commission d'examen

Président

Jacques Périé

Professeur à l'Université Paul Sabatier

Rapporteurs

Alain Gerschel

Directeur de recherche au CNRS

David Perahia

Directeur de recherche au CNRS

Examineurs

Joseph Parello

Directeur de recherche au CNRS

Jean Durup

Professeur à l'Université Paul Sabatier

Yves-Henri Sanejouand

Chargé de recherche au CNRS

ABSTRACT

The modulation of Ca^{2+} concentration is involved in many biological processes, called Ca^{2+} -dependent. Then Ca^{2+} often interacts with EF hands calciproteins. In physiological conditions Ca^{2+} is not the only divalent cation, since a high amount of Mg^{2+} is also present. One of the central questions concerning the behaviour of these calciproteins is therefore the understanding, at a molecular level, of their differences in affinity for Ca^{2+} and Mg^{2+} . Thus we have studied the difference of selectivity of several calciproteins for Ca^{2+} and Mg^{2+} . In that purpose, we have transformed Ca^{2+} into Mg^{2+} , in order to calculate differences in free energy by the thermodynamic perturbation method. We have first followed the structural evolution of the surrounding of the Ca^{2+} , during its transformation into Mg^{2+} . This was done in water and in the sites of the studied proteins. The transformations $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$ in the sites of parvalbumin (Parv) have shown a general contraction of the coordination sphere of the cation, which corresponds to a decrease in the average distance cation-oxygen. This contraction is concomitant with a decrease in the number of oxygens coordinated to the cation, from 8 for Ca^{2+} to 6 for Mg^{2+} , following the ultimate rearrangement of a bidentate ligand (Glu in relative position 12), that becomes monodentate. This rearrangement has been observed experimentally. Following the results obtained on Parv, we have studied the calciproteins troponinC and calmodulin, for which no three-dimensionnal structure of the protein complexed with Mg^{2+} is available. We have shown that these proteins follow the same pattern of structural rearrangement as that observed previously. Nevertheless our thermodynamical results have proven less satisfactory : the calculated $k_{\text{Ca}}/k_{\text{Mg}}$ ratio was not always in agreement with experience. Therefore we have developed from calculations ab initio, new sets of parameters to describe these cations.

Molecular dynamics
free energy calculation

molecular modelling
calcium

cations in solution
Magnesium

calciprotein
EF Hand

N° d'ordre : 2885

THÈSE

présentée au laboratoire de Physique Quantique
(UMR 5626 du CNRS). I. R. S. A. M. C.
118, route de Narbonne - 31 062 Toulouse Cedex

devant l'UNIVERSITÉ PAUL SABATIER - TOULOUSE III

en vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ PAUL SABATIER

Spécialité : Physico-chimie des biomolécules

par

David ALLOUCHE

Étude du comportement de calciprotéines à mains EF en présence de Ca^{2+} et Mg^{2+} à partir de trajectoires de dynamique moléculaire et de calculs de différences d'énergie libre.

Soutenue le 17 décembre 1997 devant la commission d'examen

Président

Jacques Périé

Professeur à l'Université Paul Sabatier

Rapporteurs

Alain Gerschel

Directeur de recherche au CNRS

David Perahia

Directeur de recherche au CNRS

Examineurs

Joseph Parello

Directeur de recherche au CNRS

Jean Durup

Professeur à l'Université Paul Sabatier

Yves-Henri Sanejouand

Chargé de recherche au CNRS

À mes Parents

*À Florence et Elora qui partagent mes joies et mes
peines au quotidien !*

J'exprime ici ma reconnaissance à Yves-Henri Sanejouand, qui a suivi mes travaux avec disponibilité et compétence. Qu'il soit également remercié pour l'aide précieuse qu'il m'a apporté pour la rédaction de cette thèse, de même que Romuald Poteau et Fabienne Alary.

Je tiens également à remercier Jean Durup pour ses conseils avisés ainsi que Joseph Parello, « source d'inspiration intarissable », qui est à l'origine d'une bonne part de ce travail.

David Perahia a bien voulu accepter de juger ce travail et participer à mon jury en bravant les aléas atmosphériques. Je le prie de trouver ici l'expression de mes remerciements sincères au même titre que Alain Gershel bien qu'il ait eu moins de chance dans sa lutte contre les intempéries¹!

Je remercie Jacques Périé d'avoir bien voulu présider le jury de thèse de l'élève ~~bruyant~~ brillant qu'il a contribué à former durant les cours du DEA de chimie des biomolécules.

Mes remerciements vont également à ma coturne Ginette Nicolas, pour son accueil chaleureux dans ce qui fut son territoire. Qu'elle soit sûre que je lui permettrai encore longtemps de rester dans mon bureau !

Un petit mot à l'illustre inconnu Xavier Periole, bientôt reconnu pour ces fameux jeux de paramètres permettant de décrire l'interaction cation-oxygène ! Je tiens à le remercier pour son soutien quotidien ainsi que ma petite Nathalie Guihery.

Il aurait été injuste de clore ces remerciements sans une petite pensée pour mon tuteur pédagogique Simon Mathieu², qui par sa présence de tous les instants m'a appris à me sortir des situations les plus périlleuses avec calme et pondération!

Pour finir j'aimerais remercier Jean-Pierre Daudey, qui m'a chaleureusement accueilli dans son laboratoire.

¹ C'est les intempéries qui ont gagné

² Mathieu@nagoya.jp

TABLE DES MATIERES

| | |
|---------------------|---|
| <i>INTRODUCTION</i> | 5 |
|---------------------|---|

LA SUPERFAMILLE DES CALCIPROTEINES A MAIN EF

| | |
|---|----|
| Le motif main EF | 19 |
| La Parvalbumine : un modèle pour d'autres protéines à « main EF » | 25 |
| Autres calciprotéines étudiées: | 41 |
| Références bibliographiques | 51 |

OUTILS ET METHODES

| | |
|---------------------|----|
| <i>Introduction</i> | 57 |
|---------------------|----|

| | |
|---|----|
| <i>De la notion de surface d'énergie potentielle à la dynamique moléculaire</i> | 58 |
|---|----|

| | |
|---|----|
| Notion de surface d'énergie potentielle | 58 |
| Champ de Forces | 61 |
| Dynamique Moléculaire : | 75 |
| Outils d'analyses | 88 |

| | |
|---|----|
| <i>De la trajectoire de dynamique moléculaire aux calculs d'énergie libre</i> | 92 |
|---|----|

| | |
|--|-----|
| Introduction | 92 |
| De la fonction de partition à l'expression de l'énergie libre | 95 |
| Calcul d'énergie libre par simulation de dynamique « classique » | 104 |
| Références Bibliographiques | 119 |

RESULTATS

| | |
|---|-----|
| <i>Trajectoires de dynamique moléculaire de cations en solution</i> | 128 |
|---|-----|

| | |
|---|-----|
| Protocole de calcul | 128 |
| Analyse des trajectoires de cations en solution | 134 |
| conclusion | 160 |

Calculs d'énergie libre : détermination de la différence d'énergie libre de solvation Ca²⁺ Mg²⁺ _____ 161

Protocole de calcul _____ 161

Simulation à cutoff 14 Å _____ 165

Simulation à cutoff 7,5 Å _____ 181

Conclusion _____ 188

Références Bibliographiques : _____ 193

Cations dans les calciprotéines _____ 196

Protocole de calcul _____ 201

Etude de la parvalbumine _____ 204

Echange Ca²⁺/Mg²⁺ dans d'autres protéines à Main EF . _____ 223

Discussion _____ 228

Conclusion _____ 235

Tentative avec une nouvelle paramétrisation de l'interaction cation - oxygène _____ 237

Références Bibliographiques _____ 241

CONCLUSION _____ 243



Introduction

INTRODUCTION

La modulation de la concentration intracellulaire en calcium intervient dans un grand nombre de fonctions biologiques, que l'on qualifie souvent de processus « Ca^{2+} dépendants ». Ces processus très variés - de la contraction musculaire à la libération de neurotransmetteurs - ont tous pour caractéristique de faire intervenir une étape dans laquelle la libération massive de calcium est utilisée comme message (cf. Figure 0-1). La concentration en Ca^{2+} varie alors de $\approx 10^{-7}$ M au repos, à $\approx 10^{-4}$ M après cette étape signal. Après sa libération, Ca^{2+} va former des complexes avec des calci-protéines, qui eux-mêmes vont activer d'autres maillons de ce qui formera une cascade réactionnelle, et aboutira à une réponse biologique.

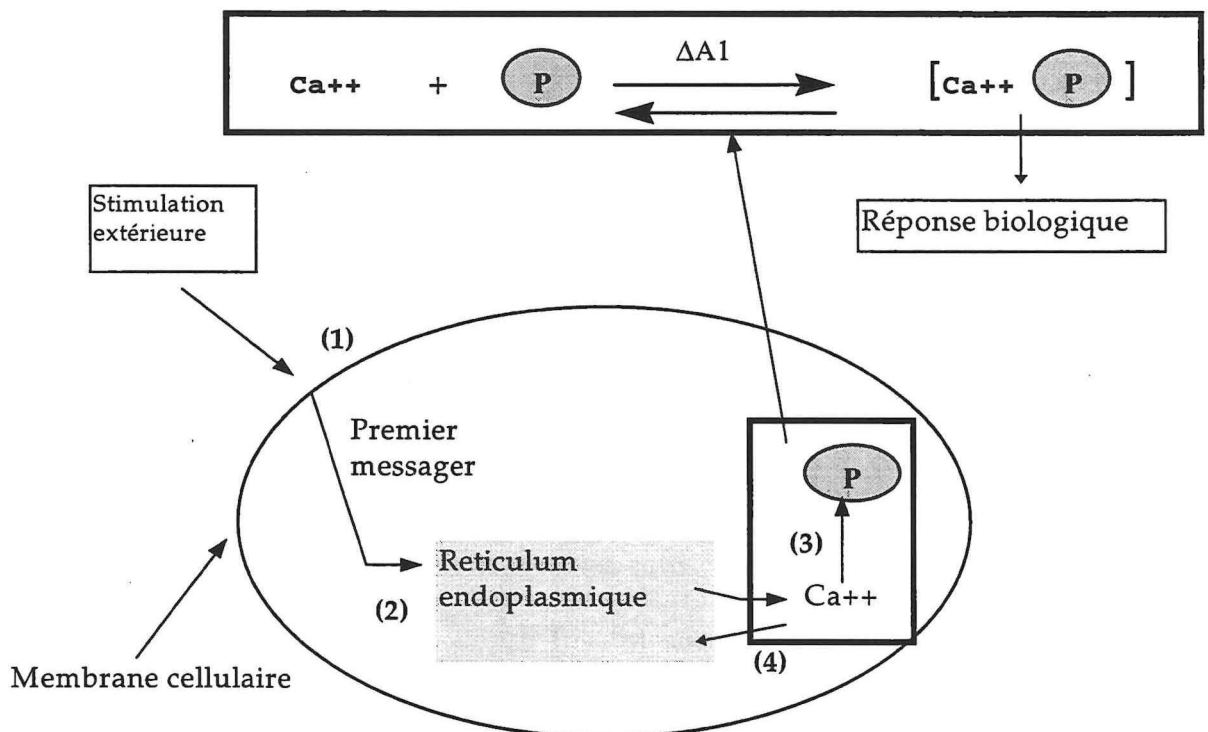


Figure 0-1 : représentation schématique d'un processus Ca^{2+} dépendant. 1) une stimulation extérieure induit l'émission dans le cytoplasme d'un premier messenger. 2) le premier messenger va induire la libération massive du Ca^{2+} stocké dans le reticulum endoplasmique. 3) le calcium va interagir avec des cibles protéiques et provoquer la réponse biologique après une cascade réactionnelle plus ou moins longue. 4) Ca^{2+} est extrait du cytoplasme par les pompes à Ca^{2+} et stocké à nouveau dans le reticulum endoplasmique.

On peut sur la base de leur séquence et de leur structure subdiviser les calci-protéines en plusieurs familles: les annexines, les calciprotéines à main EF, les protéines de stockage du calcium

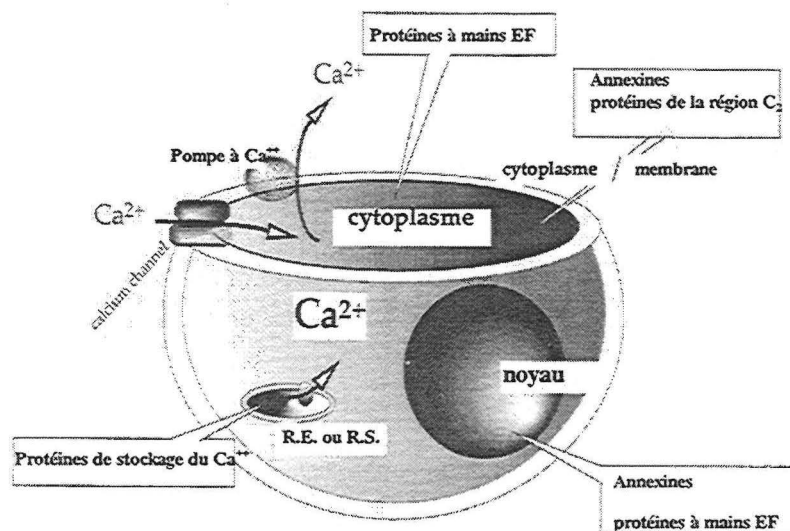
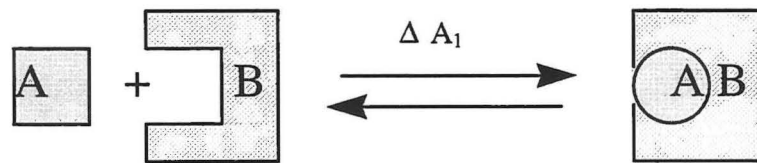


Figure 0-2 : Localisation des calciprotéines au sein d'une cellule (1).

Au cours de ce travail nous nous sommes intéressés plus particulièrement à la parvalbumine, à la troponine C et à la calmoduline qui appartiennent toutes trois à la famille des calciprotéines à main EF. Ces protéines ont entre autres la capacité de fixer de manière réversible le calcium, mais aussi le magnésium. En effet, dans les conditions physiologiques, Ca^{2+} n'est pas le seul cation divalent présent dans le cytoplasme, puisque Mg^{2+} est présent en relativement grande quantité ($\approx 10^{-3}$ M). Ces deux alcalino-terreux sont parmi les cations les plus représentés dans les organismes vivants.

De l'affinité à la reconnaissance

Les calciprotéines à main EF sont capables de discriminer le calcium et le magnésium à des degrés plus ou moins importants. Cette propriété est une des nombreuses illustrations de la capacité des macromolécules biologiques à effectuer ce que l'on peut appeler la « reconnaissance moléculaire ». Cette reconnaissance à l'échelle atomique est quantifiée au niveau macroscopique par la notion d'affinité, qui elle-même sur le plan thermodynamique est exprimée en terme de différence d'énergie libre. Ainsi, pour une réaction quelconque à l'équilibre, nous avons :



L'affinité de A pour B est donnée par la constante d'affinité k_a (mol^{-1}) à laquelle peut être associée une quantité d'énergie libre ΔA_1 obtenue selon la relation :

$$\Delta A_1 = -RT \ln k_a \quad \text{avec} \quad k_a = \frac{[AB]}{[A].[B]}$$

En fait, la notion de reconnaissance moléculaire ne prend un sens que lorsque les composés A et B sont immergés dans un milieu hétérogène contenant un grand nombre de molécules diverses autres que A et B. Ainsi, dans ce cas, la formation du complexe AB sera potentiellement en compétition avec la formation des complexes mettant en jeu les molécules A et B et les autres espèces moléculaires présentes dans le milieu.

Par conséquent, la formation du complexe AB sera fonction, d'une part, de l'affinité réciproque de A et B, mais aussi de l'affinité de ces molécules pour tous les autres composés présents dans le milieu réactionnel. Il y aura reconnaissance entre A et B si, sélectivement, le complexe AB a tendance à se former alors qu'il est en compétition avec la formation d'autres complexes, ce qui d'un point de vue thermodynamique se

traduit par le fait que l'énergie libre associée à la formation de ce complexe AB est plus basse que celle associée à la formation de tout autre complexe faisant intervenir A ou B (à concentration égale¹).

L'exemple de la formation d'un tel complexe AB est typique des phénomènes biochimiques, dans la mesure où le cytoplasme est un milieu hétérogène qui contient un grand nombre d'espèces moléculaires susceptibles d'interagir entre elles de manière non covalente et réversible. La fixation du calcium et du magnésium par les calci-protéines peut être considérée comme un exemple typique des phénomènes de reconnaissance tels que nous venons de les décrire, le degré de sélectivité de ces calci-protéines pouvant s'exprimer par le rapport des constantes d'affinité calcique et magnésique, $k_{Ca^{2+}}/k_{Mg^{2+}}$. Une des questions centrales concernant les calci-protéines est la compréhension à l'échelle moléculaire des différences d'affinité observées pour Ca^{2+} et Mg^{2+} .

De la structure à la fonction

Notre démarche se situe par ailleurs dans le cadre d'une approche plus générale qui se développe actuellement dans un grand nombre de domaines de la biochimie. Cette approche vise à faire le lien entre les événements à l'échelle atomique et les propriétés observées à des échelles plus grandes. Les premiers liens entre la forme et la fonction ont pu être établis lors de la détermination des premières structures tridimensionnelles de protéines globulaires, par diffraction de rayons X, dans les années 60 (2,3). Aujourd'hui, plusieurs milliers de structures ont été résolues par cristallographie. Du fait de leur caractère statique, ces dernières ont, dans un premier temps, conduit les biologistes à considérer les protéines comme des structures

¹ En modifiant les rapports de concentration il est bien sûr possible d'inverser la tendance.

moléculaires plutôt rigides. Mais, au cours des vingt dernières années, avec le développement de méthodes spectroscopiques toujours plus performantes (RMN, diffraction de neutrons ...), cette représentation figée a laissé de plus en plus place à une vision dynamique dans laquelle structure, dynamique et fonction des protéines sont étroitement liées. La relation structure-dynamique-fonction sera entre autre discutée dans le premier chapitre, au cours duquel nous ferons une synthèse des données expérimentales concernant les calci-protéines à mains EF que nous avons étudiées.

Parallèlement au développement des méthodes expérimentales, grâce notamment à des ordinateurs de plus en plus performants, les techniques de modélisation ont aussi considérablement progressé. Par exemple, la dynamique moléculaire a rendu possible l'étude de mouvements et de changements conformationnels se produisant sur des échelles de temps allant de la picoseconde à la nanoseconde. De plus, cette méthode a été appliquée au calcul de différences d'énergie libre en exploitant, entre autres, le fait que des chemins «non physiques » peuvent être suivis pas à pas lors d'une simulation.

C'est dans ce cadre que se situent les travaux présentés dans ce mémoire. Nous avons étudié la différence de comportement de diverses calci-protéines à mains EF en présence de calcium et de magnésium à partir de trajectoires de dynamique moléculaire et lors de calculs de différences d'énergie libre. Une telle étude est d'autant plus intéressante, sur un plan méthodologique, que les trajectoires de dynamique moléculaire décrivant des interactions de cations divalents avec leur environnement peuvent paraître à priori difficiles à réaliser, dans la mesure où les atomes sont modélisés comme des sphères dures chargées, et que l'on ne tient compte que de manière implicite de la polarisation induite par la présence d'atomes chargés. De manière plus générale, les

calculs de différence d'énergie libre donnent des résultats avec une précision difficile à estimer correctement et nous avons cherché, tout au long de ce travail, à améliorer cet aspect.

Comme l'essentiel de la différence d'énergie libre est dans ce cas d'origine enthalpique, il s'agit d'être capable de décrire correctement les interactions cation-eau, cation-protéine. A priori, la difficulté de notre entreprise ne consiste pas seulement à rendre quantitativement compte des différences d'affinité des calcioprotéines pour le calcium et le magnésium, mais aussi à pouvoir reproduire ce qui différencie les environnements des deux cations dans l'un et l'autre cas (leur nombre de coordination, notamment). Nous avons donc étudié en détails l'adaptation structurale des sites de calcioprotéines lors de la fixation de Ca^{2+} ou de Mg^{2+} .

Méthode de calcul

La méthode de calcul de différences d'énergie libre choisie est basée sur l'approche « alchimique ». Cet aspect sera développé en détail au cours du chapitre II consacré aux méthodes de simulation ainsi qu'aux outils d'analyse. Nous y décrivons successivement i) le champs de forces ii) la dynamique moléculaire proprement dite iii) quelques méthodes de calculs d'énergie libre, dont la méthode dite de « perturbation thermodynamique »(4) que nous avons utilisée, et iv) les observables que nous avons utilisées afin d'analyser nos résultats et de les comparer aux données expérimentales.

Comme nous le verrons plus tard, l'approche que nous avons utilisée fait intervenir un cycle thermodynamique(5,6), ce qui rend nécessaire le calcul de trajectoires des cations Ca^{2+} et Mg^{2+} en milieu aqueux d'une part, au sein des protéines d'autre part. C'est pourquoi nous avons subdivisé la présentation de nos résultats en deux chapitres . Le chapitre III regroupe les résultats relatifs aux simulations de cations

en solution et le chapitre IV ceux relatifs aux simulations au sein des sites de fixation des protéines.

Pour conclure, nous ferons la synthèse des résultats obtenus sur le plan structural et sur le plan énergétique et nous discuterons de ce qu'il faudrait approfondir par la suite afin d'améliorer le comportement des simulations mettant en oeuvre des interactions cation-protéines.



Références bibliographiques

1. Niki I., Hisayuki Y., Toshiki S., Masumi K., and Hiroyoshi H. (1996) *J.Biochem.* **120**, 685-698
2. Perutz M.F., Rossmann M.G., Cullis A.F., Muirhead H., Will G., and A.C.T., N. (1960) *Nature* **185**, 416
3. Kendrew J.C., Dickerson R.E., Strandberg B.E., Hart R.G., Davies D.R., Phillips D.C., and V.C., S. (1960) *Nature* **185**, 422
4. Zwanzig, R. (1954) *J. Chem. Phys.* **22**(1420), 1420
5. Van Gasteren, W. F., and Berensen, H. J. C. (1987) *Journal of Computer-Aided Molecular Design* **1**, 171-176
6. Kollman, P. A., and Merz, K. (1990) *Acc. Chem. Res.* **23**, 246-252

PARTIE I :

La super famille des calciprotéines

à main EF

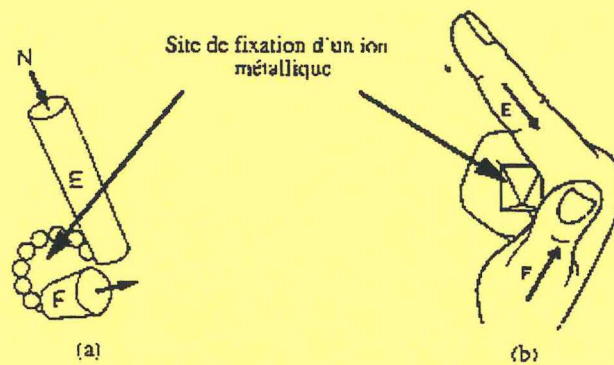


Table des matières

| | |
|---|-----------|
| I. La super famille des calciprotéines à main EF | 19 |
| A. Le motif main EF | 19 |
| 1. Structure | 19 |
| 2. Séquence primaire | 21 |
| B. La Parvalbumine : un modèle pour d'autres protéines à « main EF » | 25 |
| 1. Généralités | 25 |
| 2. Localisation et fonction | 25 |
| 3. Structure primaire | 31 |
| 4. Caractéristiques structurales : structure tertiaire | 33 |
| C. Autres calciprotéines étudiées: | 41 |
| 1. La troponine C | 41 |
| 2. la calmoduline | 45 |
| 3. protéine sarcoplasmique | 47 |
| D. Références bibliographiques | 50 |





I. La super famille des calciprotéines à main EF

A. Le motif main EF

1. Structure

La famille des calciprotéines à « main EF » tire son nom de l'organisation typique des sites de fixation du calcium. Ce motif fonctionnel ancestral, conservé au cours de l'évolution, a été appelé main EF, par Kretsinger (1), au vu de sa structure tridimensionnelle.

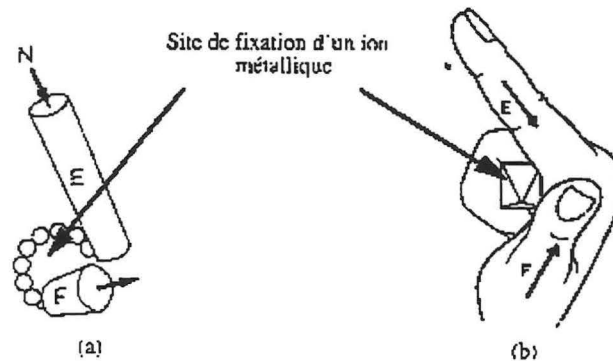


Figure I-1 : a) La structure hélice-boucle-hélice b) la « main EF »

Ce motif est constitué de 29 acides aminés, adoptant une conformation hélice- boucle- hélice. L'ensemble peut être schématisé comme le montre la Figure I-1 : le pouce et l'index de la main droite représentent chacun une hélice α tandis que le majeur replié représente la boucle liant le calcium. Le site de fixation proprement dit comprend douze acides aminés situés dans une boucle qui s'étend de part et d'autre de deux hélices α orientées à peu près perpendiculairement ($\approx 100^\circ$) l'une par rapport à l'autre. La coordination est assurée par des atomes d'oxygène provenant des groupes carboxylates

de résidus acides (Glu et Asp), de groupes carboxyles de liaisons amides, de groupes hydroxyles (sérine ou thréonine) ou de molécules d'eau.

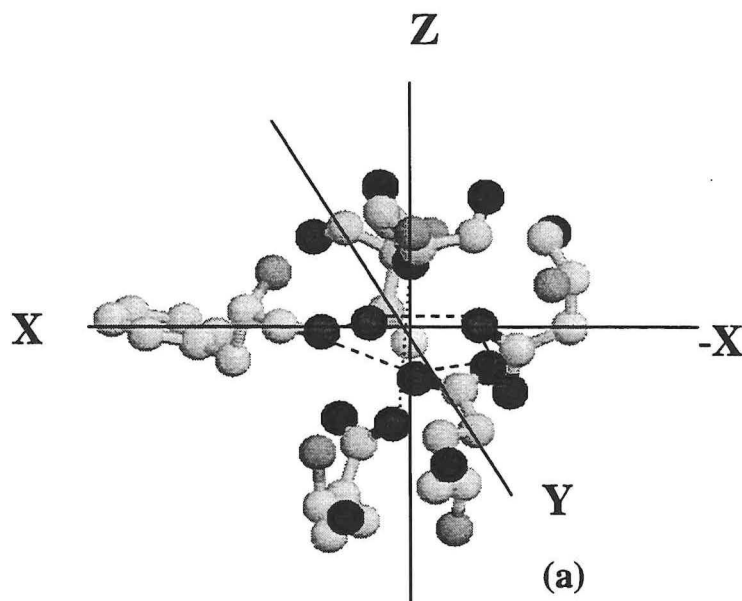


Figure I-2 : site EF_3 de la parvalbumine de brochet pI 4.10. (2) occupé par Ca^{2+} : Illustration de la disposition géométrique particulière des oxygènes selon une bipyramide à base pentagonale. Les oxygènes sont représentés par des sphères sombres.

Plus précisément, cette région de 29 résidus bien conservés inclut 9 résidus d'une première hélice, suivis de 12 résidus formant la séquence du site de fixation proprement dit et de 8 résidus d'une seconde hélice. Malgré le fait que les trois derniers résidus appartiennent au premier pas de la deuxième hélice, selon la convention définie par Kretsinger, la séquence directement impliquée dans la fixation est appelée « boucle EF » ou main EF canonique. Elle est constituée de 12 des 29 acides aminés du motif total (« main » EF). Les résidus en position relative 1-3-5-7-9-12 de la boucle EF coordonnent le cation dans les directions respectives $+X$, $+Y,+Z$, $-Y$, $-X$ et $-Z$ formant ainsi un arrangement octaédrique (voir Figure I-2). Les oxygènes coordonnant le Ca^{2+} sont au nombre de sept et forment une bipyramide, avec une légère distorsion de la base pentagonale. Cette géométrie est conservée dans toutes les structures cristallographiques connues de sites EF occupés par Ca^{2+} .



Aujourd'hui, la détermination de structures cristallographiques de près de 35 calcioprotéines à « main EF » (3) a permis de mettre en évidence un repliement tertiaire unique de ces calcioprotéines. De manière générale, toutes ces protéines possèdent au moins un domaine constitué de deux mains EF adjacentes, qui sont positionnées l'une par rapport à l'autre via un axe de pseudo-symétrie d'ordre deux.

2. Séquence primaire

Au vu des nombres de motifs « main EF » présents dans les calcioprotéines, Kretsinger a émis l'hypothèse que l'ensemble de la famille provient, chez les eucaryotes, d'un seul motif précurseur(4), responsable de la propriété de fixation du calcium. Le gène correspondant au motif main EF se serait dupliqué générant ainsi une calcioprotéine à deux motifs. Une duplication subséquente de ce gène aurait produit une protéine ancêtre des protéines à quatre domaines, ainsi que de celles à trois domaines, l'émergence de ces dernières étant due à la délétion, au hasard de l'évolution, d'un motif.

| Protéine | séquence d'acides aminés du motif EF ₃ | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|----------------------|---|---|---|--------|---|---|---|----|----|----------|----|----|----|----|----|----|----|---|----|----|---|---|---|---|---|---|---|---|
| position | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | | | | | | | | | | | | | | | | |
| parvalbumine | V | K | K | A | F | A | I | D* | Q | D* | K | S* | G | F* | I | E* | E | D | E* | L | K | L | F | L | Q | N | F | |
| calmoduline | F | K | E | A | F | S | L | F | D* | K | D* | G | D* | G | T* | I | T* | K | E* | L | G | T | V | M | R | S | L | |
| Troponine C | L | A | D | C | F | R | I | F | D* | K | N* | A | D* | G | F* | I | D* | I | E | E* | L | G | E | I | L | R | A | T |
| structure secondaire | Hélice 1 | | | boucle | | | | | | Hélice 2 | | | | | | | | | | | | | | | | | | |

Tableau I-1 : homologie de séquence entre les sites EF₃ de la parvalbumine de brochet, calmoduline humaine et troponine C de poulet. Les acides aminés notés * participent directement à la fixation du cation. Le caractère hydrophobe des acides aminés en gris est hautement conservé (exemple, en premier colonne : V,F,L).

Dans le cadre de cette hypothèse, la parvalbumine serait une protéine à trois domaines ancestraux nommés EF₂, EF₃, et EF₄. Cette dénomination a été adoptée du fait

de la forte homologie entre les sites de fixation de la parvalbumine et les sites EF3 et EF4 de la troponine C et de la calmoduline (cf. Tableau I-1), qui sont des protéines à quatre domaines. Dans le cas de la parvalbumine l'évolution aurait « œuvré » en éliminant le site EF1 et en éteignant l'activité du site calcique EF2.

La séquence consensus des 29 acides aminés qui constituent le motif main EF figure dans le Tableau I-2. Chaque portion d'hélice participe à un coeur hydrophobe de par 4 de ses acides aminés, en position relative -8,-4,-5 et -1 pour la première hélice et 13,16,17 et 20 pour la deuxième. Ces acides aminés hydrophobes sont localisés sur la même face de l'hélice et sont soit enfouis à l'intérieur du corps de la protéine, soit exposés au solvant au moment de la fixation de l'ion. Ils peuvent alors interagir avec des protéines effecteurs ou de petites molécules régulatrices (5). Dans la boucle EF, les positions 1-3-5-7-9 et 12 sont les résidus qui assurent la coordination de l'ion fixé. Trois positions parmi ces cinq sont tout particulièrement conservées. Le résidu en position 1 est dans tous les cas un acide aspartique. En position 3, l'acide aspartique survient avec une fréquence de 76% tandis que le résidu en position 12 est dans près de 92 % des cas un acide glutamique (cf. Tableau I-2). Par ailleurs, deux autres résidus sont aussi hautement conservés. La glycine en position 6 (96% des cas), assurerait un rôle dans le repliement caractéristique de cette boucle en contribuant à la formation d'un réseau de liaisons hydrogènes(5), alors que la chaîne hydrophobe du résidu 8 est enfouie, probablement pour stabiliser le système.

On dénombre actuellement, toutes espèces confondues, près de 273 séquences, identifiées comme étant membres de la super-famille des calciprotéines à mains EF(3). L'appartenance ou l'identification est relativement facile du fait que les protéines de cette famille possèdent toutes de 2 à 8 fois le domaine fonctionnel main EF.

A ce jour la famille des calciprotéines à mains EF est constituée de près de 39 sous-familles, dont bon nombre sont constituées, pour l'instant, d'une seule protéine. On trouve entre autres dans ces sous-groupes ceux de la parvalbumine, de la troponine C et de la calmoduline dont nous parlerons plus en détail ultérieurement.

Le motif main EF

| 1 ^{er} Hélice | -9 | -8 | -7 | -6 | -5 | -4 | -3 | -2 | -1 | | | |
|-----------------------------------|-----------------|-----------------|-------------------------------------|-----------------|--|----------------|---|-----------------|--|-----------------|-----------------|---------------------------|
| Préférence | Glu | Leu | Lys | Glu | Ala | Phe | Arg | Glu | Phe | | | |
| % | 58% | 27% | 32% | 41% | 30% | 50% | 21% | 26% | 29% | | | |
| Hydrophobicité | 6,4 | -2,2 | 6,7 | 5,9 | -1,3 | -2,2 | 7,3 | 2,2 | -1,9 | | | |
| volume moyen | 68,2 | 78,3 | 76,0 | 64,3 | 63,0 | 88,7 | 69,0 | 70,1 | 71,7 | | | |
| bouclé EF | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| coordination | Monodentate | | Mono | | Mono | | c=0 | | Mono | | | Bidentate |
| Préférence | Asp | Lys | Asp | Gly | Asp | Gly | Thr | Ile | Asp | Phe | Glu | Glu |
| % | 100% | 29% | 76% | 56% | 52% | 96% | 23% | 68% | 32% | 23% | 29% | 92% |
| Hydrophobicité moyenne | 9,7 | 3,6 | 0,1 | 3,4 | 7,7 | 0,4 | 3,4 | -3,8 | 6,2 | 0,8 | 5,9 | 7,9 |
| volume moyen en (Å ³) | 50,0 Asp 567 | 70,9 Lys 163 | 51,0 Asp 432 Asn 130 Ser 5 | 37,3 Gly 319 | 46,4 Asp 295 Ser 131 Asn 123 Thr 9 Gly 8 Glu 1 | 9,7 Gly 541 | 74,5 Thr 130 Phe 90 lys 70 Gln 54 Arg 27 Ser 26 Ile 15 Cys 13 Asp 11 Leu 11 Val 11 Ala 8 His 6 Met 5 Asn 3 | 79,6 Ile 384 | 45,2 Asp 181 Ser 116 Thr 79 Glu 65 Asn 57 Gln 9 Cys 4 | 73,2 phe 131 | 60,5 Glu 164 | 66,6 Glu 523 Asp 44 |
| seconde Helice | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | | | | |
| Préférence | Phe | Leu | Thr | Met | Met | Ala | Ser | Leu | | | | |
| % | 57% | 17% | 18% | 22% | 42% | 17% | 20% | 31% | | | | |
| Hydrophobicité moyenne | -2,7 | 2,8 | 3,8 | -2 | -2 | 7,5 | 53 | 0,2 | | | | |
| volume moyen | 88,7 | 68,3 | 62,9 | 72,0 | 78,0 | 64,8 | 55,9 | 77,9 | | | | |

Tableau I-2 : séquence consensus du motif main EF extrait de(5) : elle a été réalisée à partir de l'ensemble des séquences connues à l'époque (1994), soit 567 motifs EF. Pour chaque acide aminé coordonné au Ca²⁺ (position 1-3-5-7-9-12) le type et le nombre des acides aminés observés est donné.

B. La Parvalbumine : un modèle pour d'autres protéines à « main EF »

1. Généralités

L'ensemble des données expérimentales déjà accumulées, relatives à l'échange $\text{Ca}^{2+}/\text{Mg}^{2+}$, aussi bien sur le plan thermodynamique (constante d'affinité) que sur le plan structural (données cristallographiques et RMN) font de la parvalbumine un modèle de choix qui, avant d'envisager l'étude de systèmes moins connus, nous permettra dans un premier temps de confronter résultats théoriques et données expérimentales. Ainsi nous pourrons nous assurer de la pertinence de nos méthodes, avant de passer, dans un deuxième temps, à un registre plus " prédictif ", lors de l'étude d'autres calciprotéines.

2. Localisation et fonction

La parvalbumine est sans doute la calciprotéine la mieux caractérisée sur le plan expérimental. Elle a été la première de cette famille à être purifiée (6) et séquencée. Il s'agissait aussi de la première structure cristallographique (7) de calciprotéine à "main EF".

La parvalbumine a une importante affinité pour le Ca^{2+} (10^7 à 10^8 M^{-1}). Elle peut également fixer, mais dans une moindre mesure, d'autres cations comme le Mg^{2+} (10^4 M^{-1}) au niveau des sites de fixation du calcium. Comme nous allons le voir, il est fort probable que, dans les conditions physiologiques, en raison de la concentration relativement élevée en Mg^{2+} intracellulaire (1 à 3 mM), la Parvalbumine alterne entre une forme chargée en Ca^{2+} , que nous noterons, PaCa₂, et une forme chargée en Mg^{2+} que nous noterons, PaMg₂.

La parvalbumine est une protéine acide de faible poids moléculaire (11500 daltons), omniprésente chez les vertébrés (8,9) - Elle est présente en relativement grande quantité dans les muscles de Vertébrés inférieurs (10) ainsi qu'en quantité moindre dans ceux de grands Vertébrés (11) où on la trouve d'ordinaire dans une espèce donnée sous plusieurs isoformes (voir plus loin). Les parvalbumines sont absentes chez les Invertébrés. Toutefois on trouve chez ces derniers des calcioprotéines apparentées, mais de plus haut poids moléculaire (PM=20.000) appelées SCBP pour sarcoplasmic calcium binding proteins (12).

La parvalbumine est aussi présente dans bon nombre de tissus autres que musculaires (13-17), tels que la peau, les glandes endocrines (18), ou les neurones GABA- ergiques (19,20). A la fin des années 80, une hormone du thymus de poulet a été identifiée comme étant une parvalbumine (21). C'est apparemment la première fonction spécifique décrite pour une parvalbumine. On croit cependant que les parvalbumines sont essentiellement impliquées dans l'homéostasie calcique des cellules normales et tumorales (22). Ces dernières années, après sa découverte dans une sous- population de neurones du système nerveux central (19) (voir les études de localisation des parvalbumines et d'autre calci-protéines connues dans le cerveau (23,24)) l'intérêt qu'on porte à la parvalbumine s'est particulièrement accru. Même si la question du rôle musculaire de la parvalbumine reste encore ouverte, un certain nombre d'études ont largement réduit le nombre des possibilités et suggèrent que cette protéine constituerait un système régulateur non interactif au niveau du Ca^{2+} sarcoplasmique, dont le rôle est essentiel lors de la relaxation des fibres musculaires rapides.

Dans les conditions physiologiques, la Parvalbumine ressent l'environnement ionique complexe de la cellule, avec entre autres une concentration en Mg^{2+} qui, comme nous

l'avons évoqué, est de l'ordre du millimolaire, et une concentration en Ca^{2+} fluctuant entre 10^{-7} et 10^{-5} M en fonction de l'état de la fibre musculaire (25). Compte tenu de ces éléments, il est donc possible que la Parvalbumine alterne entre la forme chargée en Ca^{2+} et la forme chargée en Mg^{2+} au cours du cycle contraction-relaxation du muscle.

En effet, la Parvalbumine possède des sites de qui sont plus affines que les sites Ca-Mg de la troponine C (voir Tableau I-3). Ils sont cependant moins affines que les sites calcium spécifique des pompes du reticulum sarcoplasmique¹. La parvalbumine pourrait donc fonctionner comme une « navette à calcium », transportant ce dernier de la troponine aux pompes à Ca^{2+} pendant la relaxation musculaire(26).

| sites \ K_a (Mol) | $K_{a\text{Ca}^{2+}}$ | $K_{a\text{Mg}^{2+}}$ | Réf |
|------------------------------|-----------------------|-----------------------|------|
| Parvalbumine | 10^9 à 10^8 | 10^5 à 10^6 | (34) |
| Troponine C | | | |
| spécifiques | 10^5 | - | (35) |
| site de haute affinité Ca/Mg | 10^7 | 10^3 | (35) |

Tableau I-3: ordre de grandeur des constantes d'affinités calcique et magnésique des sites de fixation de la troponine C et de la parvalbumine.

¹ Ces dernières sont des protéines majeures du reticulum sarcoplasmique (60 à 85% des protéines de la membrane) transportant le calcium contre un gradient de concentration de 1000 environ.

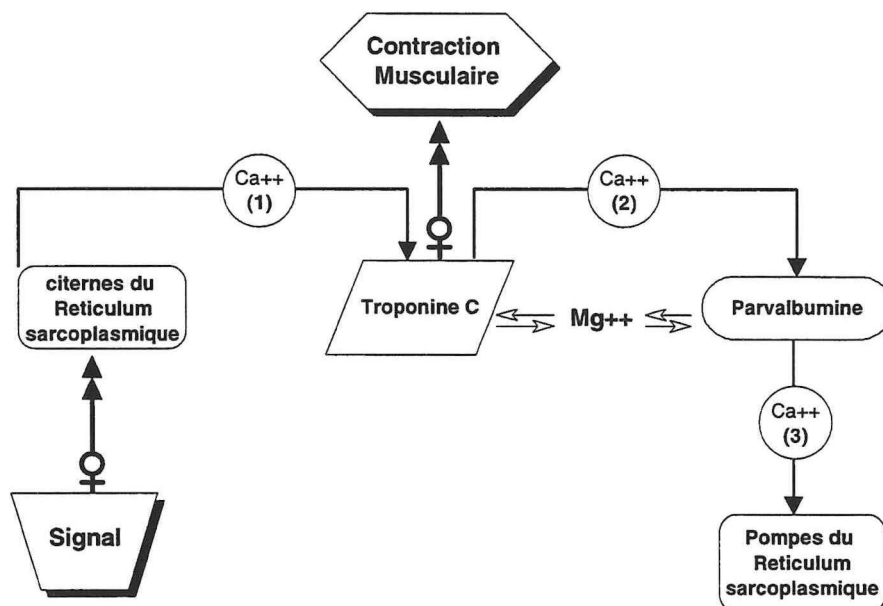


Figure I-3 : mouvement du calcium, de sa libération des citernes du reticulum sarcoplasmique(1) aux pompes à calcium du reticulum sarcoplasmique. (1)→Ca²⁺ libéré par le reticulum sarcoplasmique et fixé par les sites Ca²⁺ spécifiques de la troponine C ; (2)→ libération de la troponine C et fixation par la parvalbumine;(3)→ dissociation du complexe parvalbumine-Ca²⁺ et réintégration dans le reticulum par les pompes. Il est à noter que le Ca²⁺ libre est en équilibre avec le Ca²⁺ lié à la troponine C et à la parvalbumine

Même quand la concentration de calcium du myoplasme est abaissée par les pompes du reticulum sarcoplasmique au cours de la relaxation, sa concentration locale est encore relativement haute au voisinage de la troponine, où elle dépend alors du rapport des sites de fixation occupés et non occupés dans la molécule de troponine C. Au voisinage des pompes à Ca²⁺, la concentration en Ca²⁺ est plus réduite, le complexe calcium parvalbumine se décompose et le cation pourra être lié par les protéines translocatrices du reticulum sarcoplasmique. Comme le montre la Figure I-3, le résultat de ces diverses réactions est un flux unidirectionnel de Ca²⁺ des sites calciques de la troponine vers ceux du reticulum sarcoplasmique.

Ces considérations ont conduit Gilles et ces collaborateurs (27) à la proposition d'un modèle concernant l'action de la parvalbumine. Ce modèle est basé sur des considérations cinétiques. Une fois le Ca²⁺ libéré par le reticulum sarcoplasmique, la parvalbumine ne rentre pas immédiatement en compétition avec les sites calciques de la

troponine C pour la fixation du Ca^{2+} activateur. En première approche, il peut paraître paradoxal que l'affinité des sites de fixation de la parvalbumine soit beaucoup plus élevée que celle des sites de fixation spécifiques de la troponine C. Or on doit cependant considérer que la concentration en Mg^{2+} du muscle strié au repos est proche de 1 à 3 mM, soit 10^4 fois plus importante que celle du Ca^{2+} . Il s'ensuit que, étant donné le faible niveau de Ca^{2+} ($\approx 10^{-8}$ M) du muscle au repos, les sites de la parvalbumine doivent être presque entièrement ($\approx 90\%$) occupés par Mg^{2+} . Ainsi, à moins d'éjecter Mg^{2+} , la parvalbumine ne peut pas lier Ca^{2+} immédiatement. Comme la dissociation du complexe parvalbumine- Mg^{2+} est un processus relativement lent(28), le Ca^{2+} est donc fixé avec un certain délai après sa libération du reticulum sarcoplasmique. Ainsi, la parvalbumine exercerait son activité relaxante seulement après le temps de latence nécessaire à la dissociation du complexe Parvalbumine-magnésium. Après cette dissociation, la parvalbumine limiterait le taux de calcium libre et par voie de conséquence faciliterait la décomposition du complexe troponine C-calcium.

L'élimination du calcium par la parvalbumine ne peut se concevoir que si ce processus est plus rapide que son retour dans le reticulum sarcoplasmique. Si tel est le cas, cela déterminerait par exemple la vitesse de relaxation des muscles des petits animaux à sang froid. Ceci serait en effet particulièrement vrai à des températures ambiantes basses, l'activité des ATPases et des pompes du reticulum sarcoplasmique étant alors considérablement réduite. De fait, des poissons vivant dans l'eau froide tels que le brochet ou la carpe ont des muscles qui contiennent des taux de parvalbumine particulièrement élevés (jusqu'à $6 \cdot 10^{-4}$ M dans le brochet selon Baron et al 1985). Dans les muscles des animaux à sang chaud, l'activité des ATPases transporteurs de calcium est plus importante et la quantité de parvalbumine est habituellement plus faible, mais ici

aussi, il y a corrélation entre la concentration de parvalbumine et la vitesse de relaxation des muscles (29), ce qui suggère encore que la parvalbumine est impliquée dans leur vitesse de relaxation, en agissant comme piège à calcium temporaire.

Cependant, - comme l'ont souligné Heizman et Celio (8)- bien que la parvalbumine ait été décrite comme une protéine ancillaire du fonctionnement musculaire (25), il est bien possible que, compte tenu de la large distribution tissulaire de cette protéine, l'étendue de ses fonctions ait été sous-évaluée.

3. Structure primaire

Sur la base de leur séquence en acides aminés, les parvalbumines ont été subdivisées en deux séries phylogénétiques distinctes nommées α et β . Ainsi, les parvalbumines ont été la première sous-famille de calciprotéines pour laquelle plusieurs isoformes ont été localisées dans un même organisme (4). L'étude de leurs séquences (le plus souvent 108 à 109 résidus), a permis de mettre en évidence que ces deux isoformes se différencient l'une de l'autre par au moins onze résidus caractéristiques (30).

Ces résidus ainsi que leur position figurent dans le tableau ci-dessous :

| Positions des différences en AA | Parvalbumine α | Parvalbumine β |
|---------------------------------|-----------------------|----------------------|
| entre -1 et- 8 | Met | Ileu |
| Domaine 2 : | | |
| position 4 | Lys | Ala |
| position 9 | Phe | Cys |
| position 21 | Lys | Thre |
| position 26 | Leu | Lys |
| position 2 +1 | Lys | Ala |
| Domaine 3 : | | |
| position 7 | His | Gly ou Lys |
| position 25 | Ile | Phe |
| position 3+3 | Glu | Ser ou Asp |
| Domaine 4 : | | |
| position 5 | Leu | Phe |
| position 11 | Lys | Ser |

Tableau I-4 : différences entre les séquences des parvalbumine α et β (30). La nomenclature des positions est conforme à celle définie par Krestinger.

Ces différences de séquence se traduisent, entre autres, par la présence chez les parvalbumines de type β d'un 3^{ème} site de fixation cationique, dont l'affinité et la sélectivité sont cependant peu importantes, au moins en ce qui concerne les cations divalents (31). Nous verrons plus loin que ce troisième site est proche du site de fixation primaire à haute affinité EF3 .

Du point de vue de l'évolution, l'existence des parvalbumines α et β pourrait être due à la présence chez l'ancêtre des Vertébrés d'au moins deux formes de parvalbumine (30). Car on note leur présence aussi bien chez les Poissons, les Amphibiens et les Mammifères. La forme α est cependant absente chez les Reptiles et chez les Oiseaux. Selon Moncrief & Al (30), l'absence de deux isoformes dans ces classes d'animaux reflète soit la perte de l'une d'elles au cours de l'évolution, soit le côté parcellaire des bases de séquences actuelles. Par exemple, dans un premier temps, aucune parvalbumine β n'avait été décrite chez les Mammifères. Mais dans un deuxième temps une protéine tumorale, l'oncomoduline, a été découverte (32). Celle-ci présente une forte homologie avec les parvalbumines dans leur ensemble, mais plus particulièrement avec les parvalbumines β . Il est important de souligner que, sur le plan fonctionnel, contrairement aux autres parvalbumines, l'oncomoduline est capable, comme la calmoduline, d'activer le cycle nucléotidique de la phosphodiesterase(33), de manière beaucoup moins efficace toutefois.

4. Caractéristiques structurales : structure tertiaire

Malgré la mise en évidence de l'affinité de plusieurs calciprotéines à main EF² pour le Mg²⁺, la parvalbumine était la seule calciprotéine pour laquelle une structure cristallographique (2) chargée sous forme Mg²⁺ était disponible lorsque nous avons entrepris notre travail. Il s'agit de la structure "4pal" ci dessous. C'est cette structure que nous avons tout d'abord étudiée.

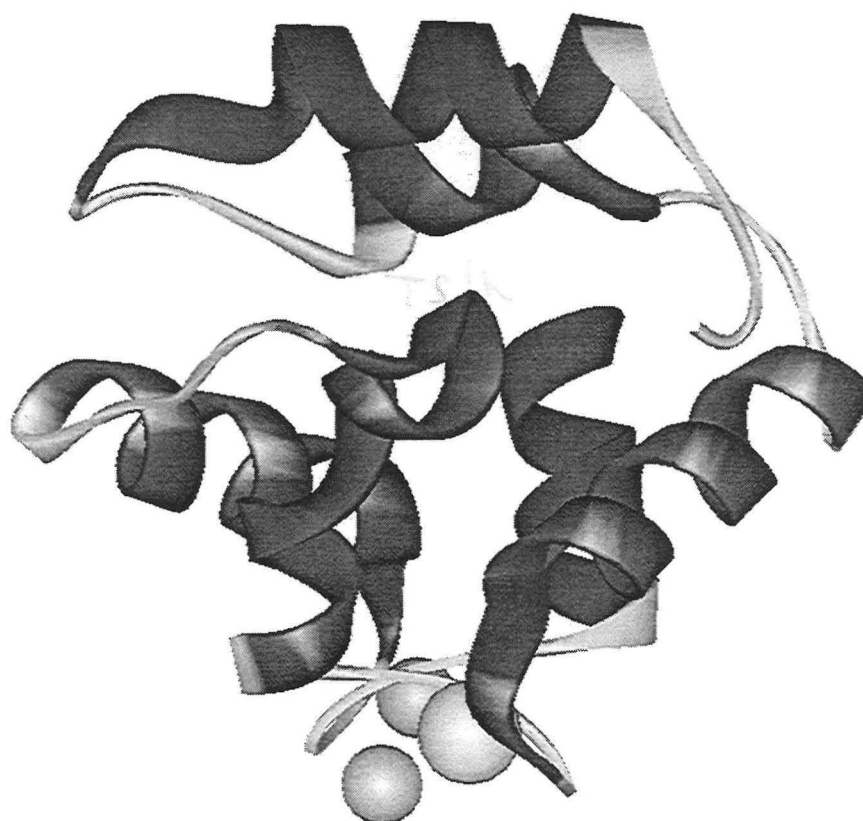


Figure I-4 : La structure cristallographique 4pal (2) : illustration du repliement de la parvalbumine

² telles que la troponine C ou la calmoduline ou la protéine sarcoplasmique

a) Forme générale du repliement

Cette structure cristallographique nous permet de donner une description détaillée du repliement typique d'une calciprotéine à main EF, et par là même de la coordination de l'ion magnésium au sein d'un site de fixation. L'étude de la structure aux rayons X de la parvalbumine de la carpe (7) et plus tard celle de parvalbumines d'autres espèces ont montré que ces protéines ont une structure secondaire constituée de six hélices α , qui adoptent le repliement schématisé ci-dessous.

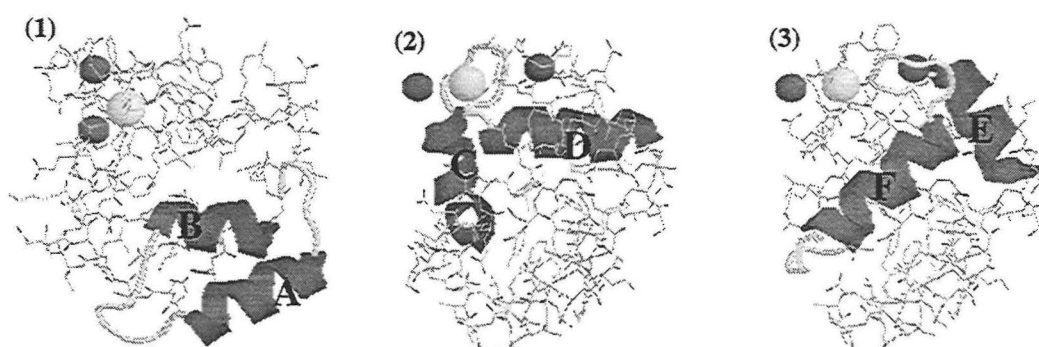


Figure I-5 : Arrangement des hélices dans la parvalbumine. Ces hélices sont nommées de A à F

Les deux hélices A et B sont antiparallèles et les deux paires d'hélices C-D et E-F adoptent la conformation caractéristique du motif « main EF », précédemment décrit. Contrairement aux boucles CD et EF, la boucle inter-hélices AB ne fixe pas le Ca^{2+} . Les deux sites primaires de fixation du calcium CD et EF sont aussi nommés EF3 et EF4. En plus de cet arrangement particulier, la parvalbumine a un cœur hydrophobe, qui implique près de 20 chaînes latérales de résidus non polaires (phénylalanine, isoleucine, leucine, valine). Des études de dichroïsme circulaire, confirmées par la suite par une étude de cristallographie, ont mis en évidence que le squelette carboné se comporte comme un corps rigide, dans lequel les chaînes latérales de phénylalanine ont une grande mobilité (des « flip-flop » autour des liaisons $\text{C}\beta\text{-C}1$). Par comparaison à d'autres

protéines globulaires la fréquence de ces flips semble très rapide dans la parvalbumine. L'analyse RMN de parvalbumines de différentes espèces a permis de conclure que la fluidité des chaînes latérales de phénylalanine du corps hydrophobe est une caractéristique structurale des parvalbumines. Au vu de ces éléments il a été émis l'hypothèse que la dynamique interne du cœur hydrophobe de la parvalbumine serait directement en relation avec sa fonction et pourrait être reliée à la capacité qu'a cette protéine de fixer les cations Ca^{2+} et Mg^{2+} (34). Ces alcalino-terreux se différencient essentiellement par leur taille³. Les variations locales imposées par la différence de taille de ces cations au niveau des sites de fixation pourraient être accompagnées par d'autres ajustements conformationnels, plus subtils, se manifestant donc dans le cadre de la dynamique de son coeur hydrophobe.

b) Adaptation de la protéine au type de cation

Declercq et ses collaborateurs (2) ont déterminé pour la première fois des structures de parvalbumine (de brochet Pi 4.10) chargées avec divers cations divalents, dont la structure "4pal". Il est à noter que cette parvalbumine est de type β et possède à ce titre un troisième site de fixation de cation. Le nom et l'occupation de l'ensemble des sites de fixation de ces structures figurent dans le Tableau I-5.

| code | Forme cristal | sites primaire | | site secondaire | Résolution |
|------|-------------------------------------|------------------|------------------|------------------------------|------------|
| | | EF3 | EF4 | 3 ^{eme} site | |
| 1PAL | Pa.Ca ₂ .NH ₄ | Ca ²⁺ | Ca ²⁺ | NH ₄ ⁺ | 1.65 |
| 3PAL | Pa.Ca ₂ .Mg | Ca ²⁺ | Ca ²⁺ | Mg ²⁺ | 2.4 |
| 4PAL | Pa.Ca.Mg ₂ | Ca ²⁺ | Mg ²⁺ | Mg ²⁺ | 1.75 |
| 2PAL | Pa.Mn ₂ . Mn | Mn ²⁺ | Mn ²⁺ | Mn ²⁺ | 1.75 |

Tableau I-5 : occupation ionique des sites de fixation de la parvalbumine de carpe (structures cristallographiques déterminées par (2))

³ rayon ionique de 0.98 et 0.65 Å pour Ca²⁺ et Mg²⁺, respectivement

Cette étude a mis en évidence une contraction de la sphère de coordination mais ceci sans grand changement notable au niveau de la structure globale de la protéine. Le rôle du cœur hydrophobe précédemment évoqué dans l'adaptation au type de cation reste donc à confirmer. En effet, comme en témoigne le Tableau I-6, malgré la substitution des ions au sein des sites, les structures sont presque parfaitement superposables. Seuls quelques ajustements relativement localisés aux alentours des sites de fixation des cations ont été mis en évidence.

| | Pa.Ca ₂ .NH ₄ | Pa.Mn ₂ .Mn | Pa.Ca ₂ .Mg | Pa.Ca.Mg.Mg | Carpe |
|-------------------------------------|-------------------------------------|------------------------|------------------------|-------------|-------|
| Pa.Ca ₂ .NH ₄ | - | 0.36 | 0.33 | 0.54 | 0.43 |
| Pa.Mn ₂ .Mn | 0.77 | - | 0.35 | 0.41 | 0.46 |
| (N° résidu) | 20 | | | | |
| Pa.Ca ₂ .Mg | 0.62 | 0.79 | - | 0.42 | 0.47 |
| (N° résidu) | 73 | 17 | | | |
| Pa.Ca.Mg.Mg | 1.17 | 0.89 | 1.15 | - | 0.59 |
| (N° résidu) | 89 | 73 | 92 | | |
| Carpe | 1.04 | 1.23 | 1.00 | 1.70 | - |
| (N° résidu) | 38 | 105 | 105 | 105 | |

Tableau I-6 : résultats des superpositions des C α des résidus 6 à 106 des 4 structures à haute résolution de parvalbumine de brochet Pi 4.10 ainsi que la structure de parvalbumine de carpe Pi4.25 également à haute résolution déterminée par Swain et al 1989. La partie supérieure du tableau donne la déviation en (Å) entre les atomes C α . La partie inférieure du tableau donne la plus importante déviation (Å) ainsi que sa position (N° résidu) en italiques. Compte tenu de l'incertitude sur les positions atomiques, les erreurs associées à ces déviations sont comprises entre 0.25 et 0.30 Å (2).

Concernant les ajustement locaux, les substitutions successives du Ca²⁺ par Mn²⁺ et Mg²⁺ au sein des sites primaires de la parvalbumine ont montré une contraction générale de la sphère de coordination des cations, qui se traduit par une diminution de la distance moyenne⁴ cation - oxygène, d'environ 2.4 Å pour le Ca²⁺, à 2.1 Å pour Mn²⁺ et Mg²⁺. Les

⁴ Ces moyennes ont été réalisées sur l'ensemble des sites des structures de la référence (2)

distances cation-oxygène relatives aux différentes structures figurent dans les Tableau I-7 et Tableau I-8 ci-dessous.

| Site CD | position | Pa Ca ₂ NH ₄ | Pa Mn ₂ . Mn | Pa Ca.Mg.Mg |
|------------------------|----------|------------------------------------|-------------------------|-------------|
| Asp51 O _{δ1} | +X | 2.21 | 1.88 | 2.17 |
| Asp53 O _{δ1} | +Y | 2.25 | 2.15 | 2.34 |
| Ser55 O _{δ1} | +Z | 2.31 | 2.32 | 2.55 |
| Phe57 O | -Y | 2.30 | 2.17 | 2.43 |
| Glu59 O _{ε1} | -X | 2.42 | - | 2.33 |
| Glu101 O _{ε1} | -Z | 2.50 | 2.37 | 2.70 |
| O _{ε2} | | 2.52 | - | 2.50 |
| < moyenne > | Å | <2.36> | <2.12> | <2.43> |
| < variance > | Å | <0.10> | <0.10> | <0.10> |

Tableau I-7 : Coordination des cations fixés au site primaire (EF3) (distances < 3 Å) (2),

| Site EF | position | Pa Ca ₂ NH ₄ | Pa Mn ₂ . Mn | Pa Ca.Mg.Mg |
|------------------------|----------|------------------------------------|-------------------------|-------------|
| Asp90 O _{δ2} | +X | 2.35 | 2.09 | 2.07 |
| Asp92 O _{δ1} | +Y | 2.37 | 2.09 | 2.09 |
| Asp94 O _{δ1} | +Z | 2.20 | 2.11 | 2.18 |
| Met96 O | -Y | 2.32 | 2.15 | 2.14 |
| H ₂ O (246) | -X | 2.47 | 2.24 | 2.15 |
| Glu101 O _{ε1} | -Z | 2.47 | 2.06 | 2.04 |
| O _{ε2} | | 2.15 | - | - |
| < moyenne > | Å | <2.33> | <2.12> | <2.11> |
| < variance > | Å | <0.10> | <0.10> | <0.10> |

Tableau I-8 : Coordination des cations fixés au site primaire (EF4) (distances < 3 Å) (2)

Comme le montrent ces tableaux, la diminution de taille du cation et la contraction de la protéine s'accompagnent également d'une diminution du nombre d'oxygènes coordonnés au cation, qui passe de 7 pour Ca²⁺ à 6 pour Mg²⁺ et Mn²⁺.

Cette modification du nombre de voisins s'effectue dans tous les cas suite à la perte d'un ligand bidentate (Glu 101 ou 62 en position relative 12), qui devient monodentate. La comparaison des différentes structures a mis en évidence que cet acide aminé s'adapte au type de cation selon deux mécanismes différents. En présence de

Mg^{2+} , la chaîne latérale du Glu 101 passe de la conformation énergétiquement favorable gauche (+) à la forme énergétiquement défavorable gauche (-) en présence de Ca^{2+} suivant le mécanisme décrit Figure I-6, à savoir que, alors que l'angle χ_2 reste invariant, l'angle χ_1 bascule de 120° , ce mouvement étant accompagné par un léger ajustement de l'angle χ_3 .

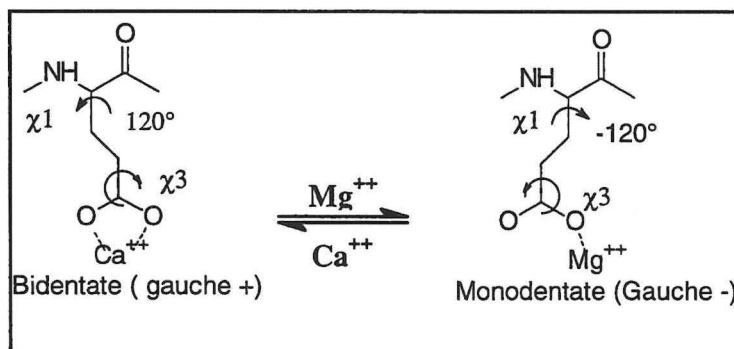


Figure I-6 : réarrangement de la chaîne latérale de l'acide glutamique en position relative 12 en fonction du type de cation complexé au site de fixation.

Une étude de RMN 2D du proton des structures PaCa2 et PaMg2 de brochet Pi 5.0 (35) a complété cette information cristallographique en mettant en évidence un arrangement symétrique des sites primaires EF3 et EF4 aussi bien dans leur forme chargée en Ca^{2+} que dans leur forme chargée en Mg^{2+} . Un examen approfondi des constantes de couplage J_3 associées aux dièdres χ_1 des deux chaînes latérales des acides glutamiques en position relative 12 (Glu 62 et 101) semble en effet indiquer un réarrangement semblable à celui observé lors de l'étude cristallographique de Declercq et Coll(2) dans le site EF (cf. Figure I-6). Il semble donc que cette adaptation du Glu en position relative 12 intervienne bien dans les deux sites primaires de la parvalbumine en présence de Mg^{2+} .

Il faut cependant noter que, bien qu'elle fasse intervenir toujours le Glu en position relative 12, dans la structure Mn^{2+} l'adaptation de la chaîne latérale est différente. Cette fois le passage monodentate bidentate fait intervenir un ajustement

concerté des trois angles dièdres de cette chaîne latérale et non pas une variation de 120° de l'angle χ_1 .

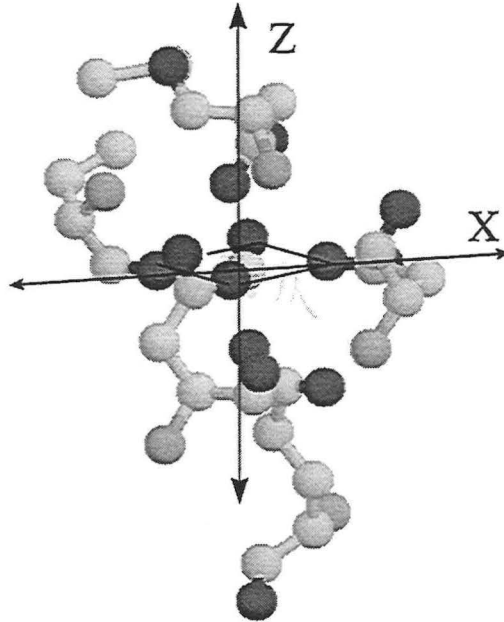


Figure I-7 : géométrie du site EF de la parvalbumine occupé par Mg^{2+} : les oxygènes forment un octaèdre quasi régulier

Quoiqu'il en soit, le réarrangement de l'acide glutamique en position 12 semble bien être un événement clé, qui permet à la protéine de maintenir avec Mg^{2+} ou Mn^{2+} le plus possible de liaisons, parmi celles participant à la fixation du Ca^{2+} . Inversement, il n'y a dans ce cas besoin d'aucune molécule d'eau supplémentaire au sein de la sphère de coordination lors de la fixation du Ca^{2+} . Ce phénomène est parfaitement mis en évidence dans la forme semi-magnésique Pa.Ca.Mg.Mg (4pal). Le site EF qui est alors occupé par un Ca^{2+} adopte une coordination à sept oxygènes, en forme de bipyramide à base pentagonale, alors que le site EF3, qui est occupé par un Mg^{2+} , a un arrangement quasi octaédrique avec six oxygènes autour du magnésium central. (cf. Figure I-7 :).

Pour finir, l'étude de Declercq et al a permis de lever les ambiguïtés concernant le troisième site de fixation associé à la parvalbumine β , puisque cette étude a permis de déterminer avec précision la position du site secondaire présent dans les parvalbumines β ainsi que son degré d'hydratation, et ceci pour les cations Ca^{2+} , Mg^{2+} , Mn^{2+} .

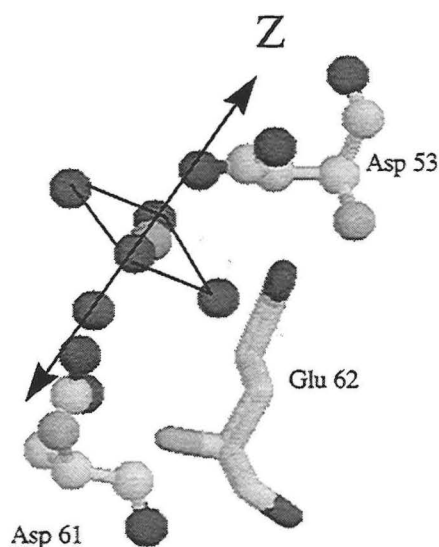


Figure I-8 : représentation du 3^{ème} site de fixation de la structure 4pal (Pa.Ca.Mg.Mg) : en l'occurrence, le Mg^{2+} est coordonné à l'acide aspartique en position 53 ainsi qu'à cinq molécules d'eau cristallographiques.

Ce 3^{ème} site s'avère être voisin du site primaire EF3. Ces sites possèdent un résidu en commun - à savoir l'acide aspartique en position 53. De plus, il est apparu que l'existence de ce 3^{ème} site est liée essentiellement à la présence dans la série β d'un acide aspartique en position 61 (il stabilise une molécule d'eau diamétralement opposée à Asp53 (cf. Figure I-8)).

C. Autres calciprotéines étudiées:

comme nous l'avons mentionné, la famille des calciprotéines à main EF est composée d'un grand nombre de protéines. Si le choix de la parvalbumine comme premier modèle est évident compte tenu du nombre de données expérimentales disponibles, le choix d'autres systèmes est, lui, plus arbitraire. Le choix des autres calciprotéines étudiées a été fait selon les critères suivants:

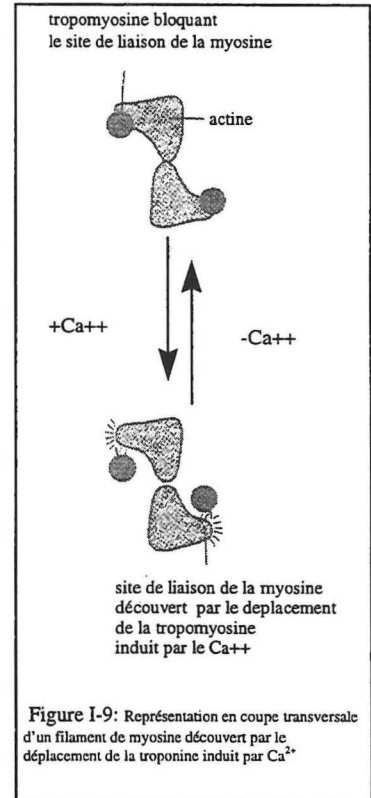
- i) la structure de la protéine complexée au calcium doit avoir une résolution inférieure ou égale à 2.0 Å. Ceci pour des raisons pratiques liées à la qualité des simulations que nous souhaitons entreprendre.
- ii) Nous nous sommes assurés que les calciprotéines choisies ont une affinité mesurable (et connue) pour le Mg^{2+} (in vitro).

La parvalbumine répond bien sûr à ces critères. Ces derniers nous ont conduit à étudier les calciprotéines suivantes.

1. La troponine C

Comme nous l'avons évoqué lors de la description du rôle de la parvalbumine, la troponine C est elle aussi impliquée dans la contraction musculaire. S. Ebashi (36) découvrit, en effet, en 1970 que l'action du calcium dans la contraction musculaire était due à la fixation de l'ion sur cette calciprotéine.

Une fois libéré, le calcium se fixe en effet sur la troponine qui constitue le récepteur spécifique du calcium responsable de la contraction de la cellule musculaire strié. Cette protéine comprend trois sous unités, nommées C, I et T(37). La troponine C, protéine d'environ 160 résidus, possède quatre sites de fixation du calcium, deux sites de haute affinité (cf. Tableau I-3) qui comme ceux de la parvalbumine fixent de manière compétitive Ca^{2+} et Mg^{2+} et deux sites de basse affinité spécifiques du Ca^{2+} . Dans le muscle strié, cette molécule joue le rôle d'un « interrupteur », en régulant l'activité de la troponine I, qui se lie à l'actine et inhibe son interaction avec la myosine



en fonction de la concentration en calcium. Lors de sa fixation sur les sites de basse affinité, le calcium induit un changement d'interaction qui assure une cohésion plus forte entre la troponine C et la troponine I, qui se sépare alors de l'actine (cf. Figure I-9 et Figure I-10)

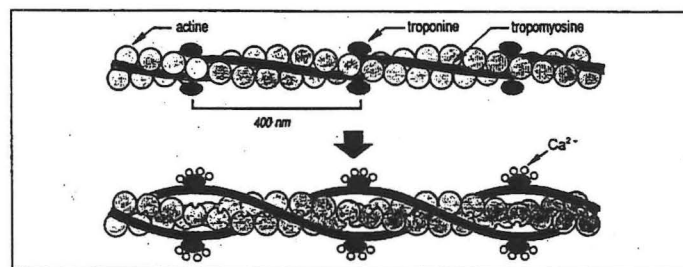


Figure I-10 : Fixation du Ca^{2+} sur la troponine C lors de la contraction musculaire

La troponine T, liée à la tropomyosine, permet la cohésion de l'ensemble troponine - tropomyosine. Les têtes de myosine se fixent ensuite sur les molécules d'actine libres en face desquelles elles se trouvent. Les deux protéines associées déclenchent l'hydrolyse de l'ATP, libérant ainsi l'énergie suffisante à la myosine pour effectuer un mouvement.

Ce mouvement élémentaire provoque le glissement des filaments les uns par rapport aux autres ou, si la résistance au mouvement est trop importante, le développement d'une force (cf. Figure I-11).

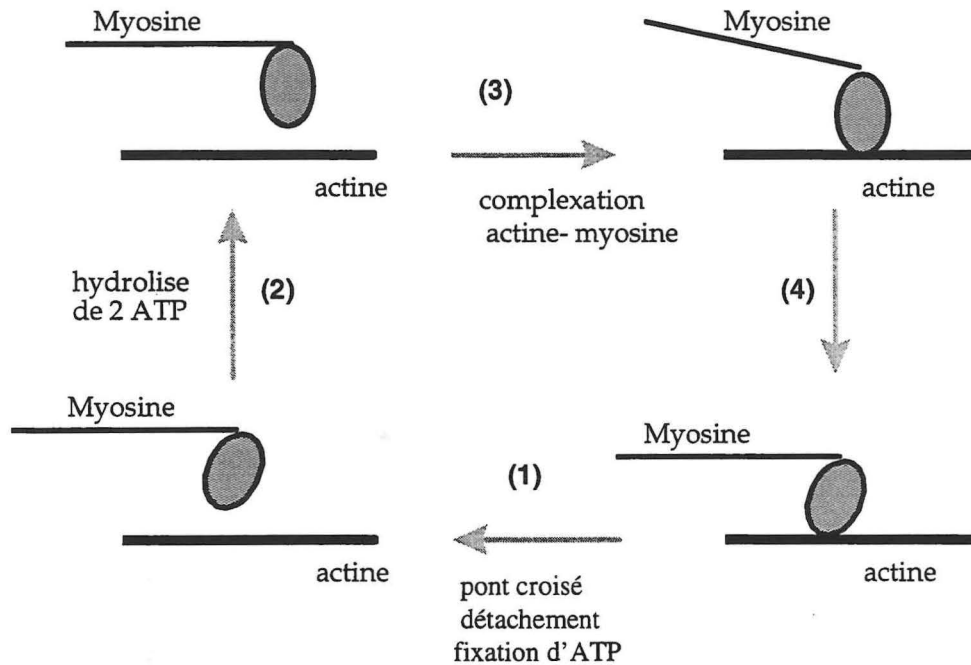


Figure I-11 : cycle élémentaire d'interaction entre tête de myosine et actine

1 → la fixation d'une molécule d'ATP induit la dissociation rapide des ponts actine-myosine 2 → alors que 1 ATP est hydrolysé les têtes de myosine libres se mettent en position pour se fixer à l'actine 3 → le pont myosine libéré avec l'énergie produite au cours de l'hydrolyse se fixe à nouveau au filament d'actine 4 → le pont croisé génère une force, et l'actine déplace les produits de l'hydrolyse (ADP et Pi) des ponts croisés de myosine ; c'est l'étape limitante de la contraction. Le pont croisé actine-myosine est alors de nouveau capable de fixer un ATP pour effectuer l'étape 1

La relaxation intervient quand la concentration en calcium s'abaisse. Le retour ou l'entrée du calcium dans le reticulum sarcoplasmique dépend des « pompes » ATPasiques dépendantes du calcium.

a) structure tridimensionnelle

La troponine C est une molécule constituée de huit hélices α qui adopte une forme d'haltère de 70 Å de long. Les lobes N et C terminaux sont reliés par une hélice α de 30 résidus (cf. Figure I-12). Chaque domaine a un rayon moyen d'environ 17 Å, leurs centres sont séparés d'environ 44Å. Chaque lobe est constitué d'une paire de mains EF. Lors de notre étude nous avons travaillé à partir de la structure cristallographique 1top (38) issue du muscle strié de poulet. Cette structure est actuellement la plus précise des structures RX de troponine C isolée disponibles. Elle a une résolution de 1.78 Å. Les sites de haute affinité, qui fixent Ca^{2+} et Mg^{2+} , sont situés dans le lobe N-terminal.

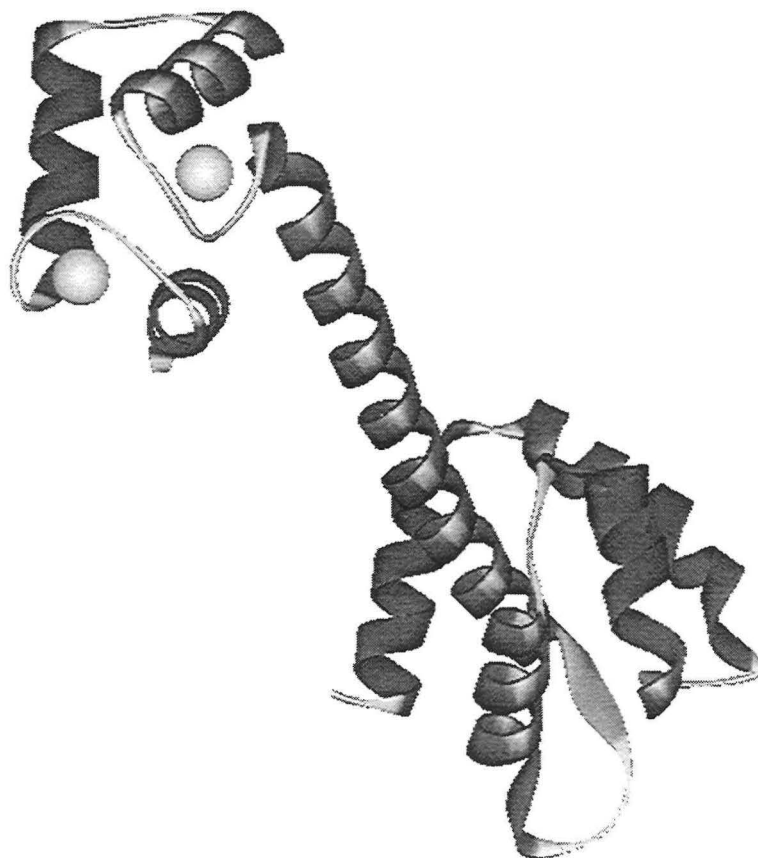


Figure I-12 : Structure: Troponine C de muscle squelettique de poulet (gallus gallus) (Résolution: 1.78Å). (38)

Il est important de noter que dans la structure 1top seuls les sites du lobe N-terminal sont occupés par un cation Ca^{2+} . D'après les auteurs, l'absence des cations des sites de basse affinité serait due aux conditions de cristallisation qui nécessitent un pH très acide.

2. la calmoduline

a) fonction

Dans les années 80, W. Y. Cheung⁽³⁹⁾ au St. Jude Children Research Hospital de Memphis et S. Kakiuchi à l'université d'Osaka montraient que l'enzyme qui dégrade l'AMP cyclique est activée par le calcium et par une calciprotéine, qui a été appelée calmoduline pour « calcium modulateur ». Depuis, on a démontré que la calmoduline était impliquée dans bon nombre de métabolismes. Si bien qu'elle nous apparaît aujourd'hui comme un senseur calcium peu spécialisé, présent dans toutes les cellules eucaryotes. La calmoduline est constituée de 148 acides aminés. Elle présente une forte homologie de séquence avec la troponine C. La fixation du calcium donne à cette protéine sa configuration fonctionnelle grâce à laquelle elle est capable de reconnaître une grande variété de protéines cibles.

b) structure tridimensionnelle

La calmoduline est en effet capable de subir de grands changements de conformation pour complexer une protéine. Sa structure est connue avec une forme ouverte et avec une forme fermée dans laquelle des acides aminés hydrophobes des deux lobes rentrent en contact grâce à une pliure de l'hélice qui relie les domaines N et C-terminaux. Lors de notre étude nous avons utilisé la structure 1cll (40) de la calmoduline humaine déterminée à la résolution de 1.7 Å (cf. Figure I-13).

Dans leur forme ouverte, les calmodulines adoptent un repliement proche de celui précédemment décrit pour la troponine C. Néanmoins, mis à part le fait que les quatre sites de la calmoduline sont occupés par Ca^{2+} , on note trois différences structurales majeures entre la calmoduline et la troponine C (41): i) Le bras N terminal en hélice de la troponine C n'a pas d'équivalent dans la calmoduline. ii) Il y a 3 résidus en moins dans l'hélice centrale inter-lobes, ce qui est responsable du fait que la taille de la calmoduline est inférieure de 5 Å à celle de la troponine C. iii) Après superposition des domaines C-terminaux de la troponine C et de la calmoduline, il y a une différence d'orientation de 60° entre les lobes N-terminaux des deux protéines.

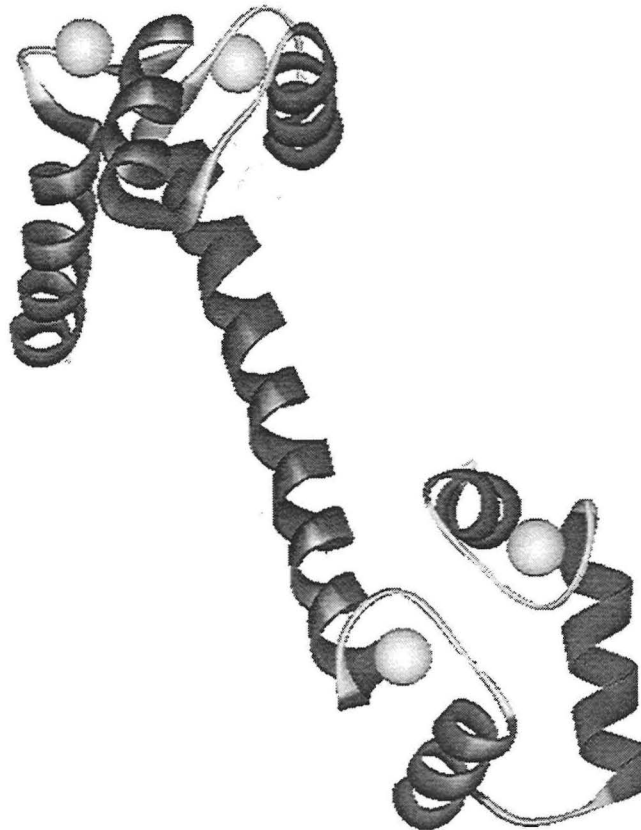


Figure I-13 : Structure de la forme recombinante de la Calmoduline Humaine (40) (Résolution 1.70Å).

Comme on peut le voir dans le Tableau I-9, de même que pour la troponine C, la calmoduline a deux types de sites : des sites à haute affinité peu spécifiques puisqu'ils fixent à la fois le Ca^{2+} et le Mg^{2+} et des sites de basse affinité hautement spécifiques, fixant uniquement le Ca^{2+} .

| sites \ K_a (Mol) | $K_{a\text{Ca}^{2+}}$ | $K_{a\text{Mg}^{2+}}$ |
|------------------------------|-----------------------|-----------------------|
| spécifiques | $\approx 10^3$ | - |
| site de haute affinité Ca/Mg | $\approx 10^4$ | 10^2 |

Tableau I-9 : constantes d'affinité des sites de fixation de la calmoduline

Il est à noter que contrairement à la troponine ce sont les sites situés dans la partie N-terminale qui sont identifiés comme fixant à la fois le Ca^{2+} et le Mg^{2+} .

3. protéine sarcoplasmique

Les protéines sarcoplasmiques (SCP) font partie des rares calciprotéines dont le site de fixation a la particularité d'avoir un acide Aspartique en position relative 12. Comme nous l'avons évoqué, lors du passage des formes calcique à magnésique de la parvalbumine le changement de conformation de la chaîne latérale du Glu en position relative 12 semble être un événement majeur. Les protéines sarcoplasmiques n'ayant pas cet acide aminé dans leur site EF1, nous avons pensé que l'étude de l'échange $\text{Ca}^{2+}/\text{Mg}^{2+}$ dans ce dernier pourrait s'avérer intéressant.

a) Fonction

Sur le plan fonctionnel, les protéines sarcoplasmiques sont moins bien connues que les précédentes. Pour l'instant connues uniquement dans les muscles d'Invertébrés on leur prête un rôle de réservoir à Ca^{2+} et Mg^{2+} homologue à celui décrit pour la

parvalbumine. Sur le plan thermodynamique elles ont des affinités relatives et absolues pour Mg^{2+} et Ca^{2+} variant de manière significative en fonction de leur origine (42).

Néanmoins une étude cinétique réalisée sur 2scp(43) a mené aux conclusions suivantes : la protéine 2scp possède trois sites Ca^{2+}/Mg^{2+} mais pas de site spécifique du Ca^{2+} ; la fixation du Mg^{2+} a une coopérativité positive marquée, mais pas celle du Ca^{2+} . La fixation du premier Ca^{2+} ou Mg^{2+} induit un changement conformationnel qu'on peut suivre par la fluorescence du tryptophane.

Dans des réactions de déplacement les changements conformationnels surviennent lors de l'étape $SCP.Mg_3 \rightleftharpoons SCP.Ca_1Mg_2$. Le déplacement du Ca^{2+} par le Mg^{2+} peut être décrit en deux réactions consécutives, la première correspondant à la dissociation du dernier Ca^{2+} ($k_1 = 2.4 \text{ s}^{-1}$) et la seconde ($k_2 = 0.45 \text{ s}^{-1}$) au changement de conformation final observé durant la fixation directe du Mg^{2+} . Le remplacement de Mg^{2+} par Ca^{2+} suit un schéma cinétique de simple compétition; le taux de dissociation du Mg^{2+} augmente avec la concentration en Ca^{2+} . En résumé, après la dissociation rapide du Ca^{2+} ou Mg^{2+} , SCP Nereis change lentement vers la configuration libre de cation, mais dans les réactions d'échange les changements conformationnels sont presque aussi rapides que la réaction de départ du cation.

b) Structure

On trouve ces protéines sous forme de monomère ou de dimère⁵. Au cours de cette étude nous avons utilisé la structure 2scp(45) extraite d'un ver (*Nereis diversicolor*).

⁵ Il a été cristallisé une autre structure de protéine sarcoplasmique la 2sas(44) sous forme de monomère, mais qui est cependant moins précise que la précédente (résolution de 2.4 Å).

Cette protéine est cristallisée sous forme de dimère. La structures tertiaire de cette protéine est représentée dans la Figure I-14.

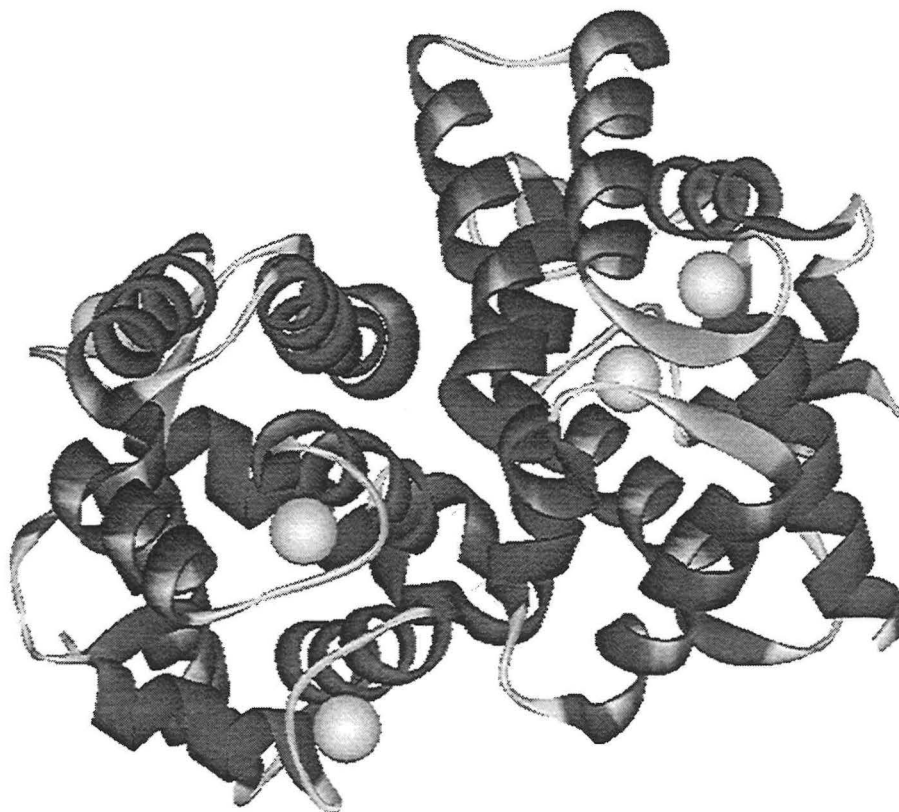


Figure I-14 : Structure du dimère: 2scp(45) Source: Nereis diversicolor, Résolution: 2.00 Å.

Il est à noter que la présence de Asp au lieu de Glu en position relative 12 du premier site a pour conséquence au niveau du repliement un rapprochement de l'hélice F1 par rapport à l'hélice E1, ce qui se traduit par une diminution de l'angle entre ces hélices(3) . En effet on constate une valeur de cet angle de 62° , qui est très inférieure à celle de 100° habituel. Ce phénomène est observé dans les deux structures connues de ce type de protéine bien que leur homologie de séquence soit faible (12 % (3)).



D. Références bibliographiques

1. Kretsinger, R. H., and Nockolds, C. E. (1973) *J.Biol.Chem* **248**(3313-3326)
2. Declercq, J. P., Tinant, B., Parello, J., and Rambaud, J. (1991) *J.Mol.Biol.* **220**(4), 1017-1039
3. Kawasaki, H., and Kretsinger, R. (1995) *protein profile* **2**(4), 333
4. Kretsinger, R. H. (1972) *Nature New biol.* **240**, 85-88
5. Falke, J. J., Drake, S. K., Hazard, A. L., and Peersen, O. B. (1994) *Q Rev Biophys* **27**, 219-290
6. Henrotte, J. G. (1952) *Nature* **169**, 968-969
7. Kretsinger, R. H., Nockolds, C. E., Coffee, J. C., and Bradshaw, R. A. (1971) *Cold spring Harbor symp Quant biol* **36**, 217-220
8. Heizman, C. W., and Celio, M. R. (1987) *Method in Enzymology* **139**(-), 552-570
9. Pechere, J.-F. (1977) *The significance of parvalbumins among muscular calciproteins. In Calcium-binding Proteines and calcium* (wasserman, R. H., corradino, R., carafoli, E., Kresinger, R. H., Maclellnann, D. H., and Siegel, F. L., Eds.), Elsevier, New York
10. Pechere, J.-F. (1968) *Comp. biochem. physiol* **24**, 910-921
11. Lehky, P., Blum, H. E., Stein, E. A., and Fischer, E. H. (1974) *J.Biol.Chem.* **24**, 4332-4334
12. Cox, J. A., and Stein, E. A. (1981) *Biochemistry* **20**, 5430-5436
13. Berchtold, M. W., Celio, M. R., and Heizman, C. W. (1984) *J.Biol.Chem.* **262**, 7314-7320
14. Kagi, U., Berchtold, M. W., and Heizman, C. W. (1987) *J.Mol.Biol.* **262**(7314-7320)
15. MacManus, J. P., Watson, D. C., and Yaguchi, M. (1985) *Biochemistry journal* **229**, 39-45
16. Hamano, K., Kiyama, H., Emson, P. C., Manabe, R., Nakaubi, M., and Tohyama, M. (190) *J. comp. Neurol.* **302**, 417-424
17. Bindels, R. L., Hartog, A., Timmermans, J. A., and van Os, C. H. (1991) *Contrib. Nephrol.* **91**, 7-13
18. Endo, T., Takazawa, K., and Onaya, T. (1985) *Endocrinology* **117**, 527-531
19. Celio, M. R. (1986) *science* **231**, 995-996
20. Kosaka, T., Katsumuru, H., Hama, K., Wu, J.-Y., and Heizman, C. W. (1987) *Brain Res.* **419**, 119-130
21. Brewer, J. M., Wunderlich, J. K., Kim, d.-H., Carr, M. Y., Beach, G. G., and Ragland, W. L. (1989) *Biochem. biophys. res. Commun.* **160**, 1155-1161
22. Heizman, C. W., and Berchtold, M. W. (1987) *Cell Calcium* **8**, 1-41

23. van Brederode, J. F., Helliesen, M. K., and Hendrickson, A. E. (1991) *Neuroscience* **44**, 157-171
24. Ohshima, T., Endo, T., and Onaya, T. (1991) *J. Neurol.* **238**, 320-322
25. Ruegg, J. C. (1989) *Calcium in muscle activation* (Verlag, S., Ed.), Berlin
26. Gerday, C., and Gillis, J. M. (1976) *J Physiol (Lond)* **258**(2), 96P-97P
27. Gillis, J. M., Thomason, D., Lefevre, J., and Kretsinger, R. H. (1982) *J Muscle Res Cell Motil* **3**, 377-398
28. Robertson, S. P., Johnson, J. D., and Potter, J. D. (1981) *Biophys J* **34**(3), 559-69
29. Heizman, C. W. (1984) *Experientia* **40**, 910-921
30. Moncrief, N. D., Kretsinger, R. H., and Goodman, M. (1990) *J.Mol.Evol.* **30**, 522-562
31. Cavé, A., Parello, J., Drakenberg, T., Thulin, E., and Lindman, B. (1982) *Mol. Cell. Biolchem. J.* **227**, 711-717
32. Gillen, M. F., Banville, D., Rutledge, R. G., Narang, S., Seligy, V. L., Withfield, J. F., and MacManus, J. P. (1987) *J.Biol.Chem.* **262**, 5308-5312
33. Mutus, B., Karuppiah, N., Sharma, R. K., and MacManus, J. P. (1985) *Biochem. Biophys. Res. Commun.* **131**, 500-506
34. Cavé, A., and Parello, J. (1981) (*Membranes et Communication intercellulaire* (Balian, r. e. A. e., ed), pp. 197-227, North Holland publishing Company, Amsterdam., Les Houches
35. Blancuzzi, Y., Padilla, A., Parello, J., and Cave, A. (1993) *Biochemistry* **32**(5), 1302-1309
36. Ebashi, S. (1970) *Nippon Ishikai Zasshi* **64**(6), 698-704
37. Greaser, M. L., and Gergely, J. (1973) *J Biol Chem* **248**(6), 2125-33
38. Satyshur, K. A., Rao, S. T., Pyzalska, D., Drendel, W., Greaser, M., and Sundaralingam, M. (1988) *J.Biol.Chem* **263**, 1628-1647
39. Cheung, W. Y. (1980) *Science* **207**, 19-27
40. Chattopadhyaya, R., Meador, W. E., Means, A. R., and Quioco, F. A. (1992) *J.Mol.Biol.* **228**, 1177-1192
41. Strynadka, N. C., and James, M. N. (1989) *Annu. Rev. Biochem.* **58**, 951-998
42. Wnuk, W., Cox, J. A., and Stein, E. A. (1981) *J Biol Chem* **256**, 11538-11544
43. Engelborghs, Y., Mertens, K., Willaert, K., Luan-Rilliet, Y., and Cox, J. A. (1990) *J Biol Chem* **265**, 18809-18815
44. Cook, W. J., Jeffrey, L. C., Cox, J. A., and Vijay-Kumar, S. (1993) *J Mol Biol* **229**(2), 461-71
45. Vijay-Kumar, S., and Cook, W. J. (1992) *J.Mol.Biol* **224**, 413

PARTIE II

OUTILS et MÉTHODES



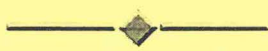


Table des matières :

| | |
|--|-----------|
| <i>I. Introduction</i> | 57 |
| <i>II. De la notion de surface d'énergie potentielle à la dynamique moléculaire</i> | 58 |
| A. Notion de surface d'énergie potentielle | 58 |
| B. Champ de Forces | 61 |
| 1. Expression de l'énergie : | 62 |
| 2. Troncature de la liste des atomes non liés: | 66 |
| 3. Limitations du champ de forces | 72 |
| C. Dynamique Moléculaire : | 75 |
| 1. Algorithme d'intégration | 76 |
| 2. Choix du pas d'intégration | 79 |
| 3. Notion de température : | 80 |
| 4. Phases d'une simulation : | 81 |
| 5. Ensemble statistique de simulation | 83 |
| D. Conditions de solvation | 85 |
| 1. Conditions périodiques | 85 |
| 2. Méthode du « mur étendu » (Extended Wall region Method) | 87 |
| E. outils d'analyses | 88 |
| 1. RMS (Root-Mean-Square) | 88 |
| 2. Courbes de distribution radiale | 89 |
| 3. Analyses géométriques | 91 |
| <i>III. De la trajectoire de dynamique moléculaire aux calculs d'énergie libre</i> | 92 |
| A. introduction | 92 |

| | |
|--|------------|
| B. De la fonction de partition à l'expression de l'énergie libre | 95 |
| C. Calcul d'énergie libre par simulation de dynamique « classique » | 104 |
| 1. Intégration thermodynamique | 105 |
| 2. perturbation thermodynamique | 107 |
| 3. Paramètre λ et type de calcul | 109 |
| IV. Références Bibliographiques | 118 |



I. Introduction

Nous avons présenté dans le chapitre précédant les calciprotéines à mains EF et la problématique qui nous intéresse, à savoir de manière générale l'étude du comportement du motif main EF lors de la fixation d'ions calcium et magnésium.

Ce chapitre est dédié à la description des concepts et outils sur lesquels sont basés les simulations effectuées lors de ce travail. Après avoir décrit les notions nécessaires à la mise en place des simulations de dynamique moléculaire : notion de surface d'énergie potentielle et champ de force, nous passerons à la technique de dynamique moléculaire proprement dite, pour enfin, à partir des bases de la thermodynamique statistique, expliciter les méthodes de calcul de différences d'énergie libre. Nous nous attarderons plus particulièrement sur la méthode que nous avons utilisée : la méthode de perturbation thermodynamique

II. De la notion de surface d'énergie potentielle à la dynamique moléculaire

A. Notion de surface d'énergie potentielle

Cette discussion ne peut débiter sans la description de l'équation de Schrödinger (Equation II-1), équation à l'origine de la description quantique des systèmes moléculaires, puisqu' idéalement la résolution exacte de cette équation conduirait à l'obtention de toutes les informations relatives au système étudié.

$$H.\Psi(R,r) = E.\Psi(R,r) \quad \text{Equation II-1}$$

Où H est l'hamiltonien du système, Ψ est la fonction d'onde, fonction des coordonnées des noyaux (R) et des électrons (r), et E, l'énergie totale du système.

Malheureusement, comme Dirac l'avait énoncé, il y a près de 50 ans: « *Les lois physiques fondamentales pour une théorie mathématique d'une grande partie de la physique et de toute la chimie sont maintenant complètement connues, et la seule difficulté est que l'application exacte de ces lois conduit à des équations beaucoup trop complexes pour être solubles.* ». En effet, bien qu'elle soit très générale l'Equation II-1 est trop complexe à résoudre pour un système moléculaire, même simple.

Cette complexité est due à l'expression de l'Hamiltonien associé à un système isolé. En effet pour un système formé de N noyaux (de masse m_μ , de charge Z_μ et positions \bar{P}_μ) et n électrons (repérés par les coordonnées $\bar{\rho}_i$), cet hamiltonien s'écrit :

$$H = \sum_{\mu=1}^N -\frac{\hbar^2 \nabla_\mu^2}{2m_\mu} + \sum_{i=1}^n -\frac{\hbar^2 \nabla_i^2}{2m_e} + \sum_{\mu < \nu = 1}^N \frac{Z_\mu Z_\nu e^2}{R_{\mu\nu}} - \sum_{i=1}^n \sum_{\mu=1}^N \frac{Z_\mu e^2}{r_{i\mu}} + \sum_{i < j = 1}^n \frac{e^2}{r_{ij}} \quad \text{Equation II-2}$$

avec $R_{\mu\nu} = \|\bar{P}_\nu - \bar{P}_\mu\|$, $r_{i\mu} = \|\bar{\rho}_i - \bar{P}_\mu\|$ et $r_{ij} = \|\bar{\rho}_j - \bar{\rho}_i\|$

Les deux premiers termes représentent respectivement les opérateurs associés aux énergies cinétiques des noyaux et des électrons, et les trois termes suivants, les opérateurs associés aux interactions électrostatiques inter-nucléaires, noyaux-électrons et inter-électroniques . Il est heureusement possible de diminuer un peu la complexité du problème, au moyen d'approximations.

Une première étape vers la simplification du problème requiert une approximation pour séparer le mouvement des noyaux de celui des électrons. Sa justification physique est fondée sur la différence de masse (donc de vitesse) entre noyaux et électrons. L'approximation de Born-Oppenheimer (1) suppose donc que, pour chaque position des noyaux \vec{P}_μ , le mouvement des électrons est suffisamment rapide pour que la solution obtenue en supposant les noyaux fixes soit proche de la solution exacte.

$$H_{elec} = \sum_{i=1}^n -\frac{\hbar^2 \nabla_i^2}{2m_e} - \sum_{i=1}^n \sum_{\mu=1}^N \frac{Z_\mu e^2}{r_{i\mu}} + \sum_{i<j=1}^n \frac{e^2}{r_{ij}} \quad \text{Equation II-3}$$

Elle introduit l'hamiltonien électronique H_{elec} ci-dessus, qui correspond aux mouvements des électrons dans le champ des noyaux fixes et dont la solution ne dépend que des coordonnées des électrons (les coordonnées des noyaux n'interviennent que paramétriquement). L'énergie totale du système, à une géométrie donnée, peut alors être calculée comme la somme de l'énergie électronique et d'une constante de répulsion nucléaire :

$$E_{tot}(R_a) = E_{elec}(R_a) + \sum_{A=1}^N \sum_{B>A}^N \frac{Z_A Z_B}{R_{AB}}$$

Cela correspond à la notion de *surface d'énergie potentielle* : il s'agit de la fonction $E_{tot}(R_m)$, qui associe l'énergie électronique ($E_{elec}(R_a)$) à chaque position des noyaux . En fait, cette énergie électronique joue le rôle d'énergie potentielle pour les mouvements nucléaires.

L'exploration de cette surface donne des informations fondamentales sur la structure moléculaire, et peut, via la thermodynamique statistique, permettre le calcul de grandeurs thermodynamiques telles que l'énergie libre ou l'entropie.

Cependant, malgré les approximations simplificatrices l'effort nécessaire à la détermination de cette surface est encore extrêmement important et compte tenu de la puissance des ordinateurs actuels ne peut être envisagée, dans le meilleur des cas, que pour des systèmes composés au plus de quelques dizaines d'atomes. De ce fait, l'étude de systèmes de la taille des macromolécules biologiques est inaccessible par ces approches.

C'est pourquoi nous sommes contraints d'utiliser, non pas la surface d'énergie potentielle exacte, impossible à obtenir, mais une surface d'énergie potentielle empirique, extrapolée à partir de données expérimentales : distances d'équilibre de liaisons issues de la diffraction des rayons X, constantes de force issues de la spectroscopie infra-rouge ou bien de calculs *ab initio* effectués sur de petites molécules organiques, ..etc. Cette approche repose sur l'approximation que des données empiriques peuvent implicitement incorporer les effets quantiques.

Les macromolécules biologiques (hétéropolymères) sont formées de fragments de molécules plus petites, qui sont connectées les unes aux autres par des liaisons covalentes. Par exemple, les protéines sont des assemblages d'acides aminés qui sont connectés via des liaisons peptidiques. Une protéine repliée peut paraître être un objet extrêmement compliqué. Pourtant, les fragments de protéines et des autres macromolécules biologiques sont assemblés à partir des mêmes types de liaisons que les petites molécules organiques. C'est ce fait qui est exploité dans la construction de la surface d'énergie potentielle de ces macromolécules .



Ces modélisations de surfaces, souvent appelées « champs de force », sont à l'origine de la mécanique et de la dynamique moléculaire. La mécanique moléculaire se préoccupe de la recherche de géométries particulières et de l'énergie associée aux points particuliers de la surface d'énergie potentielle empirique tels que : ses minima, maxima et point-selles. La dynamique moléculaire appliquée par M. Karplus et son école (M. Karplus et J. A. McCammon, 1983), de concert avec cette approche de modélisation de surfaces d'énergie potentielle de macromolécules biologiques, permet une exploration temporelle de ces surfaces sur des échelles de temps allant de la picoseconde à la nanoseconde. Cette dernière technique est la « clé de voûte » de notre travail.

B. Champ de Forces

Le champ de forces regroupe l'ensemble des expressions et des paramètres nécessaires à la description de la surface d'énergie potentielle du système moléculaire. Les champs de forces les plus communs, dont fait partie celui du programme CHARMM (2), que nous avons utilisé tout au long de ce travail, peuvent être subdivisés en deux classes de termes.

→ les termes permettant de décrire la structure covalente du système étudié, tels que les liaisons covalentes, les angles de valence, les angles dièdres.

→ les termes qui modélisent les interactions entre atomes non liés.

L'objectif d'un champ de forces est de décrire au mieux des molécules les plus diverses possibles, avec le minimum de paramètres et une précision raisonnable (cf.

Limitations du champ de forces)

1. Expression de l'énergie :

En un point de la surface de potentiel du système, l'énergie est donnée par :

$$E_{potentielle} = E_{covalente} + E_{non_liés}$$

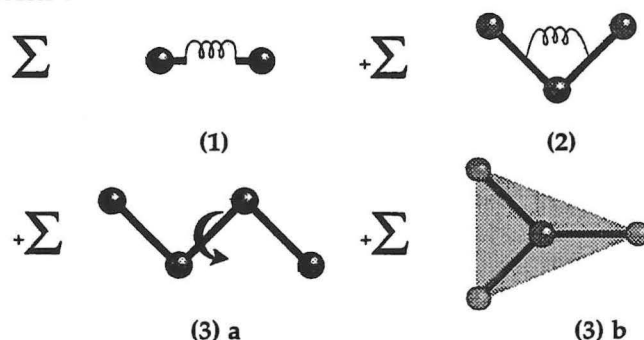
et chaque classe de termes se décompose à son tour ainsi :

$$E_{covalente} = E_{liaisons} + E_{angles} + E_{diédres} + E_{impropres} \quad \text{Equation II-4}$$

$$E_{non_liés} = E_{vdw} + E_{elec}$$

Dans CHARMM(2) , les deux classes de termes énergétiques du champ de forces ont la forme suivante.

→ Termes covalents :



Ces termes sont généralement fonction de coordonnées internes :

1) liaisons :

$$E_{liaisons} = \sum_{liaisons} K_{liaisons} \cdot (b - b_0)^2 \quad \text{Equation II-5}$$

Chaque liaison, tel un ressort, est décrite par un potentiel harmonique avec b_0 , la longueur à l'équilibre de la liaison de longueur courante b , et enfin $K_{liaison}$ la constante de force de la liaison.

2) Angles de valence :

$$E_{\text{angle}} = \sum_{\text{angle}} K_{\text{angle}} \cdot (\theta - \theta_0)^2 \quad \text{Equation II-6}$$

de même que pour les liaisons, ils sont décrits par un potentiel harmonique où les paramètres K_{angle} et θ_0 sont analogues à K_{liaison} et b_0 .

3) Angles dièdres :

Les angles dièdres sont de deux types : angle de torsion et torsion impropre.

(a) *angles de torsions*

Ces termes contrôlent les rotations autour des liaisons covalentes.

$$E_{\text{dièdres}} = \sum_{\text{dièdres}} K_{\text{dièdre}} \cdot (1 + \cos(n\phi - \delta)) \quad \text{Equation II-7}$$

Les termes énergétiques associés aux angles de torsions du système sont représentés par des fonctions périodiques, dans lesquelles $K_{\text{dièdre}}$, n et δ sont respectivement : la constante de force, le paramètre de multiplicité et la phase de l'angle dièdre.

(b) *torsions impropres :*

$$E_{\text{impropre}} = \sum_{\text{impropre}} K_{\text{impropre}} \cdot (\omega - \omega_0)^2 \quad \text{Equation II-8}$$

Ces termes ont pour fonction de maintenir la géométrie de certains groupes d'atomes (par exemple : la planéité des cycles aromatiques). Ces termes sont aussi utilisés afin de maintenir la configuration d'atomes asymétriques. Ici encore les termes de torsion impropre sont décrits par un potentiel harmonique, dans lequel K_{impropre} est la constante de force de torsion dite impropre et ω_0 est l'angle entre le plan contenant un

angle de valence (p1) et le plan (p2) contenant trois atomes coordonnés à l'atome central (cf. Figure II-1)

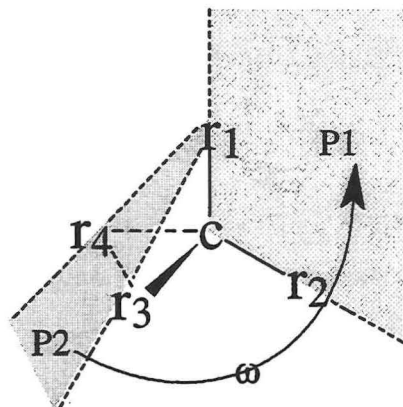
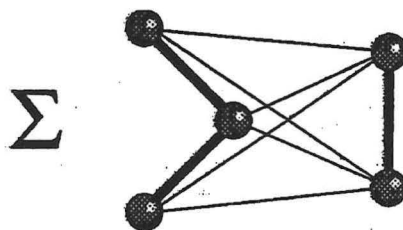


Figure II-1 : représentation d'un angle impropre ω . Les plans P1 et P2 sont respectivement formés par les atomes r1, c, r2 et r1, r3, r4.

→ Termes Non Liés :



Ces termes sont fonction des distances entre atomes non liés. Ils se décomposent en termes électrostatiques et de Lennard Jones.

* les interactions électrostatiques:

$$E_{elec} = \sum_{\text{paires-d'atomes-non-liés}} \left(\frac{q_i \cdot q_j}{\epsilon \cdot r_{ij}} \right) \quad \text{Equation II-9}$$

Ces interactions sont décrites, avec un potentiel de type Coulomb : q_i et q_j et ϵ sont respectivement les charges partielles des atomes impliqués dans l'interaction et la



constante diélectrique du milieu. Dans notre cas comme nous traiterons le milieu, c'est-à-dire le solvant, avec des molécules d'eau explicites, ϵ sera pris égal à 1.

* les interactions types van der Waals

Dans le programme CHARMM(2), ces interactions entre paires d'atomes non liés sont traitées via un potentiel de type Lennard-Jones :

$$E_{vdw} = \sum_{\text{paires-atomes-non-liés}} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) \quad \text{Equation II-10}$$

Dans cette équation r_{ij} est la distance séparant les atomes i et j . Le terme en $1/r_{ij}^{12}$ n'a pas de signification physique mais est là pour rendre compte de la répulsion à courte distance entre nuages électroniques. Le terme en $1/r_{ij}^6$ est attractif et regroupe les contributions des forces de dispersion, à moyenne et longue distance.

A_{ij} et B_{ij} sont les paramètres Lennard-Jones caractéristiques d'une paire d'atomes ij . Dans le programme CHARMM(2), ces paramètres sont calculés à partir des $E_{\min}(i)$ et $R_{\min}(i)$ caractéristiques de chaque atome, conformément aux lois de Berthelot Lorentz :

$$A_{ij} = \sqrt{E_{\min}(i) \times E_{\min}(j)} \times \left[\frac{R_{\min}(i) + R_{\min}(j)}{2} \right]^{12} \quad \text{Equation II-11}$$

$$B_{ij} = 2\sqrt{E_{\min}(i) \times E_{\min}(j)} \times \left[\frac{R_{\min}(i) + R_{\min}(j)}{2} \right]^6 \quad \text{Equation II-12}$$

$E_{\min}(i)$ et $R_{\min}(i)$ sont respectivement : la profondeur du puits de potentiel de l'interaction entre deux atomes de type (i) et la distance inter-atomique à laquelle se situe le minimum de ce puits.(cf. Figure II-2)

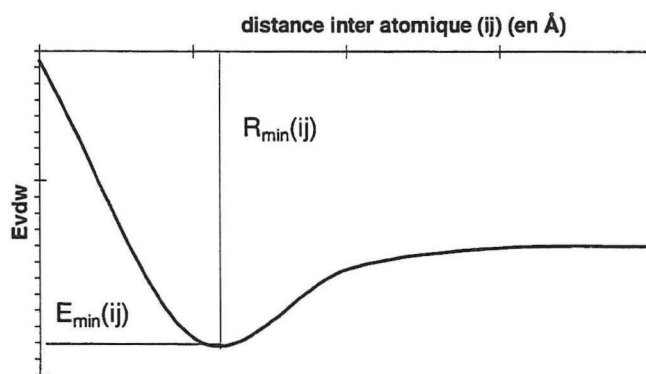


Figure II-2 : puits d'énergie van der Waals

Le champ de forces que nous utilisons n'inclut pas explicitement de termes énergétiques spécifiques des liaisons hydrogènes. Celles-ci sont traitées implicitement via la paramétrisation du potentiel Lennard-Jones 6-12 et du potentiel de Coulomb pour les atomes répertoriés comme étant susceptibles de participer à une liaison hydrogène.

2. Troncature de la liste des atomes non liés:

Compte tenu de la puissance des ordinateurs actuels, on peut envisager de traiter raisonnablement des systèmes composés de quelques milliers d'atomes, sur une station de travail ordinaire, jusqu'à quelques dizaines de milliers d'atomes, pour les « super ordinateurs » type Cray. Par exemple, un système composé d'une petite protéine de 100 à 150 résidus - comme la parvalbumine - solvatée dans deux à trois couches de molécules d'eau, représente déjà près de 5000 atomes .

La plus grande partie des ressources informatiques utilisée est monopolisée par le calcul des termes de type non liés. En effet, le temps nécessaire au calcul des termes covalents d'un système est proportionnel au nombre de coordonnées internes, et ce dernier croît linéairement avec le nombre d'atomes. Par contre, les sommes relatives aux interactions non liées (cf. Equation II-9 et Equation II-10), croissent, pour un système de N atomes, en $(N.(N-1)/2)$ soit quadratiquement lorsque N est grand. Pour un

a) Fonctions de lissage

Mais la simple coupure des interactions non-liées à une distance donnée a une contrepartie. Cela induit en effet des discontinuités dans le calcul de l'énergie et de sa dérivée. Ce qui peut provoquer des artefacts lors des minimisations et des non conservations de l'énergie totale lors de simulations de dynamique moléculaire. C'est pourquoi les programmes de dynamique moléculaire, dont CHARMM, utilisent souvent des fonctions de lissage pour modifier le profil des courbes d'interaction et ainsi atténuer les effets du cutoff.

(1) fonction de switch

Cette fonction est illustrée ci-dessous (Figure II-4). Typiquement, on l'applique aux interactions de type van der Waals. Elle est définie par deux distances : d_{cut} , pour laquelle la fonction devient nulle (le cutoff) et d_{switch} , distance pour laquelle la fonction « switch », $SW(r)$, commence à être appliquée.

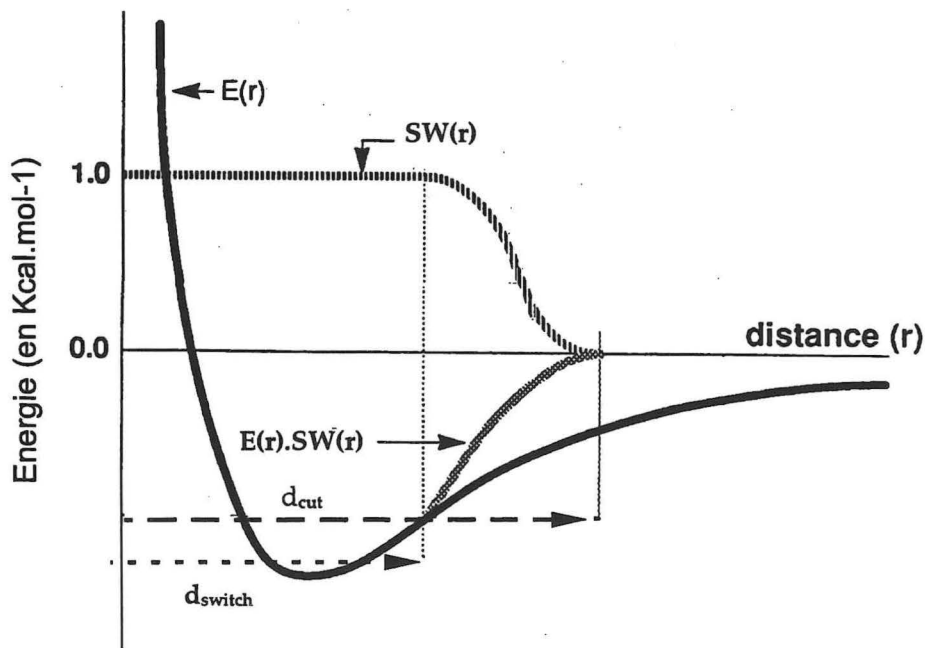


Figure II-4. Application de la fonction de switch. Courbe de potentiel van der Waals non modifiée $E(r)$ (courbe noire) ; la fonction de switch $S(r)$ (courbe striée) et la courbe de potentiel « switchée » $E(r).SW(r)$.

système de 1000 atomes, la liste des interactions non liées comprendrait près de 500 000 termes et près de 50 millions pour un système de 10.000 atomes.

Pour diminuer le coût en calcul de ces sommes, une distance de troncature, dite rayon de coupure ou de «cutoff », est introduite à partir de laquelle les interactions entre atomes du système ne sont plus prises en compte. Sur le plan physique, cette approximation peut se justifier par le fait que, si le rayon de cutoff est suffisamment grand, les interactions avec l'ensemble des atomes situés à grandes distances d'un atome donné ont un effet total quasi nul sur l'atome considéré (si cet atome fait partie d'un groupe d'atomes dont la charge totale est nulle).

Du point de vue temps de calcul le résultat est radical, puisque comme nous pouvons le voir sur la Figure II-3, le nombre d'interactions et donc le temps nécessaire à leur calcul est plus de dix fois moins important lorsque l'on introduit un cutoff à 8 Å, qu'en l'absence de toute distance de troncature.

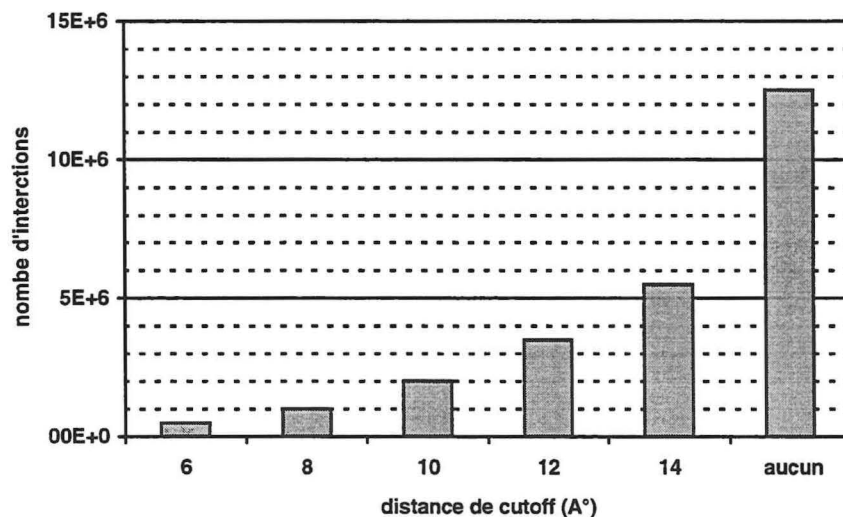
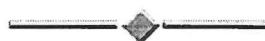


Figure II-3 : nombre d'interactions entre atomes non liés en fonction de la distance de cutoff (Å) pour un système de 5000 atomes. Le temps nécessaire pour évaluer l'énergie du système est à peu près proportionnel à ce nombre d'interactions



type «shift» qui va modifier le potentiel électrostatique pour toutes les valeurs de la distance (cf. Figure II-6), de telle sorte que l'énergie et la force associées deviennent progressivement nulles pour une distance inter-atomique égale au cutoff.

La fonction de shift que nous avons utilisée est donnée par l'expression suivante :

$$S(r) = \left(1 - \frac{r^2}{\text{cutoff}^2}\right)^2$$

Pour $r < \text{cutoff}$ nous avons : $E'(r) = E(r) * S(r)$. Pour $r > \text{cutoff}$, $E'(r)=0$

Lorsque r tend vers le cutoff, $S(r)$ tends vers 0, et $E'(r)$ devient nulle .

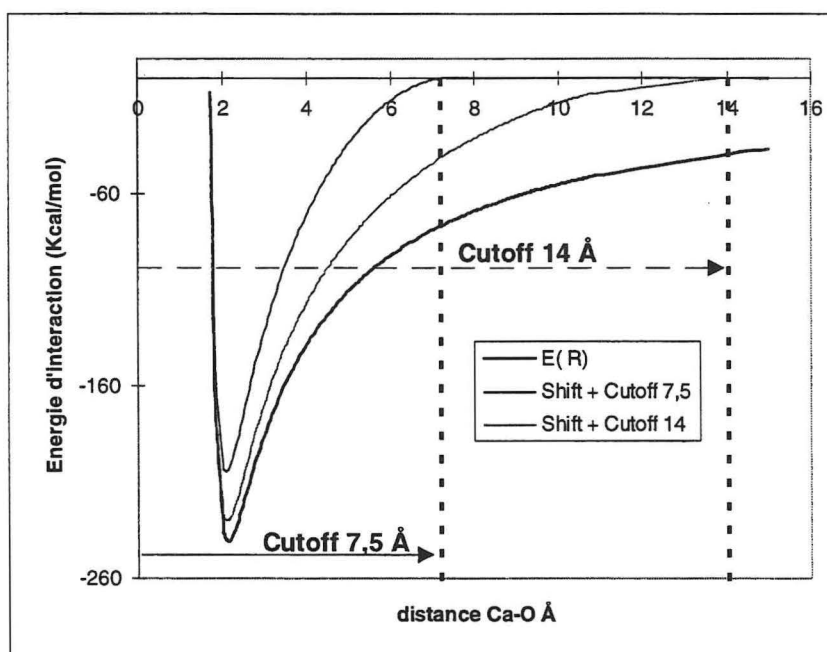


Figure II-6 : application d'une fonction de shift (S (R)) avec cutoff 7.5 et 14 Å, à une énergie d'interaction totale E(R) (van der Waals + électrostatique) entre un Calcium et l'oxygène d'une molécule d'eau .

Au cours de notre étude nous avons appliqué alternativement ce type de fonction Shift avec un cutoff de 7.5 et 14 Å. L'application de cette fonction de lissage aux interactions van der Waals est peu courante. Cependant, comme on peut le voir dans la figure ci-dessous, lorsque le cutoff a une valeur importante dans notre cas, cela n'induit que de faibles modifications .



Pour $r < d_{switch}$ $SW(r) = 1$, pour $r > d_{cut}$ $SW(r) = 0$ et enfin pour $r \in [d_{switch}; d_{cut}]$ $SW(r)$

est donnée par l'expression suivante :

$$SW(r) = \frac{(d_{cut} - r)^2 * (d_{cut} + 2r - 3d_{switch.})}{(d_{cut} - d_{switch})^3}$$

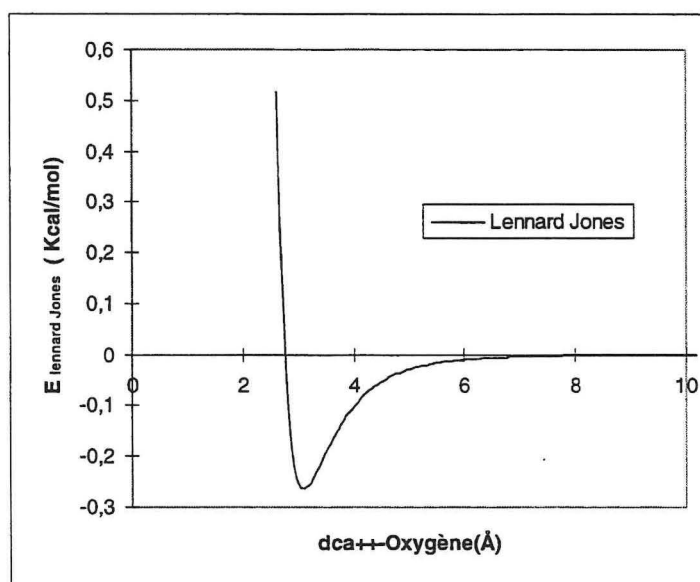


Figure II-5 : interaction Lennard-Jones Ca²⁺ oxygène. On peut constater que pour des distances inter atomiques supérieures à 7 Å l'interaction est quasi nulle

Cette fonction est à priori plus adaptée au traitement des interactions Lennard-Jones qu'à celui des interactions électrostatiques. En effet, le potentiel Lennard-Jones a une portée relativement courte (il décroît en $1/r^6$), et c'est pourquoi appliquer cette fonction pour des intervalles de distances de 6 à 10 Å peut paraître raisonnable.

Contrairement au potentiel Lennard-Jones, le potentiel Coulombien utilisé pour décrire les termes électrostatiques a un champ d'action nettement plus important dans la mesure ou il décroît en $1/r$. Par exemple, un cation divalent tel que Ca²⁺ ou Mg²⁺ et l'oxygène d'une molécule d'eau séparés par une distance de 10 Å ont encore une énergie d'interaction électrostatique d'environ -55 Kcal/mol. On comprend donc aisément qu'une fonction de switch soit peu appropriée, c'est pourquoi on utilise une fonction de

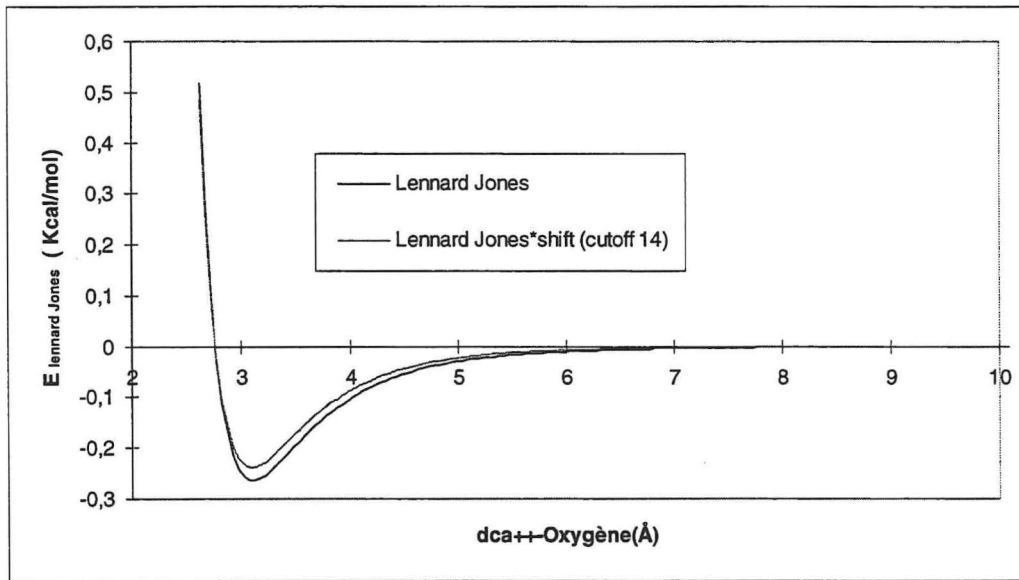


Figure II-7 : interaction Lennard-Jones entre un calcium et un oxygène, avec et sans fonction shift

3. Limitations du champ de forces

Compte tenu de la taille des systèmes étudiés et de la diversité des atomes qui les constituent, l'expression du champ de forces est un compromis visant à décrire au mieux la surface d'énergie potentielle du système de manière, d'une part à tenir compte de la limitation des ressources informatiques et, d'autre part, à limiter le nombre de paramètres nécessaires pour décrire les divers types d'atomes dans les multiples contextes chimiques dans lesquels ils apparaissent dans les systèmes biologiques.

C'est dans un tel but de simplification que, dans les simulations de dynamique moléculaire, les atomes sont décrits comme des sphères chargées reliées par des ressorts interagissant entre elles suivant une fonction énergie potentielle de forme relativement simple. Ces fonctions ont pour vertu d'être individuellement peu gourmandes en temps de calcul et en paramètres nécessaires à leur expression. Toutefois, ces simplifications ne sont pas sans contreparties.

- Structure covalente

La description harmonique des liaisons est une bonne description au voisinage de la distance d'équilibre. Néanmoins, cette forme de description quadratique exclut toute possibilité de formation ou de rupture des liaisons. Elle n'est donc pas adaptée à l'étude de problèmes mettant en jeu une modification de la structure covalente du système tels que les actes catalytiques ou les transferts de protons. De tels phénomènes nécessitent une description plus détaillée de la surface d'énergie potentielle, dans laquelle au moins une partie des électrons du système est traitée de manière explicite (par une approche quantique). Initialement développée par Warshel (3) ainsi que Karplus & Gô, les approches hybrides entre mécanique quantique et mécanique moléculaire classique sont

actuellement en plein essor. Pour une présentation succincte de ces méthodes voir la référence (4). Si l'on considère les phénomènes qui nous intéressent, le traitement harmonique de la structure covalente est à priori peu préjudiciable. En effet la fixation de cations par les calcioprotéines ne fait intervenir que des interactions non covalentes.

• Interactions non liées

La présence dans les simulations de cations types calcium ou magnésium (ou plus généralement d'atomes chargés), n'est pas facile à traiter. En effet ces atomes, de par le champ électrostatique qu'ils génèrent, polarisent fortement leur environnement.

| Origine | intensité ¹ | portée | expression dans le champ de forces |
|---|--------------------------------|---------------------------|------------------------------------|
| <i>Coulombienne</i> | | | |
| • ion - ion ⊕ ⊖ | ≈300 k _b .T intense | $\frac{1}{r}$ très longue | $\frac{qq'}{\epsilon.r}$ |
| <i>Polarisation</i> | | | |
| • ion-molécule non polaire ⊕ ○ | ≈k _b .T moyenne | $\frac{1}{r^4}$ Moyenne | ! |

Tableau II:1 y figurent en fonction de leur origine physique, respectivement, l'intensité et la portée des termes(5), ainsi que leur expression éventuelle dans le champ de forces de CHARMM, (!) indiquant que le terme n'y est pas explicitement décrit .

Comme on peut le voir dans le Tableau II:1, certains termes énergétiques induits par la présence de cations ne sont pas inclus explicitement dans la description énergétique utilisée par CHARMM. Des termes d'intensité et de portée non négligeables, tels que les termes associés à la polarisation, du type ion - dipôle induit, sont censés être englobés implicitement dans les paramètres du champ de force, en l'occurrence la

¹ k_bt/molécule est environ égal à 0.6 kcal.mol⁻¹ pour T=300K .

charge et les $E_{\min}(i)$ et $R_{\min}(i)$. Le choix des paramètres adéquats incluant au mieux, de manière efficace, tous ces termes « négligés » n'est donc pas facile, mais semble néanmoins possible en ce qui concerne les cations en milieu aqueux (ce que les résultats de notre travail justifieront à posteriori). Lors du traitement des interactions cation-oxygène, les paramètres que nous avons utilisés ont été déterminés selon deux approches. Ils ont été extraits, soit à partir de calculs ab initio (6,7), soit à partir de données physico-chimiques via des calculs d'énergie libre de solvation (8).

Si cette opération a, comme nous le verrons par la suite, déjà été menée avec succès pour des cations en solution(8), elle semble à priori plus difficile pour les cations au sein de protéines. En effet, contrairement à l'eau pure², dans les sites de fixation les cations peuvent interagir avec des groupements de natures différentes (carbonyles et carboxylates) dont la polarisabilité électronique est à priori beaucoup plus élevée que celle de l'eau.

Quoiqu'il en soit, compte tenu de la quantité de données expérimentales disponibles, il sera possible de mesurer la validité des paramètres utilisés pour décrire l'interaction cation-oxygène. Par ailleurs il sera aussi intéressant de voir comment se comportent des cations au sein de protéines en partant d'une description la plus simple possible.

² L'eau est polarisée en grande partie par son orientation et ses déplacements dans le champ du cation, ce phénomène est pris en compte explicitement dans la dynamique, alors que pour les groupement carbonyles et carboxylates la polarisation est essentiellement due au mouvement des électrons le long des doubles liaisons.

successives, une trajectoire dite de dynamique moléculaire est générée. Dans la mesure où la physique classique ne peut fournir une solution exacte que pour des systèmes d'au plus deux particules indépendantes, la résolution de l'Equation II-13 nécessite, pour nos systèmes, l'emploi de méthodes numériques. Ces méthodes numériques, sont basées sur le fait que le déplacement d'un atome peut être exprimé, au même titre que n'importe quelle fonction, sous la forme d'une série de Taylor. Ainsi, si l'on note $r(t)$, la position d'un atome au temps t , sa position après un intervalle très court Δt peut être écrite sous la forme :

$$\vec{r}_i(t + \Delta t) = \vec{r}_i(t) + \frac{d \vec{r}_i(t)}{d t} \cdot \frac{\Delta t}{1!} + \frac{d^2 \vec{r}_i}{d t^2} \cdot \frac{\Delta t^2}{2!} + \dots + \frac{d^n \vec{r}_i}{d t^n} \cdot \frac{\Delta t^n}{n!} \quad \text{Equation II-14}$$

1. Algorithme d'intégration

La résolution numérique des équations du mouvement dépend, d'une part, de la connaissance à l'instant t de la position, $\vec{r}_i(t)$, la vitesse, $\frac{d \vec{r}_i}{d t}$ et l'accélération, $\frac{d^2 \vec{r}_i}{d t^2}$; et d'autre part d'approximations appropriées, à savoir la prise en compte, uniquement, des contributions des termes d'ordres les moins élevés du développement II-14. En pratique, les termes supérieurs à l'ordre trois seront négligés.

Dans CHARMM cette étape est réalisée selon l'algorithme d'intégration de Verlet (9). Ce dernier est issu de la démonstration suivante : selon l'Equation II-14 la position $r(t+\Delta t)$ d'un atome à l'instant $t+\Delta t$ est à l'ordre trois :

C. Dynamique Moléculaire :

Les méthodes de dynamique moléculaire vont permettre, lors de simulations, d'explorer la surface d'énergie potentielle du système. Cette méthode est la clé de voûte de notre étude, puisque d'une telle exploration seront extraites les informations structurales ou énergétiques qui nous intéressent.

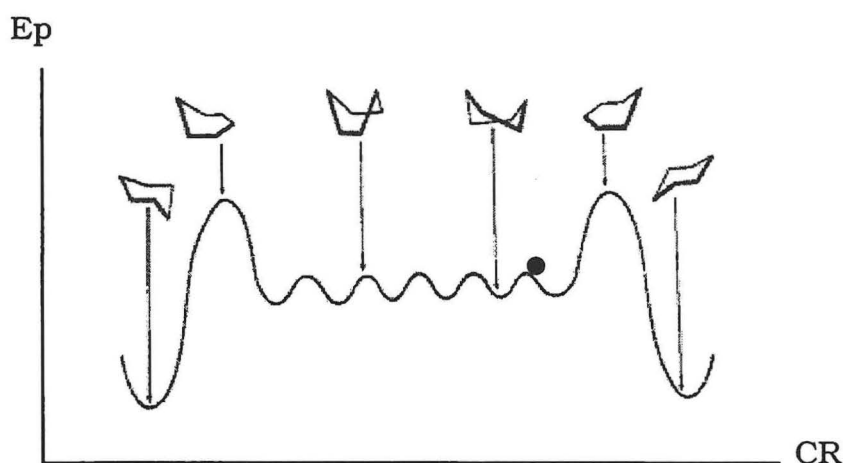


Figure II-8 illustration schématique de l'exploration d'une surface de potentiel du cyclohexane en coupe:

Cette méthode est basée sur la résolution, pour chaque atome du système, des équations de Newton: $\vec{F}_i = m_i \cdot \vec{a}_i$ où \vec{F}_i , m_i et \vec{a}_i sont respectivement la force subie, la masse et l'accélération d'un atome i . La force subie par un atome peut être calculée directement, à partir de la dérivée de l'énergie potentielle E par rapport aux coordonnées \vec{r}_i de l'atome i .

$$\vec{F}_i = - \frac{\partial E}{\partial \vec{r}_i} \quad \text{Equation II-13}$$

Ainsi, connaissant l'expression de l'énergie potentielle et les masses des atomes, il est possible d'intégrer l'équation de Newton pour chacun des atomes du système, et ainsi de déterminer leurs positions futures. De proche en proche, par itérations

$$\vec{r}_i(t + \Delta t) = \vec{r}_i(t) + \frac{d \vec{r}_i(t)}{d t} \cdot \frac{\Delta t}{1!} + \frac{d^2 \vec{r}_i}{d t^2} \cdot \frac{\Delta t^2}{2!} + \frac{d^3 \vec{r}_i}{d t^3} \cdot \frac{\Delta t^3}{3!}$$

une relation similaire peut être écrite pour sa position à $t - \Delta t$:

$$\vec{r}_i(t - \Delta t) = \vec{r}_i(t) - \frac{d \vec{r}_i(t)}{d t} \cdot \frac{\Delta t}{1!} + \frac{d^2 \vec{r}_i}{d t^2} \cdot \frac{\Delta t^2}{2!} - \frac{d^3 \vec{r}_i}{d t^3} \cdot \frac{\Delta t^3}{3!}$$

Si l'on somme maintenant les deux expressions précédentes, on obtient :

$$\vec{r}_i(t + \Delta t) + \vec{r}_i(t - \Delta t) = 2\vec{r}_i(t) + \frac{d^2 \vec{r}_i}{d t^2} \cdot \Delta t^2$$

ce que l'on peut écrire sous la forme :

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

Equation II-15

En résumé : un cycle élémentaire d'intégration comprend les deux étapes suivantes:

1. → une étape de calcul des accélérations courantes à partir des forces (Equation II-13)
2. → une étape de calcul des nouvelles coordonnées ($t + \Delta t$) à partir des coordonnées présentes (instant t) et antérieures $t - \Delta t$ selon l'Equation II-15

Typiquement, pour les systèmes biologiques, les coordonnées initiales sont issues de structures cristallographiques.

Comme on peut le voir ci-dessus les coordonnées à t et $t - \Delta t$ sont nécessaires pour déterminer les coordonnées à $t + \Delta t$. C'est pourquoi un cycle particulier est nécessaire pour initialiser l'algorithme . Pendant ce dernier, pour chaque atome du système, est calculé selon l'Equation II-16 un deuxième ensemble de coordonnées à $t + \Delta t$ à partir de sa vitesse et de ses coordonnées initiales \vec{v}_{0i} et \vec{r}_{0i} :

$$\vec{r}_{oi}(t + \Delta t) = \vec{r}_{oi}(t) + \vec{v}_{oi}\Delta t + \vec{a}_i(r) \cdot \frac{\Delta t^2}{2} \quad \text{Equation II-16}$$

Il est à noter que les vitesses atomiques n'apparaissent explicitement que dans ce cycle d'initialisation. Ces vitesses initiales sont assignées de manière aléatoire, en fonction d'une distribution de Maxwell-Boltzmann correspondant à la température désirée. Ce premier cycle réalisé, la trajectoire de dynamique peut commencer.

2. Choix du pas d'intégration

Les algorithmes standards de dynamique moléculaire, tels que celui décrit précédemment, sont basés sur l'intégration des équations du mouvement de Newton pour chacun des atomes du système. A partir de $\vec{r}_i(t)$, les coordonnées à l'instant t , et de la force $\vec{F}_i(t)$ subie par l'atome (i) au temps t , ils permettent de calculer la nouvelle position et la nouvelle force au temps suivant $t+\Delta t$ ou Δt est le pas d'intégration. Ce dernier doit être suffisamment petit, de façon à ce que les forces ne varient pas trop entre chaque itération du calcul, ceci afin de préserver la stabilité de l'algorithme. C'est pour cette raison que Δt doit être inférieur d'un ordre de grandeur à la période des mouvements de plus haute fréquence du système. Pour les molécules organiques, Δt doit donc être de l'ordre du dixième de la période de vibration des liaisons covalentes faisant intervenir l'hydrogène (≈ 10 fs).

Cependant, il est possible d'intégrer les équations du mouvement en présence de contraintes, qui vont permettre de « geler » les mouvements vibrationnels les plus rapides et, ainsi, d'utiliser un pas d'intégration plus grand, sans que le comportement dynamique du système soit notablement modifié(10). Nos calculs ont été réalisés avec un pas d'intégration de 2 fs, grâce à l'utilisation de l'algorithme SHAKE (11), qui permet d'appliquer ce type de contraintes sur la longueur des liaisons covalentes durant les trajectoires.

3. Notion de température :

La notion de température dans les simulations de dynamique moléculaire doit être définie. L'énergie cinétique interne totale du système est décrite par l'expression.

$$E_c = \sum_{i=1}^N \frac{1}{2} m_i \cdot v_i^2 = \frac{1}{2} \cdot (3N - 6) \cdot k_b \cdot T \quad \text{Equation II-17}$$

où N est le nombre d'atomes et k_b la constante de Boltzmann. Selon l'Equation II-17, la température est proportionnelle à l'énergie cinétique du système. Cependant une température donnée ne correspond pas à une vitesse atomique précise, mais à un ensemble de vitesses plus ou moins probables, que l'on peut exprimer sous la forme d'une distribution de vitesses atomiques, qui, à l'échelle microscopique, permet de relier température et vitesse atomique par la relation suivante:

$$F(v) \cdot dv = \left(\frac{m}{2\pi \cdot k_b T} \right)^{3/2} \cdot e^{-mv^2/2k_b T} \cdot 4\pi \cdot v^2 dv \quad \text{Equation II-18}$$

Cette équation donne la densité de probabilité, $F(v) \cdot dv$, d'avoir une molécule de masse m à une vitesse de module v, à la température T. C'est cette densité de probabilité, en fonction du module de la vitesse, qui constitue la distribution dite de Maxwell-Boltzmann. La Figure II-9 montre cette distribution pour un ensemble de molécules d'eau à différentes températures.

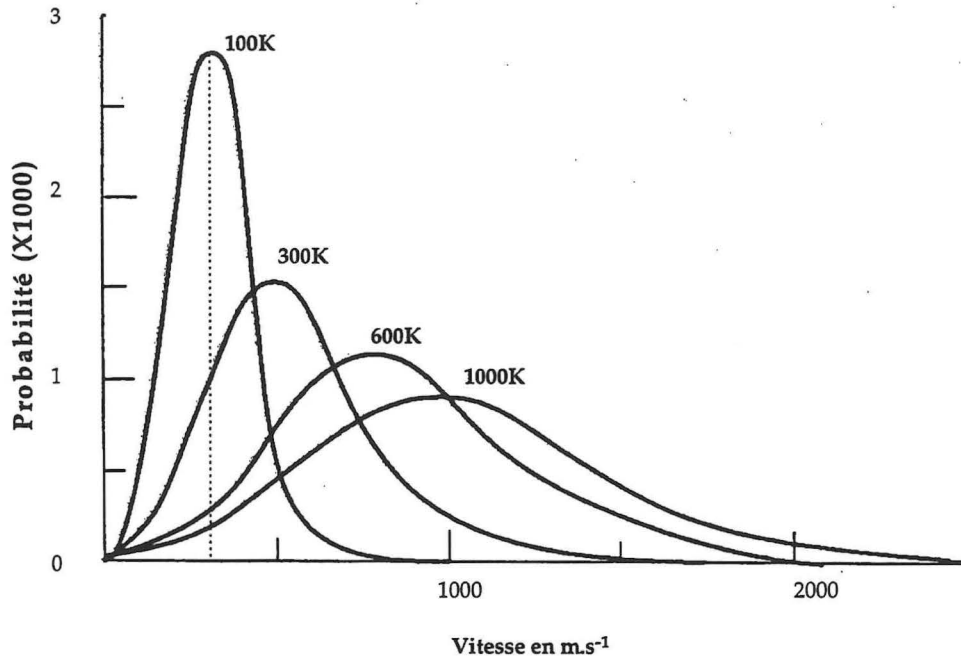


Figure II-9 : Distribution des modules des vitesses moléculaires à l'équilibre, pour des molécules d'eau à différentes températures, déterminée au moyen de l'Equation II-18 de Maxwell- Boltzmann.

Ainsi, lors du cycle d'initialisation d'une trajectoire de dynamique moléculaire, les vitesses des atomes sont assignées de manière aléatoire, de telle sorte que leur module respecte la distribution de Maxwell-Boltzmann associée à la température initiale désirée.

4. Phases d'une simulation :

a) étape de thermalisation

Les vitesses sont typiquement initialisées à une température de l'ordre de quelques Kelvins. Puis les équations du mouvement sont résolues selon l'Equation II-15, tout en chauffant peu à peu le système jusqu'à ce que la température désirée soit atteinte. Pour ce faire, les vitesses sont réassignées périodiquement, en les multipliant par un facteur : $\sqrt{\frac{T^\circ}{T(t)}}$ où T° est la nouvelle température désirée et $T(t)$ est la Température

courante calculée selon l'Equation II-17. Cette étape constitue l'étape de thermalisation. Lorsque la température escomptée est atteinte, la période d'équilibration du système peut alors débiter.

b) étape d'équilibration

Durant cette période, la simulation est poursuivie avec un contrôle strict de la température. En effet, après l'initialisation des vitesses selon une distribution de Maxwell-Boltzmann pour une température donnée, des fluctuations de température ont lieu. Ces fluctuations peuvent provenir d'une part des approximations utilisées lors de l'intégration numérique des équations du mouvement, d'autre part de la diminution de l'énergie potentielle totale dans le paysage conformationel exploré lors de la simulation.

Or, les lois de Newton et la forme de la fonction énergie utilisée sont telles que l'énergie totale du système doit rester constante et être équirépartie entre énergie cinétique et énergie potentielle. Ainsi la diminution de l'énergie potentielle a pour conséquence immédiate une augmentation de l'énergie cinétique et donc de la température. C'est pourquoi, durant la période d'équilibration, à intervalles fixes, la température moyenne est recalculée et comparée à la température désirée pour la simulation. Le cas échéant, c'est à dire en cas de fluctuation de la température en dehors d'une « fenêtre » donnée, les vitesses des atomes du système sont corrigées afin que la température calculée à partir de ces nouvelles vitesses demeure la température escomptée lors de la simulation.

Il faut noter, cependant, qu'après une période plus ou moins longue selon le système et les coordonnées initiales utilisées, le système évoluera spontanément vers une

zone de température stable dans laquelle les corrections des vitesses sont rares voire inutiles. On considérera alors que la période d'équilibration est achevée.

c) étape de fluctuation

L'étape d'équilibration étant effectuée, peut alors débiter la dynamique de fluctuation. Cette phase est aussi parfois appelée étape de production, car c'est durant cette période que sont collectées les coordonnées, les vitesses, à partir desquelles pourront être extraites des informations physico-chimiques.

5. Ensemble statistique de simulation

| <i>simulation</i> | <i>nom de l'ensemble</i> | <i>paramètres macroscopique constants</i> |
|-----------------------|--------------------------|---|
| libre | microcanonique | (N, V, E) |
| Température constante | canonique | (N, V, T) |

Tableau II:2

De manière naturelle, les simulations de dynamique moléculaire se font dans l'ensemble microcanonique. Cependant, des fluctuations ou des dérives faibles de l'énergie totale sont toujours observées. Ces perturbations sont dues à l'accumulation d'erreurs numériques relatives aux approximations effectuées, liées pour partie aux approximations relatives à l'intégration des équations de Newton (développement limité tronqué au 3^{ème} ordre, et taille finie du stockage des réels en informatique) et pour une part plus importante à la troncation des listes de paires d'atomes non liés (12).

La conservation de l'énergie totale est donc un critère important pour s'assurer de la validité de la trajectoire obtenue lors d'une simulation. Afin de tester cette dernière, le rapport entre les fluctuations de l'énergie totale et de l'énergie cinétique du système constitue une bonne mesure :



$$\sqrt{\frac{\langle \Delta E_{tot}^2 \rangle}{\langle \Delta E_{cin}^2 \rangle}}$$

relation 1

Compte tenu des erreurs numériques inhérentes aux simulations, on considère d'ordinaire que ce rapport doit être inférieur à 10 % (13).

Par ailleurs, en dehors des variations dues aux approximations numériques, il est à noter que dans un ensemble statistique, qu'il soit microcanonique ou canonique, la fluctuation d'une observable sera inversement proportionnelle à la racine du nombre de degrés de liberté du système. Il s'en suit que les fluctuations, selon l'ensemble statistique considéré, de la température dans le microcanonique, ou de l'énergie totale dans l'ensemble canonique, seront faibles pour des systèmes de grande taille. Dans ce cas, les équations de l'ensemble canonique pourront alors avantageusement remplacer celles de l'ensemble microcanonique pour décrire les propriétés d'un système isolé et ceci bien qu'elles ne s'appliquent en toute rigueur qu'à un système en équilibre avec un thermostat. (ref : *Statistical and thermal Physics by Reif, p 94-99*).

D. Conditions de solvation

Il est maintenant largement admis que la présence explicite du solvant est indispensable dans les simulations. Autrefois, pour des raisons pratiques, les simulations de dynamique moléculaire s'effectuaient dans le vide, le solvant étant traité de manière implicite par l'intermédiaire de la constante diélectrique ou bien sous forme de continuum.

1. Conditions périodiques

Les conditions périodiques sont un artifice de calcul permettant de simuler un système « infini » à partir d'un nombre limité de particules. Elles sont surtout utilisées lors de simulations avec un solvant explicite ou bien de simulations de systèmes cristallins. En général, le système à étudier (soluté + solvant) est contenu dans une boîte parallélépipédique (14) ou un octaèdre tronqué. Cette boîte est entourée de 26 copies ou « images » d'elle-même, obtenues par translations. Les équations du mouvement sont intégrées uniquement pour les molécules appartenant à la boîte centrale, toutes les positions de leurs copies dans les cellules voisines évoluant en fonction de celles-ci. Si l'une des molécules sort de la boîte centrale, elle est immédiatement remplacée par la molécule image qui pénètre du côté opposé.

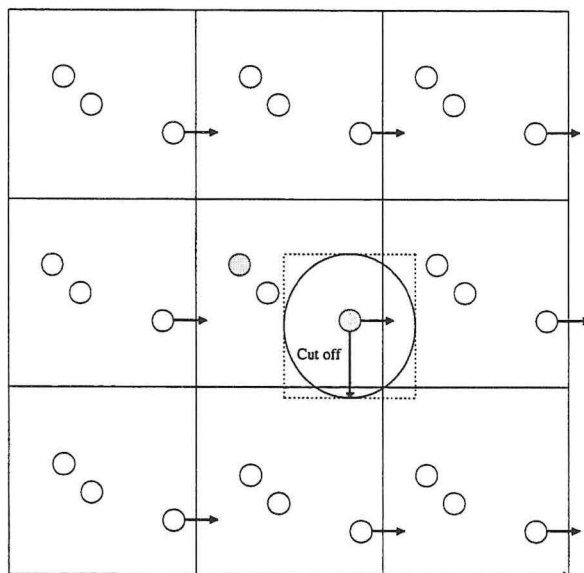


Figure II-10 : Représentation bidimensionnelle d'un système en conditions périodiques ; le cutoff correspond à la distance de troncature des interactions non liées.

Les forces agissant sur les atomes de la boîte centrale sont calculées à partir de la liste de tous les atomes remplissant le critère de cutoff. Cette liste prend alors en compte indifféremment les atomes de la boîte centrale et ses images. Afin de limiter les effets liés à la périodicité du système, le soluté de la boîte centrale ne doit pas interagir avec ses propres images et, de même, les molécules de solvant ne doivent pas interagir à la fois avec un atome du système et avec une image de celui-ci. Ainsi la taille de la boîte doit être suffisamment importante. Cette taille doit être adaptée, en fait, au cutoff utilisé durant la simulation.

C'est pourquoi, pour des solutés de grande taille, cette approche est difficile à mettre en œuvre car la forme parallélépipédique nécessite l'introduction d'un grand nombre de molécules d'eau. Il faut cependant noter que l'utilisation de conditions périodiques sur une boîte en forme d'octaèdre tronqué (15) ou d'oeuf (16) peut permettre de diminuer le nombre de molécules d'eau explicites de manière substantielle. Dans ce cas, les boîtes images précédemment décrites ont pour équivalent des zones concentriques.

2. Méthode du « mur étendu » (Extended Wall region Method)

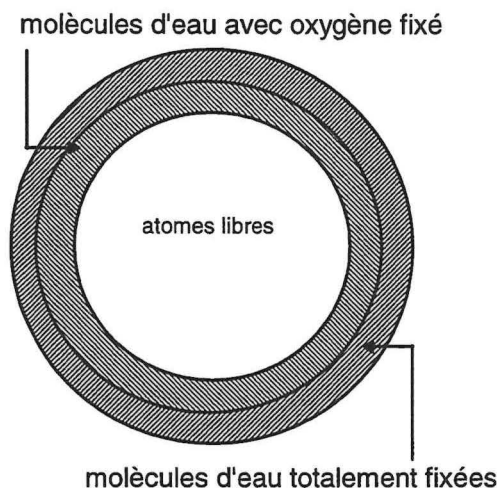


Figure II-11 sphère dont les frontières sont traitées selon la méthode du mur étendu

Dans ce travail nous avons utilisé une autre méthode de solvation. Cette technique présente l'avantage d'être l'une des plus économiques. Elle consiste à entourer le soluté d'une certaine épaisseur de molécules de solvant. Pour éviter que les molécules situées à la surface du système ne s'évaporent, la couche la plus périphérique est contrainte à rester immobile. De plus, afin d'éviter des effets de frontière trop marqués, une zone plus interne est partiellement maintenue fixe. Les atomes d'oxygène sont alors contraints.

Du fait de son moindre coût en temps de calcul, ce type d'approche est souvent utilisé (17), pour solvater tout ou partie³ de protéines, notamment lors de calculs d'énergie libre. Ainsi, il est possible, à moyens informatiques constants, d'explorer plus longtemps la surface d'énergie potentielle du système avec ce type de solvation qu'avec des conditions périodiques.

³ typiquement le site de fixation ou le site catalytique

E. outils d'analyses

Les simulations mettant en jeu des protéines incluent un nombre de particules très important et génèrent un volume de données très important. Une des difficultés majeures concernant les trajectoires de dynamique moléculaire est certainement leur analyse. Pour ce faire, il est essentiel de pouvoir simplifier ce flot d'informations et de trouver des observables permettant de dégager les propriétés intéressantes, de préférence celles comparables à des données expérimentales.

Les méthodes d'analyse les plus couramment utilisées sont présentées ci-dessous.

1. RMS (Root-Mean-Square)

a) L'écart quadratique moyen

Il permet d'estimer la déviation moyenne existant entre deux ensembles de coordonnées A et B d'une même structure. Sa valeur est donnée par :

$$RMS_{A-B} = \sqrt{\frac{\sum_{i=1}^N (x_i^A - x_i^B)^2}{N}}$$

ou N représente le nombre total d'atomes et x_i leurs coordonnées.

b) Les fluctuations atomiques

Ces grandeurs permettent d'observer la déviation des atomes par rapport à leurs positions moyennes et donc de mettre en évidence la mobilité des atomes du système durant la simulation. Elles sont définies par :

$$\langle \Delta x_i^2 \rangle_2^1 = \sqrt{\frac{\sum_{t=1}^n (x_i(t) - \langle x_i \rangle)^2}{n}}$$

ou n est le nombre de points de la trajectoire étudiée et ou $\langle x_i \rangle$ est la valeur moyenne de la coordonnée i estimée sur les n étapes de la simulation de dynamique moléculaire. Le profil des fluctuations atomiques est calculé lors de la période de production de la dynamique. Dans le cas d'une protéine, la comparaison de ce profil avec les facteurs expérimentaux de Debye-Waller⁴ permet de juger de la validité de la simulation. Ces facteurs, aussi appelés facteurs B, sont obtenus lors de l'affinement cristallographique.

2. Courbes de distribution radiale

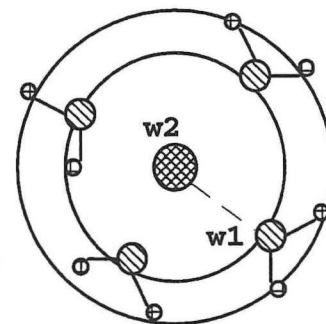
Les courbes de distribution radiale permettent d'obtenir des informations structurales concernant l'arrangement moyen des atomes du système. Ces courbes reflètent la probabilité de trouver deux atomes de type considéré à une distance donnée, (ex: O-O, O-H, pour le solvant, Ca-O, Mg-O, pour les ions). Elles permettent de caractériser l'arrangement spatial moyen des molécules de solvant entre elles ou autour de solutés tels que Ca²⁺ ou Mg²⁺. Ces fonctions de distribution, $g(r)_{w1-w2}$ s'expriment ainsi

:

$$g(r)_{w1,w2} = \frac{1}{4\pi\rho_{w1}r^2} \times \frac{\Delta N_{w1}(r)}{\Delta r}$$

⁴ Ces facteurs sont exprimés en Å² et sont, à une constante près, équivalents aux fluctuations atomiques, qui ont pour expression : $B = \frac{8}{3} * \pi^2 \langle \Delta x_i^2 \rangle$

où $\Delta N_{w_1}(r)$ est le nombre moyen d'atomes de type w_1 (moyenne effectuée sur l'ensemble des coordonnées enregistrées lors de la trajectoire de dynamique) entre deux sphères de rayon r et $r+\Delta r$ centrées sur l'atome w_2 . ρ est la densité d'atomes w_1 dans le solvant pur. Le facteur $\frac{1}{4\pi\rho_{w_1}\cdot r^2\cdot\Delta r}$ permet de normaliser cette



fonction lorsque r tend vers l'infini. Ce type de distribution est d'autant plus intéressant qu'il est directement comparable aux courbes expérimentales issues de diffractions de neutrons ou de rayons X, (voir Figure II-12). De plus, l'aire délimitée par le premier pic donne le nombre moyen de premiers voisins pour un atome w_2 donné. Ce nombre est aussi appelé nombre de coordination.

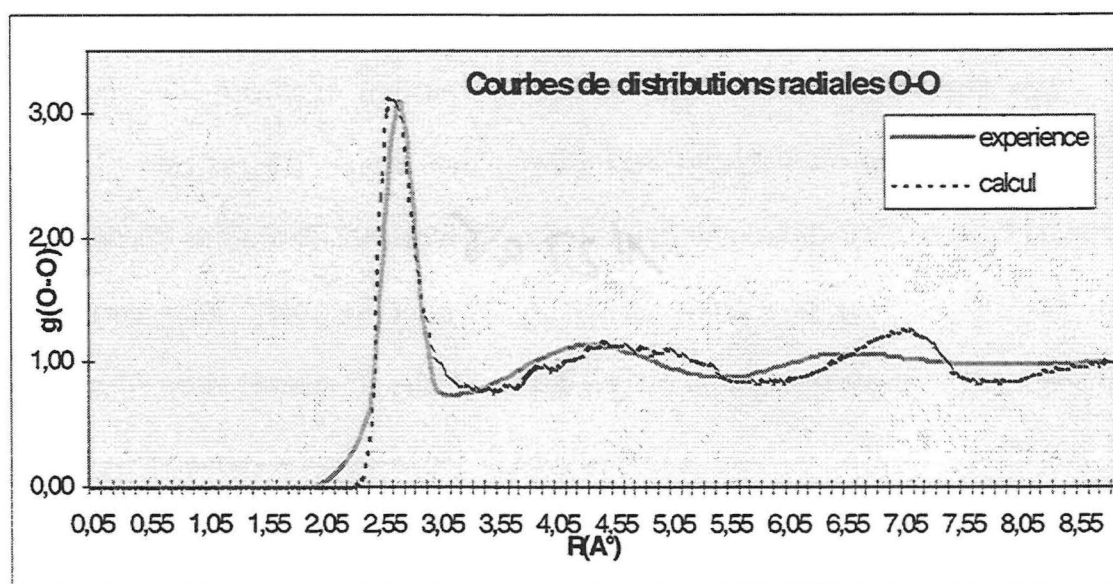


Figure II-12 exemple de courbe de distribution radiale : courbes expérimentale O-O de l'eau pure à 25°C d'après Soper et Philips(18). Les positions du premier pic et du deuxième pic donnent les distances moyennes séparant respectivement un atome d'oxygène, ses premiers voisins (2,875 Å) et ses seconds voisins, (4,5 Å). Une intégration de l'aire du premier pic (4,5 oxygènes) suggère pour l'eau liquide un arrangement moyen quasi tétraédrique des molécules d'eau. La courbe calculée est relative à une simulation de calcium en solution . cette simulation a été réalisée avec un cutoff de 7.5 Å.

Au cours de notre travail, nous avons calculé ce type de courbes pour diverses paires d'atomes: O-O, Ca-O, et Mg-O à partir de simulations de dynamique moléculaire.



3. Analyses géométriques

Au cours des simulations, il est souvent intéressant de suivre des éléments de géométrie simples tels que des distances interatomiques, des dièdres, des liaisons hydrogènes... et ceci d'autant plus qu'avec le développement de méthodes spectroscopiques de plus en plus performantes, en termes de résolution spatiale et temporelle, de plus en plus d'informations structurales et dynamiques sont disponibles. Par exemple des dièdres et des distances moyennes pourront être le cas échéant facilement comparés à leurs homologues expérimentaux déterminés en RMN à partir de constantes de couplage J_3 et de distances NOE.

Les liaisons hydrogènes, intramoléculaires ou intermoléculaires, sont extrêmement importantes car elles peuvent être à l'origine de modifications des caractéristiques dynamiques ou conformationnelles d'une molécule. Sur le plan pratique, la caractérisation de ce type de liaison est cependant difficile dans la mesure où elles sont de nature non covalente, et peuvent avoir des durées de vie assez courtes (exemple : 4 ps pour l'eau liquide à température ordinaire). C'est pourquoi ces liaisons sont souvent définies via des critères géométriques : distance $X\dots H$ et angle $X \cdots H - O_w$.

III. De la trajectoire de dynamique moléculaire aux calculs d'énergie libre

A. introduction

L'énergie libre est un des concepts les plus importants de la thermodynamique. L'énergie libre des systèmes moléculaires décrit leur tendance à s'associer et à interagir.

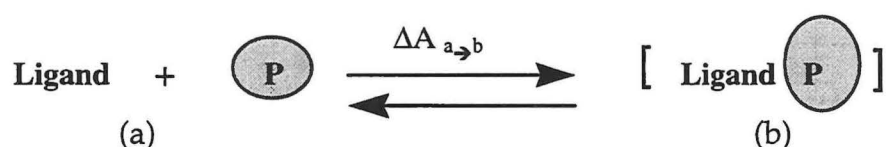


Figure III-1

C'est pourquoi, être capable de prédire cette grandeur au moyen de théories moléculaires serait une avancée extrêmement importante.

A priori, deux types d'approche permettent d'étudier un processus de reconnaissance. En effet, on peut faire le choix de calculer directement l'énergie libre associé à la formation ou la dissociation d'un complexe ligand-protéine . On qualifie cette approche de « directe » car l'énergie libre est déterminée en suivant une trajectoire de dissociation. La coordonnée de dissociation a un sens physique clair. Il existe une deuxième approche, que nous avons utilisée tout au long de ce travail. Cette approche de calcul dite « alchimique » permet de déterminer des différences d'énergie libre en



suivant des chemins non physiques. Elle est basée sur l'utilisation d'un cycle thermodynamique (19). L'énergie libre étant une fonction d'état, le changement d'énergie libre ΔA associé est indépendant du chemin suivi et ne dépend que des points de départ et d'arrivée. Ainsi, le long d'un chemin fermé nous avons $\Delta A = 0$. L'approche « alchimique » repose sur cette propriété. En effet il est possible d'imaginer un cycle dans lequel le processus spécifique étudié ne représente qu'une partie d'un cycle et permet d'accéder à la quantité d'énergie libre qui nous intéresse par le calcul de l'énergie libre associée à la partie complémentaire du cycle.

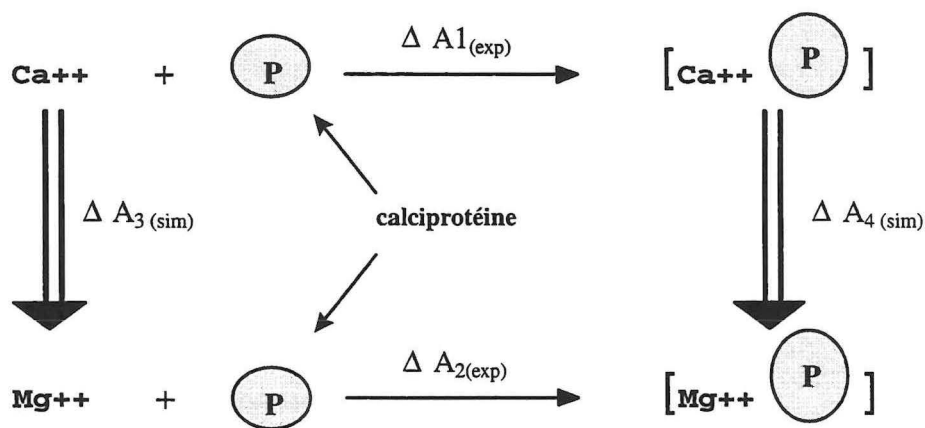


Figure III-2 : cycle thermodynamique relatif à la fixation d'une calciprotéine (P) à deux cations type Ca^{2+} et Mg^{2+} . Dans la partie droite la notation [] indique la formation d'un complexe entre le cation et la protéine.

Le cycle thermodynamique ci-dessus illustre l'application de cette approche à la fixation de cations Ca^{2+} et Mg^{2+} par une calciprotéine. Dans ce cycle nous avons : $\Delta A_{1(exp)} + \Delta A_{4(sim)} = \Delta A_{3(sim)} + \Delta A_{2(exp)}$

On peut donc accéder à la variation d'énergie libre expérimentale, $[\Delta A_{2(exp)} - \Delta A_{1(exp)}]$, par le calcul de la différence d'énergie libre $[\Delta A_{4(sim)} - \Delta A_{3(sim)}]$.

Ce cycle est en fait très général. Il est applicable à tout phénomène de reconnaissance. De manière générale, les étapes $\Delta A_{3(sim)}$ et $\Delta A_{4(sim)}$ correspondent aux transformations des ligands (en l'occurrence, les ions calcium et magnésium), dans le milieu de solvation, d'une part, et dans le site de fixation protéique, d'autre part.

Une fois déterminées, ces variations d'énergie libre nous permettent de déterminer le rapport des constantes d'affinité⁵ $k_{Ca^{2+}}/k_{Mg^{2+}}$, qui rendra compte de la sélectivité du système protéique pour ces deux ligands.

Ces étapes nécessitent donc la transformation de la nature du ligand, ce qui pour le moins est un processus « non physique ». C'est pourquoi cette approche a été qualifiée d'alchimique .

Ce chapitre est consacré à la description des techniques de calcul d'énergie libre, via des simulations de dynamique moléculaire. Nous rappellerons tout d'abord le concept de fonction de partition, qui est fondamental dans le sens où cette fonction contient toutes les informations nécessaires à l'expression de l'énergie libre. A ce titre, il s'agit d'un concept incontournable si l'on souhaite comprendre les méthodes de calcul d'énergie libre. Dans un second temps, nous décrirons plus précisément les principales méthodes de calcul d'énergie libre, ainsi que leur mise en œuvre en termes d'approches « directe » et « alchimique ». C'est dans cette partie que nous parlerons de la méthode de perturbation thermodynamique que nous avons utilisée lors de ce travail.

⁵ La constante relative de fixation (où rapport des constantes d'affinité) s'obtient ainsi :

$$\frac{K_{Ca^{2+}}}{K_{Mg^{2+}}} = e^{-\frac{(\Delta A_2 - \Delta A_1)}{RT}}$$

B. De la fonction de partition à l'expression de l'énergie libre

Définissons tout d'abord la notion de fonction de partition. Pour ce faire intéressons-nous, d'une manière générale, à la répartition d'un ensemble donné d'atomes ou de molécules sur leurs différents niveaux d'énergie potentielle, à une température donnée. Imaginons, par exemple, une mole d'un composé moléculaire M dont on connaisse parfaitement les niveaux d'énergie, que nous noterons $E_0, E_1, \dots, E_i, \dots$. Certains de ces niveaux peuvent être dégénérés, c'est à dire correspondre à plusieurs états de même énergie. On désigne par $g_0, g_1, \dots, g_i, \dots$ les multiplicités des différents niveaux, c'est à dire le nombre d'états de chacun d'eux. Si le niveau i est non dégénéré, $g_i = 1$. La connaissance des E_i et g_i (qui est exclusivement du ressort de la physique moléculaire, ou si on préfère de la chimie quantique sous l'aspect théorique et de la spectroscopie sous l'aspect expérimental), va nous suffire pour calculer toutes les grandeurs thermodynamiques du système macroscopique constitué par cette mole M.

Simplifions maintenant à l'extrême le problème. Pour ce faire considérons une transition entre deux niveaux quelconques (E_i, g_i) et (E_j, g_j). Bien que très simple, cet exemple est représentatif de bon nombre de processus physiques (collisions, absorption et émission de rayonnement).

Considérons donc des transitions entre des niveaux i et j :



A l'équilibre la loi d'action de masse nous permet d'écrire :

$$\frac{[j]}{[i]} = K = e^{-\Delta A^0 / RT} \quad \text{Equation III-1}$$

où ΔA^0 et K sont respectivement, l'enthalpie libre molaire⁶ et la constante d'équilibre standard. En thermodynamique classique ΔA^0 s'exprime ainsi :

$$\Delta A^0 = \Delta U^0 - T\Delta S^0$$

Calculons-la terme à terme. L'énergie totale molaire ΔU^0 associée à cette transformation correspond à l'expression :

$$\Delta U^0 = N_A (E_j - E_i)$$

où N_A est le nombre d'Avogadro. Il ne reste à définir que le terme entropie, que nous obtenons au moyen de la formule de Boltzmann :

$$S_i = k_b \cdot \ln \Omega_i \quad \text{Equation III-2}$$

Dans laquelle Ω_i est le nombre d'états microscopiques (ou micro-état) compatibles avec l'état macroscopique d'une mole de molécules dans le niveau i . Dans l'exemple qui nous intéresse, Ω_i est proportionnel à la multiplicité g_i de ce niveau élevée à la puissance

N_A puisque chaque molécule a le choix entre g_i possibilités. Ainsi, $\frac{\Omega_j}{\Omega_i} = \frac{g_j^{N_A}}{g_i^{N_A}}$

$$\text{d'où d'après l'Equation III-2 : } \Delta S^0 = k_b \cdot \ln \frac{\Omega_j}{\Omega_i} = k_b \cdot N_A \ln \frac{g_j}{g_i}$$

puisque par définition la constante de Boltzmann

$$k_b = \frac{R}{N_A} \quad \text{Equation III-3}$$

(R étant la constante des gaz parfaits)

⁶ Il est à noter que l'énergie libre se différencie de l'enthalpie libre par un terme PV. Ce terme étant négligeable, dans la plupart des problèmes mettant en jeu des biomolécules, il s'en suit que énergie libre et enthalpie libre sont souvent assimilées l'une à l'autre dans la littérature.

Nous obtenons donc finalement :

$$\Delta A^\circ = N_A(E_j - E_i) - RT \ln \frac{g_j}{g_i} \quad \text{Equation III-4}$$

A partir des Equation III-1 et Equation III-4 nous avons donc :

$$\frac{[j]}{[i]} = \frac{g_j}{g_i} e^{-\frac{E_j - E_i}{k_b T}} \quad \text{Equation III-5}$$

On va maintenant introduire les probabilités p_i et p_j pour que nos molécules soient respectivement dans le micro-état i et dans le micro-état j .

Ainsi on aura :

$$\frac{p_j}{p_i} = \frac{[j]}{[i]} = \frac{g_j \cdot e^{-E_j/k_b T}}{g_i \cdot e^{-E_i/k_b T}}$$

que l'on peut écrire sous la forme :

$$\frac{p_j}{g_j \cdot e^{-E_j/k_b T}} = \frac{p_i}{g_i \cdot e^{-E_i/k_b T}} = Cte \quad \text{Equation III-6}$$

En fait on peut démontrer que l'Equation III-6 a un caractère général, valable pour un micro-état quelconque (20). Cette équation introduit une constante, notée $1/Q$. Autrement dit, nous avons associé à un micro-état i une probabilité p_i :

$$p_i = \frac{1}{Q} \cdot g_i \cdot e^{-E_i/k_b T} \quad \text{Equation III-7}$$

La distribution de probabilité (Equation III-7) est appelée distribution de Boltzmann, et le terme exponentiel est le facteur de Boltzmann. Comme la somme des probabilités de tous les états possibles est un, on voit que Q n'est autre qu'un facteur de normalisation, qui sera donc :

$$Q = \sum_i g_i \cdot e^{-E_i/k_b T} \quad \text{Equation III-8}$$

L'expression de Q donnée par l'équation ci dessus est appelée fonction de partition du système. Il s'agit d'un concept fondamental de la thermodynamique statistique car toutes les fonctions thermodynamiques macroscopiques peuvent se déduire de façon simple de cette fonction de partition (20). Par exemple, l'énergie libre absolue d'un système est reliée à cette dernière par la relation suivante:

$$A = -RT \ln Q \qquad \text{Equation III-9}$$

La fonction de partition d'un système exprimée par l'Equation III-8 fait appel à l'énergie et la dégénérescence de micro-états issues de la mécanique quantique. Pour pouvoir développer une théorie statistique à partir de la mécanique classique, il nous faut introduire la notion de micro-état « classique ». Pour ce faire, il est nécessaire d'introduire, au préalable, la notion d'espace des phases. L'espace des phases associé à une particule est un espace cartésien à six dimensions dont les axes sont associés aux variables x, y, z et p_x, p_y, p_z . Une configuration d'un système correspond à la donnée explicite des quantités \vec{r} et \vec{p} que l'on représente dans l'espace des phases par un point ayant ces coordonnées. Pour un système constitué de N particules indépendantes cet espace est pour le moins difficile à représenter dans la mesure où il a alors $6N$ dimensions. Au cours de son évolution le système décrit une trajectoire dans cet espace.

Il faut noter par ailleurs que l'énergie d'un système E est toujours définie avec une incertitude δE , celle-ci étant certes très petite mais non nulle. L'énergie du système est donc comprise entre E et $E + \delta E$. La trajectoire du système dans l'espace des phases sera donc comprise entre deux hypersurfaces d'énergie correspondant à E et $E + \delta E$. On peut alors à ce stade définir la notion de micro-état classique. Pour ce faire considérons, comme exemple, une coupe de l'espace des phases à deux dimensions.



- L'approche « directe » qui permet de déterminer l'évolution de l'énergie libre au cours d'un processus moléculaire tel qu'une dissociation.
- L'approche « alchimique » que nous avons utilisée et qui, comme nous l'avons vu, permet d'accéder à des variations d'énergie libre par l'intermédiaire d'un cycle thermodynamique

Considérons tout d'abord les méthodes de calcul.

1. Intégration thermodynamique

L'« intégration thermodynamique »(21) utilise le fait que :

$$A_1 - A_0 = \int_{\lambda_1}^{\lambda_2} \frac{\partial A}{\partial \lambda} . d\lambda \quad \text{Equation III-18}$$

En supposant qu'un paramètre approprié λ peut permettre de décrire la conversion continue entre deux états d'énergie libre A_1 et A_0 , alors l'Equation III-18 peut être intégrée numériquement. Pour montrer comment intervient ce paramètre λ , récrivons III-18 en exprimant l'énergie libre A à partir de la fonction de partition :

$$A_1 - A_0 = \int_{\lambda_1}^{\lambda_2} \frac{\partial [-k_b T \ln Q(\lambda)]}{\partial \lambda} . d\lambda = -k_b T \int_{\lambda_1}^{\lambda_2} \frac{\partial [\ln Q(\lambda)]}{\partial \lambda} . d\lambda \quad \text{Equation III-19}$$

ainsi sachant que la dérivée d'une fonction logarithmique peut s'écrire sous la forme :

$$\frac{\partial [\ln F(x)]}{\partial x} = \frac{1}{F(x)} \cdot \frac{\partial F(x)}{\partial x}$$

et substituant ceci dans l'Equation III-19, on obtient :

$$A_1 - A_0 = -k_b T \int_{\lambda_1}^{\lambda_2} \frac{1}{Q(\lambda)} \frac{\partial Q(\lambda)}{\partial \lambda} . d\lambda \quad \text{Equation III-20}$$

Nous avons, en fonction de λ et à une constante multiplicative près (cf l'Equation I-12) :

$$Q(\lambda) = \int_{\text{espace}} e^{-\frac{V(q,\lambda)}{k_b T}} . dq$$

Par conséquent, l'Equation III-20 peut être réécrite sous la forme :

$$A_{\lambda} - A_{\lambda_0} = \int_{\lambda_0}^{\lambda} \left[\frac{\int \frac{\partial V(q,\lambda)}{\partial \lambda} . e^{-V(q,\lambda)/kT} dq}{\int e^{-V(q,\lambda)/kT} dq} \right] . d\lambda$$

que l'on peut noter plus simplement ainsi:

$$A_{\lambda} - A_{\lambda_0} = \int_{\lambda_0}^{\lambda} \left\langle \frac{\partial V(q,\lambda)}{\partial \lambda} \right\rangle_{\lambda} . d\lambda \quad \text{Equation III-21}$$

Ainsi, le problème du calcul d'une différence d'énergie libre par la méthode d'intégration thermodynamique se résume à celui de la détermination de la valeur moyenne de la dérivée de l'énergie potentielle par rapport à λ . Par ailleurs, l'énergie potentielle du système étant constituée d'une somme de termes (cf. champs de force), l'Equation III-21 peut être écrite comme une somme de composantes d'énergie libre associées à chacun de ces termes. Il est donc possible d'entreprendre, via la méthode d'intégration thermodynamique, l'analyse des composantes d'énergie libre associées à une transformation. En effet la dérivée partielle moyenne du potentiel total par rapport à λ est égale à la somme des dérivés partielles moyennes des termes constitutifs de ce potentiel, d'où par exemple :

$$A_{\lambda} - A_{\lambda_0} = \int_{\lambda_0}^{\lambda} \left\langle \frac{\partial V_{vdw}(q,\lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda + \int_{\lambda_0}^{\lambda} \left\langle \frac{\partial V_{elec}(q,\lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda + \int_{\lambda_0}^{\lambda} \left\langle \frac{\partial V_{bond}(q,\lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda + \dots$$

Equation III-22

Dans le même esprit, il est possible de pousser plus loin la décomposition et d'étudier plus particulièrement la contribution d'une partie du système, solvant, acide

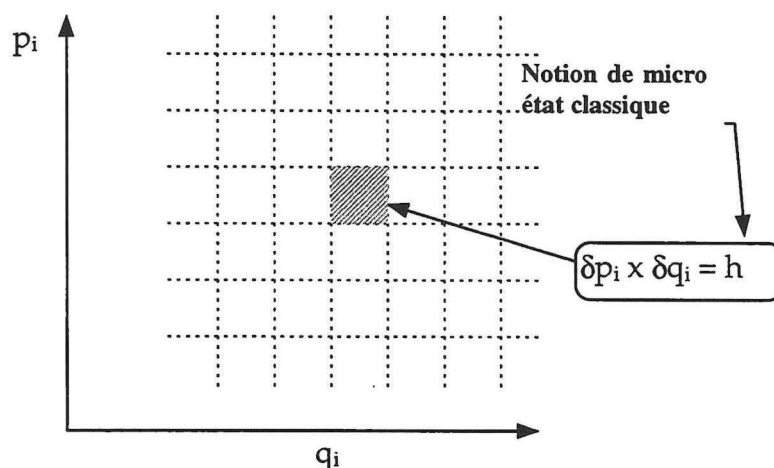


Figure III-3 : illustration dans un plan $\delta q_i \times \delta p_i$ de la subdivision de l'espace des phases en cellules élémentaires

Comme nous l'avons vu, à partir des positions et des vitesses des éléments du système il est possible de calculer l'énergie totale de ce dernier de manière quantique ou classique. Ainsi, à la place des niveaux d'énergie et des états, nous aurons simplement les positions et les vitesses de tous les atomes, ou plus exactement les coordonnées généralisées q_i et leurs moments conjugués p_i . Le niveau d'énergie E sera facile à transposer du quantique au classique: ce sera l'énergie totale correspondant à un jeu de coordonnées et de moments : c'est l'hamiltonien $H(q,p)$. Quant à la multiplicité g_i , elle sera évaluée grâce au raisonnement semi-classique suivant. Les relations d'incertitude de Heisenberg nous disent que les incertitudes sur une coordonnée q_i et sur un moment conjugué p_i ne sont pas indépendantes mais que leur produit est toujours au moins égal à la constante de Planck.

$$\delta q_i \times \delta p_i \geq h \quad \text{Equation III-10}$$

Il s'en suit que si l'on considère un petit rectangle de côté δq_i et δp_i tels que sa surface soit juste égale à h , deux états du système situés à l'intérieur de ce rectangle sont alors indiscernables l'un de l'autre. Ainsi à chaque portion de surface égale à h dans le plan



(q_i, p_i) correspondra un état classique (cf. Figure III-3). Nous appellerons micro-état classique l'ensemble des configurations qui sont dans une cellule de taille $\delta q \cdot \delta p$.

En généralisant ce raisonnement à tous les degrés de liberté de notre système, un micro-état classique correspond au contenu d'une cellule de volume h^N , de l'espace des phases du système. Ainsi, à un volume élémentaire dV de cet espace des phases il correspondra une multiplicité g_i égale à dV/h^N . Tout cela nous permet de réécrire l'Equation III-8 (relative à la fonction de partition) sous la forme suivante :

$$Q = \iint_{\text{espace}} e^{-\frac{H(q,p)}{k_b T}} \cdot \frac{dq \cdot dp}{h^n} \quad \text{Equation III-11}$$

Dans le cas de simulations ⁷, l'expression de la fonction de partition peut être encore simplifiée, car à la température ordinaire le terme énergie cinétique contenu dans $H(q,p)$ est à peu près constant (il vaut $\frac{1}{2}k_b T$ par degré de liberté) et seule l'énergie potentielle $V(q)$ varie au cours de l'exploration du relief de l'hypersurface d'énergie potentielle. On pourra donc écrire, à un facteur multiplicatif constant près la fonction de partition sous la forme:

$$Q = Cste. \int_{\text{espace}} e^{-\frac{V(q)}{k_b T}} \cdot dq \quad \text{Equation III-12}$$

A priori, d'un point de vue pratique, deux méthodes d'échantillonnage sont envisageable pour déterminer Q :

- i) La méthode de Monte Carlo qui permet de générer un ensemble de systèmes choisis de manière relativement aléatoire dans l'espace des phases.

⁷ Elle sont souvent réalisées dans l'ensemble statistique micro-canonique. Mais nous avons vu que pour N grand cet ensemble peut être considéré comme très proche de l'ensemble canonique

- ii) La dynamique moléculaire, qui, comme nous l'avons dit, permet d'explorer l'espace des phases du système en fonction du temps.

A partir de cet échantillonnage on pourra calculer des différences d'énergie libre. Dans le cas de la dynamique moléculaire, mais aussi dans le cas d'une « trajectoire » Monte Carlo, les résultats obtenus ne pourront avoir un sens que si l'hypothèse ergodique est correcte. Cet aspect, extrêmement difficile, n'est qu'imparfaitement traité à l'heure actuelle. Des progrès importants ont été faits, en particulier, ces dernières années, par des physiciens soviétiques, notamment, par Y. Sinai, qui a démontré la validité de l'hypothèse ergodique dans le cas particulier d'un gaz formé de sphères impénétrables.

D'autre part, on sait maintenant, pour un « petit » système de quelques centaines de particules, calculer numériquement les moyennes obtenues via (i) et (ii), et on les trouve effectivement égales à la précision des calculs près ; en un sens certes restreint, ces calculs sur ordinateur constituent une « démonstration ». On montre ainsi que la moyenne dans le temps d'un paramètre quelconque de ce système est égale à la moyenne de ce paramètre prise sur un ensemble de systèmes.

L'hypothèse ergodique est la justification première de l'emploi de la dynamique moléculaire lors de calculs d'énergie libre. Car on peut, à température constante, suivre l'évolution temporelle d'une seule réplique du système, au lieu de considérer un ensemble de systèmes à un instant donné, décrivant « systématiquement » les états accessibles au système. En toute rigueur, le calcul idéal exact doit faire intervenir une simulation de durée infinie, ou pour le moins une durée très longue à l'échelle des mouvements microscopiques. Or les simulations ont, en pratique, une durée d'au mieux quelques nanosecondes. Une des principales sources d'erreurs sera donc cette durée finie et relativement courte. Comme nous le verrons des aménagements sont nécessaires pour gérer au mieux cet handicap.

La méthode Monte Carlo permet aussi de générer un ensemble de configurations représentatif de l'espace des phases du système. Cette technique d'échantillonnage est cependant moins bien adaptée que la dynamique pour l'exploration de surfaces de potentiel de macromolécules biologiques autour d'une conformation donnée. Par contre la méthode de Monte Carlo permet d'observer à différents « grains » une hypersurface mal connue.

Considérons maintenant le cas de la fixation d'un cation par une calcioprotéine.



Le changement relatif d'énergie libre entre les états (a) et (b) peut s'exprimer en terme de rapport des fonctions de partition de ces systèmes. En effet d'après l'Equation III-9 :

$$\Delta A_{ab} = A_b - A_a = -k_B T \cdot \ln \frac{Q_b}{Q_a} \quad \text{Equation III-13}$$

Et selon l'Equation III-12 nous avons:

$$\frac{Q_b}{Q_a} = \frac{\int e^{-V_b(q)/k_b T} dq}{\int e^{-V_a(q)/k_b T} dq} \quad \text{Equation III-14}$$

ce qui peut s'écrire :

$$\frac{Q_b}{Q_a} = \frac{\int e^{-V_b/k_b T} \cdot e^{V_a/k_b T} * e^{-V_a/k_b T} dq}{\int e^{-V_a/k_b T} dq} \quad \text{Equation III-15}$$

expression que l'on peut noter sous la forme ⁸:

$$\frac{Q_b}{Q_a} = \langle e^{-[\Delta V_{ab}/k_b T]} \rangle_a \text{ avec } \Delta V_{ab} = v_b - v_a \quad \text{Equation III-16}$$

Dans l'Equation III-16, $\langle \rangle_a$ signifie que la moyenne est prise sur l'espace des phases correspondant à l'état (a) : $(e^{-[\Delta V_{ab}/k_b T]})$ est ainsi moyenné, avec pondération, sur l'ensemble des états énergétiques accessibles au système (a). Le terme ΔV est la différence entre l'énergie potentielle de l'état (a) et celle de l'état (b) en un point quelconque de l'espace des phases. Ainsi le rapport des fonctions de partition peut être déterminé à partir d'un ensemble de différences d'énergie potentielle entre l'état de référence (a) et l'état « perturbé » (b)

La variation d'énergie libre peut donc être mise sous la forme :

$$\Delta A_{ab} = -k_B T \cdot \text{Ln} \langle e^{-[\Delta V_{ab}/k_b T]} \rangle_a \quad \text{Equation III-17}$$

D'un point de vue fondamental, le calcul d'une différence d'énergie libre peut donc se faire via la détermination du rapport $\frac{Q_b}{Q_a}$. Notons que L'Equation III-17 est exacte et peut donc être appliquée à une transformation quelconque.

⁸ Compte tenu de l'Equation III-7, qui donne la probabilité d'un état, l'Equation III-15 peut être réécrite : $\frac{Q_b}{Q_a} = \int e^{-(V_b - V_a)/k_b T} \cdot P_a dq_a$ où P_a est la probabilité d'un état de référence (a)

C. Calcul d'énergie libre par simulation de dynamique « classique »

Dans le paragraphe précédent, nous avons vu qu'à partir d'un échantillonnage de l'espace des phases du système, il est possible de déterminer une différence d'énergie libre. En pratique si l'état de référence et l'état perturbé sont relativement différents⁹- ce qui est le cas, de loin, le plus fréquent - on ne peut pas espérer déterminer directement le rapport $\frac{Q_b}{Q_a}$ via une simulation. En effet, les phénomènes de reconnaissance impliquent d'importantes réorganisations du solvant et de la protéine. Ces changements sont lents. Il s'en suit que d'un point de vue général, la variation d'énergie libre associée aux processus de reconnaissance ne peut être calculée directement à partir de l'échantillonnage réalisé lors d'une seule simulation de dynamique moléculaire mettant en jeu le ligand et le système hôte.

Deux grandes méthodes de calcul des différences d'énergie libre sont principalement utilisées. Elles font intervenir, de manière générale, un paramètre noté λ qui décrit une transition continue d'un état à un autre. L'utilisation et la nature de ce paramètre varie en fonction de la méthode de calcul mais aussi selon l'approche choisie.

Nous traiterons successivement l'aspect méthodologique avec d'une part la méthode dite d'intégration thermodynamique et d'autre part celle de perturbation thermodynamique, puis, l'aspect mise en œuvre des deux approches que nous avons évoquées en introduction de ce chapitre, à savoir :

⁹ c'est-à-dire pour toute transformation mettant en jeu une variation d'énergie nettement supérieure à $k_b.T$

aminé particulier, etc. L'équation III-22 est exacte, mais il serait dangereux de l'interpréter comme la somme de termes d'énergie libre relatifs chacun à un type d'interaction. En effet chaque moyenne est calculée sur un même espace des phases, dont l'échantillonnage fait intervenir le potentiel dans son ensemble. Cette réserve étant faite, la décomposition I-22 peut fournir une idée qualitative de l'importance des différents termes d'interaction. Cette question, est cependant aujourd'hui encore en cours de discussion (3)(22)(4).

2. perturbation thermodynamique (23)

Comme nous l'avons évoqué plus haut, les différences d'énergie libre ne peuvent être calculées directement à partir de l'Equation III-17 pour des systèmes de grande taille. Cependant, l'énergie libre étant une fonction d'état, et donc indépendante du chemin suivi, il est possible pour la calculer de subdiviser la transformation étudiée en plusieurs étapes. Pour ce faire, une variable λ décrivant un système est associée à la transformation. On fait en général varier λ de 0 à 1. Dans l'état initial de référence λ prend la valeur 0 et dans l'état final la valeur 1.

On peut découper l'intervalle (0,1) en n petits tronçons. Chacun de ces tronçons est équivalent à une micro transformation à laquelle sont associés un λ initial et un $\lambda+\Delta\lambda$ final. Si λ est assez petit il s'agira d'un traitement dit de perturbation. L'équation ci-dessous résulte directement de l'application de l'Equation III-17 au calcul de la différence d'énergie libre associée à une de ces micro transformations :

$$\Delta A_{\lambda \rightarrow \lambda \pm \Delta \lambda} = -k_B T \cdot \text{Ln} \left\langle e^{-\frac{V(\lambda) - V(\lambda \pm \Delta \lambda)}{k_B T}} \right\rangle_{\lambda} \quad \text{Equation III-23}$$

Elle permet à partir de l'échantillonnage des états du système réalisé à une valeur de λ de déterminer la différence d'énergie libre $\Delta A_{\lambda \rightarrow \lambda + \Delta\lambda}$. De plus, cette équation étant symétrique, on pourra déterminer à partir du même échantillonnage la quantité $\Delta A_{\lambda \rightarrow \lambda - \Delta\lambda}$.

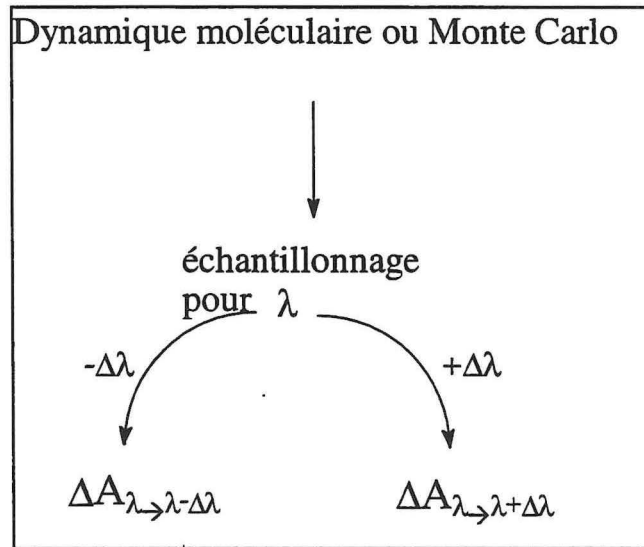


Figure III-5 : technique du double fenêtrage

On pourra donc, à partir d'un seul échantillonnage réalisé à une valeur de λ donnée, calculer les variations d'énergie libre associées à deux fenêtres de taille $\Delta\lambda$ centrées sur λ (cf. Figure III-5). D'où si par exemple $n=10$

$$\Delta A_{(ab)} = -\Delta A_{(\lambda=0.1; \lambda=0)} + \Delta A_{(\lambda=0.1; \lambda=0.2)} \dots \dots -\Delta A_{(\lambda=0.9; \lambda=0.8)} + \Delta A_{(\lambda=0.9; \lambda=1.0)}$$

Equation III-24

Il suffira donc d'échantillonner les états du système pour n valeur de λ , puis en sommant les variations d'énergie libre déterminées selon Equation III-23 sur les $2n$ petites variations calculées, on pourra exprimer la variation d'énergie libre totale associée à la transformation par :

$$\Delta A_{ab} = A(\lambda_1) - A(\lambda_0) = -k_B T \cdot \sum_{i=1}^{2n} \ln \left\langle e^{-\left[\frac{V(\lambda_i) - V(\lambda_i \pm \Delta\lambda)}{k_B T} \right]} \right\rangle_{\lambda_i} \quad \text{Equation III-25}$$

3. Paramètre λ et type de calcul

On peut remarquer que l'Equation III-25 est parfaitement générale au même titre que l'Equation III-21 à l'origine de la méthode d'intégration thermodynamique. Ces équations sont exactes et donc valables quelle que soit la nature du système étudié. Cependant, jusqu'à présent nous n'avons discuté ni de la nature de λ , ni de la relation liant λ au potentiel du système. Nous allons voir que c'est la nature de λ qui fait la différence entre les deux approches directe et alchimique que nous avons évoquées.

Considérons ces deux approches.

a) Approche directe

Selon cette approche on considère le paramètre λ comme une véritable coordonnée de dissociation.

Ainsi, dans le cas d'un processus de fixation, une coordonnée appropriée est la distance entre les centres de masse des entités à séparer, qui varie alors d'un état initial, où le ligand et la protéine sont associés, à un état final où ils sont « totalement » dissociés.

Si le système a été orienté de sorte que cette coordonnée soit alignée sur l'axe z , alors, à partir d'un échantillonnage réalisé pour une distance inter-centres de masse (z) fixe, il est possible de calculer la différence d'énergie libre associée à une petite variation de z selon la méthode de perturbation thermodynamique, ou bien la dérivée de l'énergie libre à cette valeur de z avec les équations d'intégration thermodynamique. Pour ce faire, le paramètre λ précédemment évoqué est remplacé par la coordonnée de chemin z dans les équations relatives à cette approche.

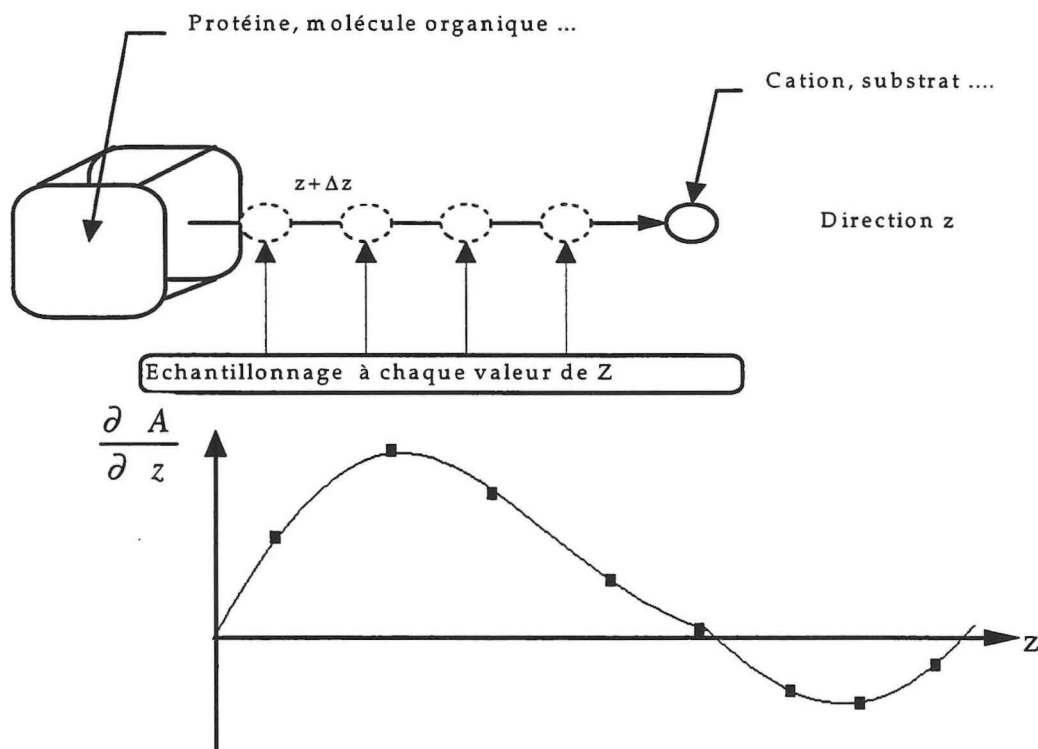


Figure III-6 : représentation schématique d'un calcul d'énergie libre direct associé à la dissociation d'un complexe moléculaire. A partir de l'échantillonnage effectué à chaque valeur de la coordonnée z , le terme $\partial A / \partial z$ est calculé, permettant de calculer un profil d'énergie libre

Par exemple, si on effectue cette substitution dans le cas de la méthode d'intégration thermodynamique nous obtenons :

$$\frac{\partial A}{\partial z} = \left\langle \frac{\partial V(q,z)}{\partial z} \right\rangle_z \quad \text{Equation III-26}$$

Dans cette équation $\frac{\partial A}{\partial z}$ est la dérivée partielle du potentiel suivant z . Il est à noter que cette quantité correspond à la force le long de cette direction. C'est pourquoi, on appelle parfois l'énergie libre ainsi obtenue « potentiel de force moyenne ». Par calcul

successif de la force moyenne pour différentes valeurs de z entre z_0 et z_1 on obtient donc le profil de la force moyenne, en fonction de z (cf. Figure III-6). L'intégration de ce profil donnera le profil d'énergie libre de la dissociation selon z . Cette intégration est réalisée par intégration numérique de l'Equation III-27.

$$A_{z_1} - A_{z_0} = \int_{z_0}^{z_1} \left\langle \frac{\partial V(q, z)}{\partial z} \right\rangle_z . dz \quad \text{Equation III-27}$$

Ce type d'approche a émergé au début des années 90, avec quelques études de processus de dissociation relatifs à des systèmes pouvant avoir des tailles et des complexités très différentes, depuis des édifices moléculaires relativement petits et simples comme dans le cas de la dissociation d'une paire d'ions en solution (24) ou de complexes cation-éther-couronne (25-27) jusqu'à un système de la taille d'un complexe antigène anticorps en solution (soit 40000 atomes) (28) . Cette approche a par exemple aussi été appliquée à l'étude de la fusion d'agrégats de sodium et d'hélium (29).

Il faut noter que cette technique permettant d'obtenir un profil d'énergie libre, elle a pour avantage, en plus de permettre de déterminer la variation d'énergie libre totale, de permettre d'accéder à l'énergie libre d'activation¹⁰ du processus étudié. Dans la majorité des études, l'accord avec l'expérience est souvent qualitatif en ce qui concerne la variation d'énergie libre totale. De même, dans la plupart des cas, ce type d'approche n'a pas permis de reproduire quantitativement les valeurs expérimentales des hauteurs de barrières d'activation.

¹⁰ Il s'agit de la hauteur de la barrière du profil d'énergie libre, c'est à dire du maximum de la courbe $\Delta A=f(\text{coordonnée de dissociation})$

Cette approche est de plus sur un plan pratique extrêmement exigeante en temps de calcul. En effet, par essence, les simulations d'énergie libre, en général, ont besoin d'une durée d'échantillonnage longue. Or dans l'approche « directe », il faut étudier un système de taille suffisamment grande afin de décrire convenablement le solvant autour du ligand et de la protéine et ce pour chaque valeur de z .

De plus, dans le cadre de notre travail, cette approche impliquerait obligatoirement deux simulations indépendantes relatives l'une à la dissociation du Calcium et l'autre à celle du Magnésium et ce pour chacun des sites des calciprotéines envisagées. C'est pourquoi nous avons préféré utiliser une autre approche.

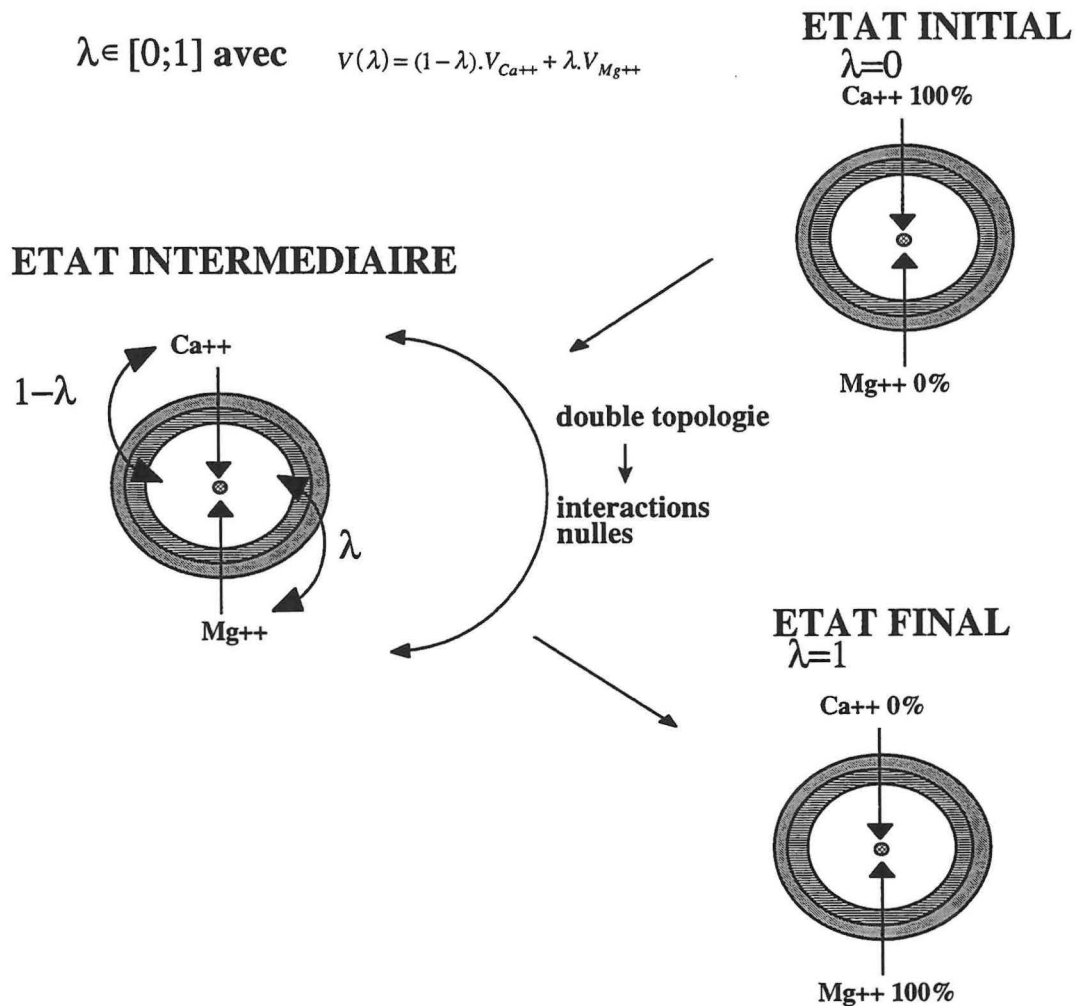
b) L'approche « alchimique ».

Comme nous l'avons vu au début de ce chapitre, cette approche fait intervenir un cycle thermodynamique, qui permet d'accéder à la différence d'énergie libre expérimentale, par le calcul de deux variations d'énergie libre alchimique (dans l'eau et dans la protéine¹¹). La détermination de ces quantités nécessite la transformation progressive de la nature du cation ($\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$) durant une simulation. L'énergie libre étant une fonction d'état, le chemin choisi est arbitraire. En conséquence, la dépendance entre λ et le potentiel du système peut être traitée d'une infinité de façons. La méthode la plus simple et la plus répandue consiste à définir le potentiel du système comme une combinaison linéaire du potentiel de l'état de départ et de celui du point d'arrivée. Dans le cas qui nous intéresse nous aurons donc :

$$V(\lambda) = (1 - \lambda).V_{\text{Ca}^{++}} + \lambda.V_{\text{Mg}^{++}} \quad \text{Equation III-28}$$

¹¹ Contrairement à l'approche directe, l'approche alchimique nécessitera de faire une seule simulation par site de calciprotéine étudié.

Comme on peut le voir sur la figure ci-dessous, la transformation progressive se traduit par la variation du coefficient de "perturbation" λ , de 0 à 1. A $\lambda=0$, le système interagit avec Ca^{2+} et à $\lambda=1$, avec Mg^{2+} . Dans l'intervalle, le système interagit avec un cation "hybride".



Regardons de plus près le potentiel du cation « muté ». Selon notre champ de force, les cations interagissent ainsi avec le reste du système:

$$V_{\text{inter}} = \sum_{i \in \{\text{paires-cation-atomes-non-lié}\}} \left[\left(\frac{A_{\text{ion}-i}}{r_{\text{ion}-i}^{12}} - \frac{B_{\text{ion}-i}}{r_{\text{ion}-i}^6} \right) + \frac{q_{\text{ion}}q_i}{\epsilon \cdot r_{\text{ion}-i}} \right]$$

si on introduit l'Equation III-28 dans cette expression nous avons:

$$V_{inter}(\lambda) = \sum_{i \in \{cation-atome_i\}} \left[(1-\lambda) \left(\frac{A_{Ca-i}}{r_{Ca-i}^{12}} - \frac{B_{Ca-i}}{r_{Ca-i}^6} + \frac{q_{Ca}q_i}{\epsilon \cdot r_{Ca-i}} \right) + \lambda \left(\frac{A_{Mg-i}}{r_{Mg-i}^{12}} - \frac{B_{Mg-i}}{r_{Mg-i}^6} + \frac{q_{Mg}q_i}{\epsilon \cdot r_{Mg-i}} \right) \right]$$

La transformation du cation peut d'un point de vue pratique être décrite de deux manières.

1) l'approche topologique double, qui correspond en fait à l'apparition et la disparition de deux cations indépendants. La topologie double met en œuvre dans la simulation deux fragments, avec chacun leurs coordonnées, mais ces fragments interagissent avec l'environnement selon λ .

2) l'approche topologique simple, qui fait intervenir un seul cation dont on modifie progressivement les paramètres. La topologie simple ne fait intervenir qu'un seul fragment, ce qui implique qu'un atome i du système non muté voit un seul cation hybride à une distance r_{ion-i} . De plus, le calcium et le magnésium ayant même charge, l'énergie d'interaction totale de ce cation hybride est donc:

$$V_{inter}(\lambda) = \sum_{i \in \{cation-atome_i\}} \left[\left(\frac{A_{Ca-i}}{r_{ion-i}^{12}} - \frac{B_{Ca-i}}{r_{ion-i}^6} + \frac{q_{ion}q_i}{\epsilon \cdot r_{ion-i}} \right) + \lambda \left(\frac{B_{Ca-i} - B_{Mg-i}}{r_{ion-i}^6} - \frac{A_{Ca-i} - A_{Mg-i}}{r_{ion-i}^{12}} \right) \right]$$

Equation III-29

Et seule la composante van der Waals de cette énergie d'interaction est dépendante du paramètre λ .

Par ailleurs, l'approche alchimique peut permettre de suivre le comportement de l'environnement du cation lors de sa transformation progressive du Ca^{2+} au Mg^{2+} . Car



bien que le chemin associé à λ soit non physique, l'état initial et l'état final n'en demeurent pas moins pertinents.

(1) perturbation ou intégration thermodynamique?

Nous avons vu qu'à partir d'un échantillonnage adéquat, on peut déterminer la variation d'énergie libre soit par intégration thermodynamique, soit par perturbation thermodynamique. Cependant nous n'avons pas encore discuté de l'efficacité de ces méthodes. C'est ce à quoi nous allons nous attacher à présent. Récemment, Pearlman (30) a réalisé une étude comparative des méthodes d'intégration et de perturbation thermodynamiques. Ces dernières ont été appliquées alternativement avec une description topologique double (utilisée par des programmes de dynamique moléculaire largement usités tel que CHARMM (2)) ou avec une approche topologique simple (utilisée en particulier par un programme non moins usité tel que AMBER (31)). Pearlman a étudié une transformation modèle simple, dont la variation d'énergie libre associée est connue très exactement. Elle est nulle : il s'agit de la transformation de l'éthane en éthane en solution aqueuse (cf. Figure III-7)

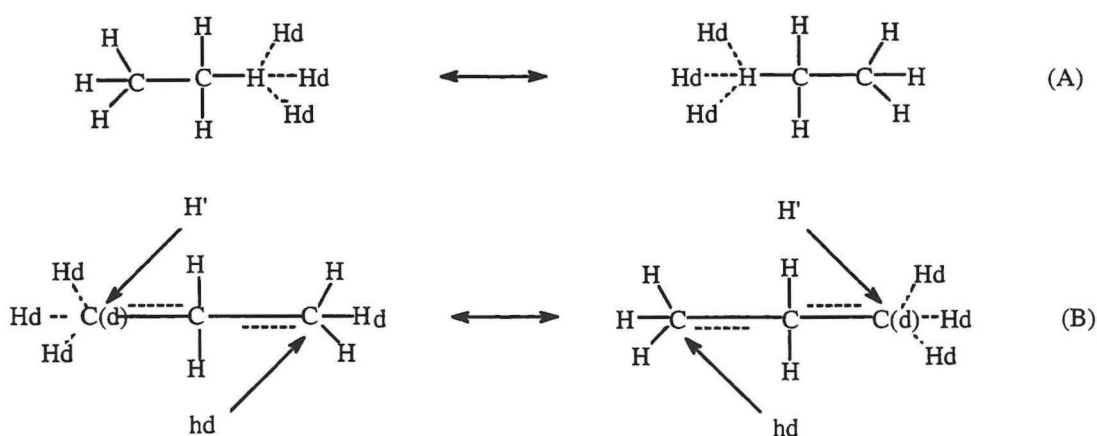


Figure III-7 :représentation schématique du système éthane-éthane utilisé par Pearlman(30) pour représenter topologie simple (A) et topologie duale (B). les lignes pointillées représentent les liaisons aux atomes qui disparaissent - noté avec l'indice « d »-.

Au vu des résultats obtenus, Pearlman arrive aux conclusions suivantes :

- i) Concernant l'approche topologique, quelle que soit la méthode de calcul utilisée (intégration ou perturbation thermodynamique), l'approche topologique simple semble la plus efficace.
- ii) L'intégration et la perturbation thermodynamique sont d'une efficacité comparable lorsqu'elles sont utilisées avec une topologie simple. Néanmoins, l'intégration thermodynamique est plus efficace, lors de l'utilisation d'une topologie double.

Notons que, comme le souligne Kollman (32), un des principaux inconvénients de la méthode de perturbation thermodynamique semble être que l'énergie libre totale (cf. Equation III-25) ne peut être décomposée en somme de composantes d'énergie libre, car le logarithme d'une exponentielle avec différentes composantes énergétiques n'est pas égal à la somme des logarithmes des composantes individuelles. Alors que comme nous l'avons vu, la méthode d'intégration thermodynamique permet d'entreprendre l'analyse des composantes d'énergie libre associées à la transformation.

Pour clore cette discussion, il faut mentionner que, bien sûr, outre les choix de la méthode de calcul et de l'approche topologique, les calculs d'énergie libre sont aussi sensibles à la précision de la description énergétique du système ainsi qu'aux paramètres relatifs au protocole d'échantillonnage, tels que : le cutoff , le nombre d'étapes de calculs et la durée de la simulation. En fait, le choix de ces paramètres de calcul est beaucoup plus critique que celui de la méthode de calcul. Il n'y a pas ou peu de règles générales.

Par exemple, lors de l'étude de Pearlman, pour un système aussi simple que l'éthane en solution dans l'eau plus de 700 ps d'échantillonnage au total ont été

nécessaires pour obtenir une bonne convergence du résultat. Ceci n'a cependant pas un caractère général, car la durée d'échantillonnage est fortement dépendante de la nature du système étudié ainsi que de la transformation considérée. Par exemple, le calcul de l'énergie de solvation de Ca^{2+} en solution (soit près de 380 kcal/mol), a été déterminée à une précision de ~ 1 kcal en effectuant la transformation $\text{Ne} \rightarrow \text{Ca}^{2+}$ lors d'une simulation de 80 ps (8). L'étude de Pearlman est sans doute biaisée par l'utilisation de conditions périodiques qui accroissent considérablement les temps de relaxation de son système.

Pour chaque problème considéré, il faut donc mettre au point un protocole adéquat. Cette optimisation consiste à trouver le bon compromis entre le découpage de la transformation (nombre d'étapes), la durée d'échantillonnage et le coût total en calcul, sachant que d'une part le système doit être équilibré à chaque étape et que d'autre part la convergence du résultat sera d'autant meilleure que la durée d'échantillonnage sera plus importante.



IV. Références Bibliographiques

1. Born, M., and Oppenheimer, R. (1927) *Ann. Physik* **84**, 457
2. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) *J. Comput. Chem* **4**, 187
3. Warshel, A., and Levitt, M. (1976) *J. Mol. Biol.* **103**, 227
4. Field, M. J., Bash, P. A., and Karplus, M. (1990) *J. Comput. Chem.* **11**(6), 700-733
5. Gerschel, A. (1995), pp. 96-97, Interéditions CNR éditions, Paris
6. Shiratori, Y., and Nakagawa, S. (1991) *J. Comput. Chem.* **12**, 717-730
7. Periole, X., Allouche, D., Daudey, J.-P., and Sanejouand, Y.-H. (1997) *J.Phys.Chem.* **101**, 5018-5025
8. Aqvist, J. (1990) *J.Phys.Chem.* Vol **94**(Iss 21), pp 8021-8024
9. Verlet, L. (1967) *Phys. Rev.* **159**, 98-103
10. Van Gusteren, W. F., and karplus, M. (1982) *Macromolecules* **15**, 1528-1543
11. Rickaert, J. P., Ciccotti, G., and Berensen, H. J. C. (1977) *J. Comput. Chem* **23**, 327-336
12. Sanejouand, Y. H., and Tapia, O. (1995) *J Phys Chem* **99**(15), 5698-5704
13. Van Gusteren, W. F., and Berensen, H. J. C. (1977) *Mol. Phys.* **34**, 1311
14. Theodorou, D. N., and Suter, U. W. (1985) *J.Chem.Phys.* **82**, 955
15. Adams, D. J. (1979) *Chem. Phys. Lett.* **62**, 329
16. Alary, F., Durup, J., and Sanejouand, Y.-H. (1993) *J.Phys.Chem.* **97**, 13864-13876
17. van Gusteren, W. F., and Berendsen, H. J. C. (1985) *In Molecular Dynamics and protein Structure*, Hermans, J. Ed. Poly cristal book Serv (Hermans, J., Ed.), III. III vols., Springer, Western
18. Soper, A. K., and Philips, M. G. (1987) *J.Chem.Phys.* **107**, 47
19. Van Gusteren, W. F., and Berensen, H. J. C. (1987) *Journal of Computer-Aided Molecular Design* **1**, 171-176
20. Diu, B., Guthmann, C., Lederer, D., and Roulet, B. (1989) *Physique statistique (arts, E. d. s. e. d., Ed.)*, Herman, Paris
21. Kirkwood, J. (1935) *J. Chem. Phys.* **3**(300), 300

22. Smith, P. E., and Van Gunsteren, W. F. (1994) *J.Phys.Chem* **98**, 13735-13740
23. Zwanzig, R. (1954) *J. Chem. Phys.* **22**(1420), 1420
24. Dang, L. X., Rice, J. E., and Kollman, P. A. (1990) *J. Chem. Phys.* **93**, 7528-7529
25. Dang, L. X., and Kollman, P. A. (1990) *J.A.C.S.* **112**(15), 5716-5720
26. Dang, L. X., and Kollman, P. A. (1995) *J.Phys.Chem.* **99**(1), 55-58
27. Kowall, T., and Geiger, A. (1995) *J.Phys.Chem.* **99**(15), 5240-5246
28. Durup, J., and Alary, F. (1995) *Molecular dynamics study of the dissociation of an antigen-antibody Complex in solution*. Modelling of Biomolecules Structures and Mechanisms ((eds), P. a., Ed.), Kluwer Academic Publishers, Netherlands
29. Calvo, F., and Spiegelmann, F. (1996) *Phys Rev B-Condensed Matter* **54**(15), 10949-10958
30. Pearlman, D. A. (1994) *J.Phys.Chem.* **98**, 1487-1493
31. Pearlman, D. A., Case, D. A., Caldwell, J. C., Seibel, G. L., Singh, U. C., Weiner, P., and Kollman, P. A. (1991) *AMBER 4.0*, University of California, San Francisco
32. Kollman, P. A. (1993) *Chemical Reviews* **93**, 2395-2417

PARTIE III :

RÉSULTATS





A) CATIONS EN SOLUTION

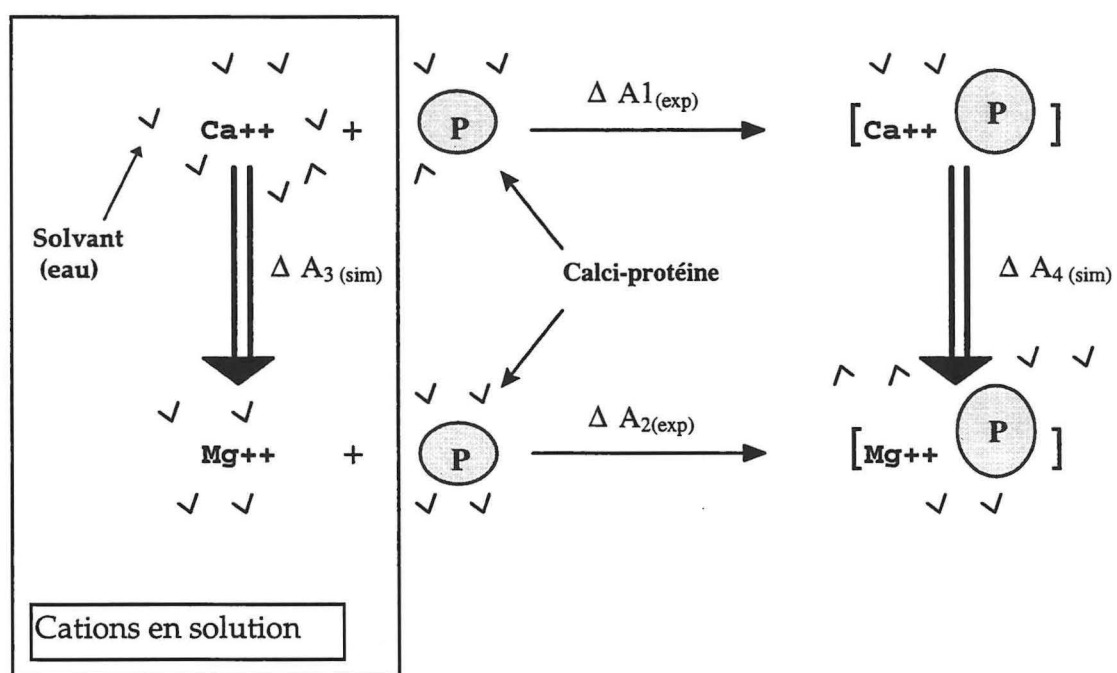




Table des matières

| | |
|--|------------|
| I. TRAJECTOIRE DE DYNAMIQUE MOLÉCULAIRE DE CATIONS EN SOLUTION | 128 |
| A. PROTOCOLE DE CALCUL | 128 |
| 1. CONDITIONS DE SOLVATATION | 129 |
| 2. CHOIX DES PARAMÈTRES | 130 |
| B. ANALYSE DES TRAJECTOIRES DE CATIONS EN SOLUTION | 134 |
| 1. COMPORTEMENT DU SOLVANT : RDF O-O | 134 |
| 2. ETUDE STRUCTURALE DES COUCHES DE SOLVATATION DES CATIONS | 139 |
| 3. INFLUENCE DU CUTOFF | 147 |
| 4. GÉOMÉTRIE DE LA PREMIÈRE COUCHE DE SOLVATATION | 152 |
| C. CONCLUSION | 159 |
| II. CALCULS D'ÉNERGIE LIBRE : DÉTERMINATION DE LA DIFFÉRENCE D'ÉNERGIE LIBRE DE SOLVATATION CA²⁺ MG²⁺ | 160 |
| A. PROTOCOLE DE CALCUL | 160 |
| B. SIMULATION À CUTOFF 14 Å | 164 |
| 1. INFLUENCE DU NOMBRE D'ÉTAPES DE LA TRANSFORMATION | 164 |
| 2. ANALYSES APPROFONDIES | 167 |
| C. SIMULATION À CUTOFF 7,5 Å | 179 |
| 1. INFLUENCE DE $\Delta\lambda$ | 179 |
| 2. ANALYSES APPROFONDIES | 180 |
| D. CONCLUSION | 186 |
| III. RÉFÉRENCES BIBLIOGRAPHIQUES : | 191 |



Introduction

Comme nous avons pu le voir précédemment, la stratégie la plus adaptée pour l'étude du processus qui nous intéresse semble être l'approche « alchimique ». Cette voie nécessite le passage par un cycle thermodynamique pour déterminer le rapport $K_{Ca^{2+}}/K_{Mg^{2+}}$. Nous sommes ainsi amenés à effectuer la transmutation $Ca^{2+} \rightarrow Mg^{2+}$ dans l'eau et dans les sites de fixation de cations de calcioprotéines.

Il est important de noter, comme nous l'avons évoqué dans le paragraphe relatif au champ de force, que la description des atomes dans nos simulations est basée sur une interaction effective constituée de termes électrostatiques et van der Waals. Cette description pourrait paraître problématique pour des atomes chargés. Avant d'effectuer les simulations d'énergie libre proprement dites, nous avons fait des simulations « test » de cations en solution afin de pouvoir nous rendre compte de la pertinence de nos paramètres et des conditions de solvation.

Puis, dans un deuxième temps nous considérerons les simulations relatives aux calculs de différences d'énergie libre de solvation du calcium et du magnésium dans l'eau. Ces simulations n'ont pas en soi un caractère prévisionnel particulier, car cette quantité est relativement bien connue via l'expérience. Cependant, ce système est, là encore, un bon modèle pour mettre au point notre protocole de calcul.

I. Trajectoire de dynamique moléculaire de cations en solution

Dans un premier temps, nous avons évalué la capacité des paramètres du Ca^{2+} et du Mg^{2+} à reproduire les éléments de la géométrie des premières couches de solvation de ces cations. Cette évaluation a été effectuée à partir des courbes de distribution radiale cation atomes de solvant déterminées à partir de simulations de dynamique moléculaire de cations en solution dans l'eau (à température ambiante).

A. Protocole de calcul

Ces simulations font intervenir le protocole suivant :

- Un pas d'intégration de 2 fs grâce à l'utilisation de l'algorithme SHAKE
- Une fonction de lissage type shift est utilisée dans le traitement des interactions électrostatiques et van der Waals.
- 3 ps de thermalisation, durant laquelle la température est progressivement amenée de 0 à 300 K, par incrémentation de température de 5 K tous les 25 pas .
- Une période de relaxation de 18 ps est réalisée afin de laisser le système atteindre l'équilibre.
- Puis, 20 ps de trajectoire sont effectuées durant lesquelles nous avons enregistré un ensemble de 500 configurations. De cet ensemble seront extraites les informations.
- Pendant la période d'équilibration et ensuite, les vitesses sont corrigées tous les 500 pas d'intégration si la température moyenne du système n'est pas comprise dans une fenêtre de plus ou moins 5K centrée sur la température escomptée (300K).
- afin d'étudier l'influence du cutoff, pour chacun des cations, suivant le même protocole et avec les mêmes coordonnées et vitesses de départ, nous avons réalisé des simulations à deux valeurs de cutoff, 7,5 et 14 Å.

Avec 41 picosecondes, la durée totale de nos simulations peut paraître courte, toutefois elle est suffisamment importante pour nous permettre d'étudier la géométrie de la première couche de solvation. En effet, compte tenu de la forte intensité du champ électrostatique généré par les cations divalents, on peut penser que l'orientation des molécules d'eau autour d'eux est rapide. Par contre, l'étude de la dynamique d'échange entre première et deuxième couche de solvation est hors de portée de tels calculs puisque expérimentalement les temps de résidence d'une molécule d'eau en première couche est respectivement de l'ordre de 10^{-5} s et 10^{-8} s pour le magnésium et pour le calcium (1).

1. conditions de solvation

Nous avons fait le choix d'utiliser les conditions de solvation « extended wall system » (2) qui ont pour vertu d'être les moins onéreuses, du point de vue des ressources informatiques. Notre système est constitué d'un cation Ca^{2+} ou Mg^{2+} - ou "hybride" lors des calculs d'énergie libre - solvato au centre d'une sphère d'eau de 15 Å de rayon (soit 460 molécules d'eau). La limite de cette sphère est faite de deux barrières successives, l'une constituée de molécules d'eau totalement fixées (celles situées à plus de 11 Å du cation), l'autre, constituée de molécules d'eau dont l'oxygène est fixé (celles situées à plus de 9 et à moins de 11 Å du cation) formant ainsi une frontière semi-rigide. Les molécules d'eau situées dans une sphère de 9 Å autour du cation sont donc totalement libres de bouger (cf. Figure I-1). Dans toutes nos trajectoires, le cation est maintenu fixe au centre de la sphère de solvation.



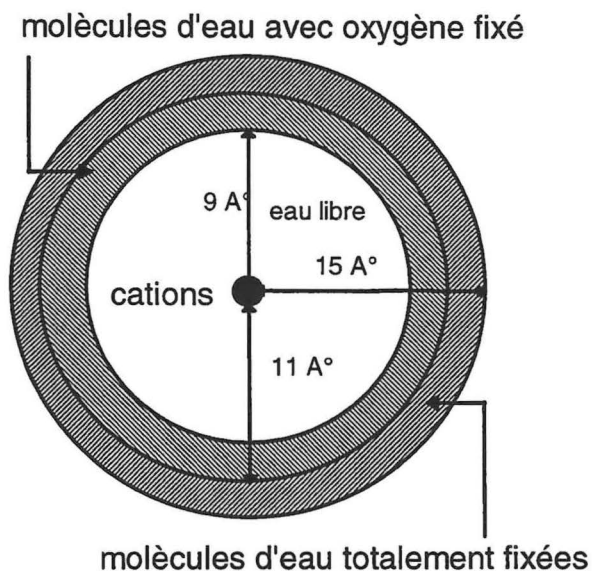


Figure I-1 : représentation schématique des conditions de solvation.

2. Choix des paramètres

Concernant le solvant, nous avons utilisé un des modèles d'eau les plus courants, à savoir, le modèle d'eau "TIP3P"⁽³⁾. Ce dernier fait intervenir un potentiel à deux corps, qui a été déterminé de manière à reproduire, lors de simulation, la densité ainsi que la chaleur latente de vaporisation de l'eau pure.

Concernant les paramètres des cations, nous avons utilisé les paramètres développés par Åqvist (4). Ces paramètres ont été obtenus, selon la procédure illustrée Figure I-2. Le principe est le suivant: de manière itérative, les paramètres des cations sont adaptés de façon à ce que l'énergie libre de solvation et la position du premier pic de la courbe de distribution radiale cation - oxygènes calculés lors de simulations soient en accord avec les données expérimentales.

¹ TIP3P pour Transferable Intermolecular Potential 3 points. Une charge électrostatique est disposé sur chacun des atomes . Ce modèle est rigide, c'est à dire que la géométrie de la molécule d'eau est fixe.

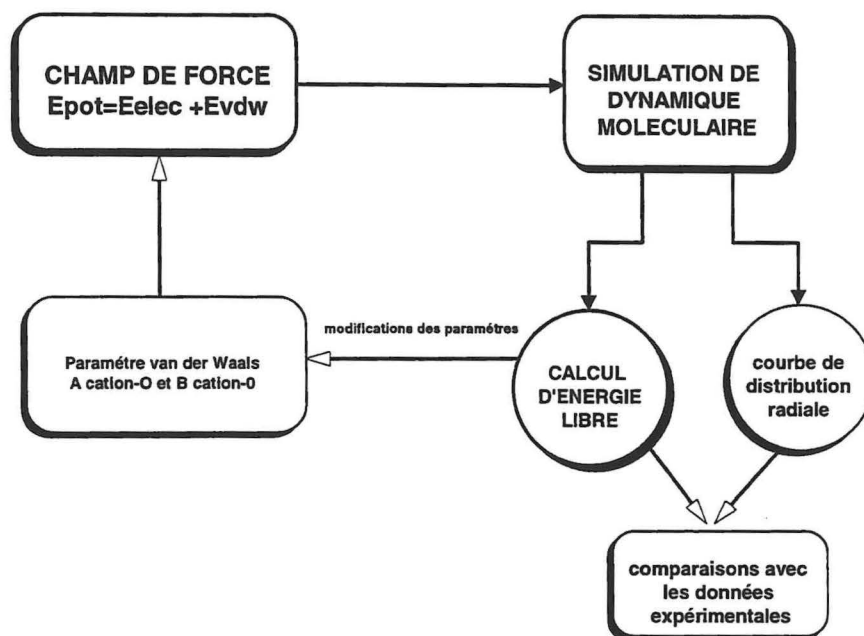


Figure I-2 : la méthode de paramétrisation développée par Åqvist pour l'obtention des paramètres $A_{\text{cation-O}}$ et $B_{\text{cation-O}}$.

Il est à noter que les paramètres obtenus par Åqvist avec cette approche purement empirique, donnent une énergie d'interaction cation molécule d'eau proche de celle obtenue à partir de calcul *ab initio*, notamment en ce qui concerne les cations monovalents. Par ailleurs, ces paramètres ont été obtenus avec des conditions de solvation autres que celles que nous avons choisies, à savoir qu'initialement les simulations d'Åqvist utilisaient le modèle d'eau SPC (5) dans une sphère SCAAS² (6). Néanmoins, Åqvist a montré que SPC et TIP3P donnent des résultats comparables. Par exemple, le calcul de la différence d'énergie libre de solvation entre Na⁺ et K⁺ selon plusieurs conditions de solvation - conditions périodiques et sphère SCAAS avec modèle d'eau SPC ou TIP3P- donne des résultats très similaires (7).

² SCAAS est un modèle sphérique dans lequel la frontière est constituée d'une couche d'eau maintenue par des contraintes harmoniques angulaires et radiales. La contrainte radiale impose une densité de l'eau appropriée, alors que la contrainte angulaire, qui est appliquée au vecteur moment dipolaire, force le système à avoir une polarisation moyenne correspondant à celle attendue pour un système de taille infinie.

Sur un plan pratique, les paramètres d'Åqvist ont été déterminés avec le programme MOLARIS (8). Ce dernier utilise, concernant les paramètres Lennard Jones, des règles légèrement différentes de celles utilisées par le programme CHARMM (9). En effet pour une paire d'atomes i j donnée, les paramètres A_{ij} et B_{ij} associés au potentiel Lennard Jones, sont calculés de la manière suivante :

$$\begin{aligned} A_{ij} &= A_i \cdot A_j \\ B_{ij} &= B_i \cdot B_j \end{aligned} \quad \text{Equation I-1}$$

on peut montrer que, ce mode de calcul des paramètres A_{ij} et B_{ij} correspond aux règles de combinaison suivantes:

$$E_{\min}(ij) = \sqrt{E_{\min}(i) \times E_{\min}(j)} \quad \text{Equation I-2}$$

$$\bar{R}_{\min}(ij) = \sqrt{R_{\min}(i) \cdot R_{\min}(j)} \quad \text{Equation I-3}$$

A la différence du programme CHARMM, qui calcule $\bar{R}_{\min}(ij)$ selon une moyenne arithmétique (règle de Berthelot Lorentz), le programme MOLARIS calcule donc $\bar{R}_{\min}(ij)$ comme une moyenne géométrique des R_{\min} relatifs aux types des atomes i et j . Pour transférer des paramètres d'un programme à l'autre, il faut donc extraire $\bar{R}_{\min}(ion-ow)$ à partir de A_{ion-ow} et B_{ion-ow} obtenu par Åqvist selon l'équation ci-dessous :

$$\bar{R}_{\min}(ij) = \sqrt[6]{\frac{2 \cdot A_{ij}}{B_{ij}}}$$

puis poser $R_{\min}(ion) = 2 \cdot \bar{R}_{\min}(ion-ow) - R_{\min}(ow)$ de sorte qu'après calcul des paramètres A_{ion-ow} et B_{ion-ow} CHARMM utilise bien les paramètres originaux d'Åqvist. $R_{\min}(ion-ow)$ étant connu le paramètre $E_{\min}(ion)$ peut donc ensuite être extrait à partir de l'équation ci dessous.



$$E_{\min}(ion) = \left(\frac{B_{ij}}{2.(\bar{R}_{\min}(ion-ow))^6} \right)^2 \cdot \frac{1}{E_{\min}(ow)} \quad \text{Equation I-4}$$

Ceci peut paraître trivial, néanmoins si par mégarde on extrait $R_{\min}(ion)$ à partir de paramètres $A_{ion-ion}$ et $B_{ion-ion}$ déterminés selon les Equation I-1 on obtient alors un paramètre $R_{\min}(ion)$, qui après introduction dans CHARMM donnera $R_{\min}(ion-ow)$ calculé par moyenne arithmétique de $R_{\min}(ion)$ et $R_{\min}(ow)$. Cette procédure d'extraction des R_{\min} conduit, comme l'ont malheureusement expérimenté Marrone et Merz pour le Na^+ et le K^+ (7,10,11), à l'utilisation de « pseudo paramètres» Åqvist erronés. Les paramètres Lennard-Jones que nous avons utilisés figurent dans le tableau suivant :

• Paramètres Åqvist

| ATOME | E_{\min} | R_{\min} | charge |
|-------|-----------------|------------|--------|
| Ca | -0,449 kcal/mol | 1,325 Å° | +2 |
| Mg | -0,875 kcal/mol | 0,784 Å° | +2 |

Il faut noter que ne figurent pas dans le tableau ci-dessus de paramètres pour l'hydrogène de l'eau, car dans son étude, Åqvist a négligé cette interaction.

Par ailleurs, il est vrai qu'a priori concernant le choix des paramètres des cations, nous avons le choix entre plusieurs sources: les paramètres distribués avec CHARMM22, ceux développés par Shiratori et Nakagawa(12) pour le calcium à partir de calcul Ab initio, ceux dont nous venons de parler, etc . Cependant, comme on peut le voir dans l'étude de Obst et Bradaczek (13) les paramètres CHARMM donnent lors de simulations des distances moyennes cation-oxygène trop importantes dans l'eau, de même que les paramètres Shiratori pour la distance calcium oxygène. De toutes façons,

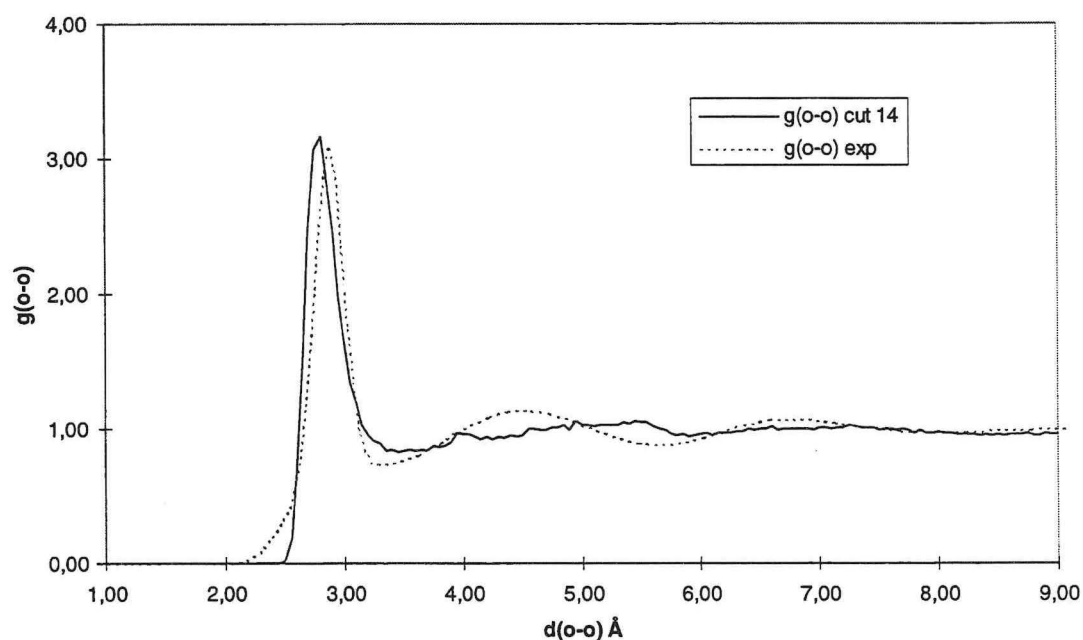


les paramètres d'Åqvist sont les seuls pour lesquels nous soyons sûrs que le calcul de la différence d'énergie libre de solvation donnera un résultat correct, puisqu'ils ont été déterminés à cet effet.

B. Analyse des trajectoires de cations en solution

Dans un premier temps, nous nous sommes intéressés à l'arrangement moyen des atomes du système. Nous avons donc calculé les courbes de distribution radiale³ pour : O-O, Ca-O, et Mg-O à partir des 20 dernières ps de nos simulations de dynamique moléculaire.

1. Comportement du solvant : RDF o-o

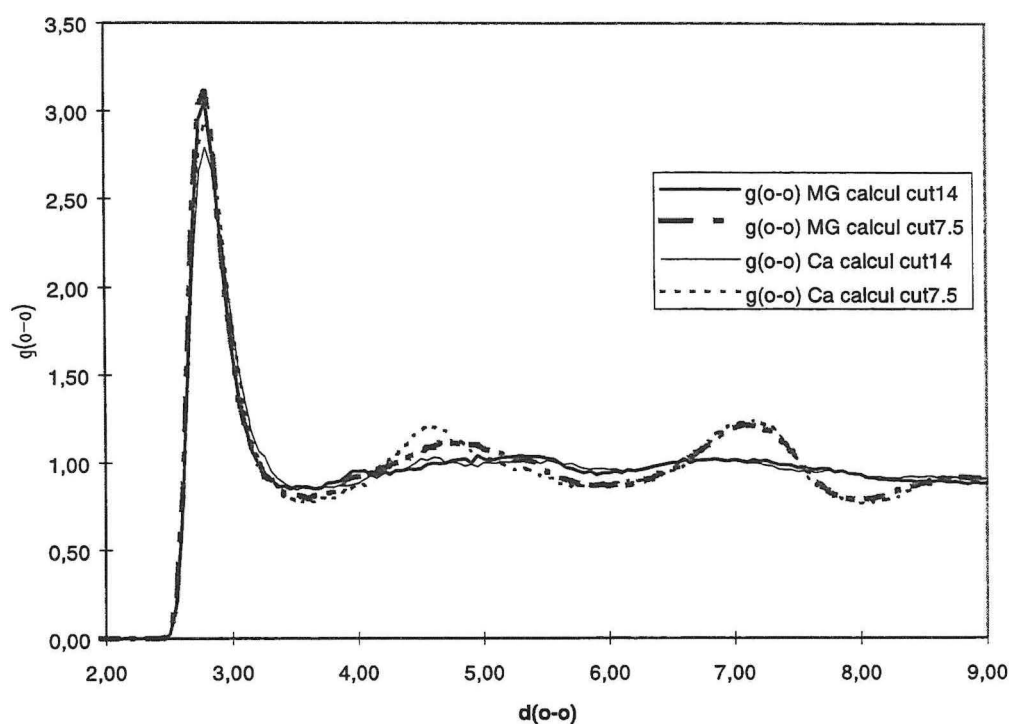


courbes I-1 : comparaison des courbes de distribution radiale oxygène - oxygène expérimentale obtenue par diffraction de rx sur l'eau pure (14) avec celle calculée lors d'une simulation d'un Mg²⁺ en solution, réalisée à cutoff 14 Å.

³ On les note souvent RDF ,pour en anglais , radial distribution function.

Comme lors des simulations initiales de Jorgensen mettant en œuvre une boîte d'eau TIP3P en conditions périodiques (3), dans toutes nos trajectoires⁴, lors de la comparaison des courbes de distributions radiales O-O expérimentales et théoriques (cf. courbes I-1), on observe une quasi superposition des premiers pics. L'absence de deuxième pic, qui est caractéristique du modèle TIP3P(3) n'est, elle, constatée que lors des simulations à cutoff 14 Å. Comme on peut le voir sur les courbes I-2, des 2^{ième} et 3^{ième} pics sont présents seulement dans les simulations réalisées à cutoff 7.5 et ceci aussi bien pour la simulation mettant en œuvre le calcium que celle mettant en œuvre le magnésium. Le deuxième pic de la courbe issue de la simulation avec calcium a une intensité légèrement plus importante. La présence d'un deuxième pic pour TIP3P a été observée précédemment lors d'une simulation d'une protéine en solution(15). Néanmoins, dans notre cas, la présence de ces pics étant corrélée à la valeur du cutoff, il est possible que l'apparition de ces pics soit un artefact relatif au modèle ou au protocole que nous avons utilisés.

⁴ Pour ne pas surcharger les courbes I-1, seule une rdf o-o relative à nos simulations est superposée à la courbe expérimentale. On peut cependant noter sur les courbes I-2, qui représentent l'ensemble des courbes calculées, que la position du premier pic est invariante.



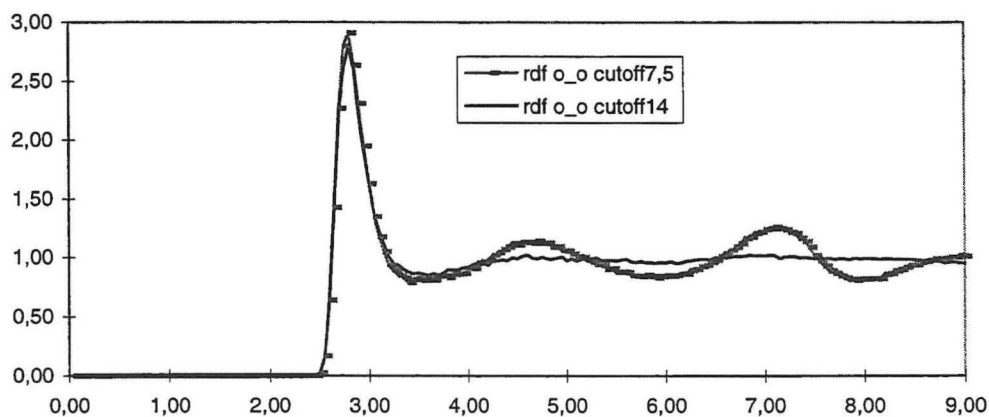
courbes I-2 : Courbes de distribution radiale O-O déterminées à partir des simulations de chaque cation, effectuées à cutoff 7.5 et 14.

L'origine de ces pics n'est donc pas claire. On pourrait suspecter que leur émergence est imputable aux cations, mais alors pourquoi seraient ils absents lors des simulations à grand cutoff. A priori, plusieurs hypothèses semblent plausibles :

- i) La réponse réside peut-être dans la durée de la simulation. En effet, comme le soulignent par exemple Sanejouand et Tapia (16) l'augmentation du cutoff induit des temps de relaxation plus longs. Ainsi l'apparition des deuxième et troisième pics nécessiterait des durées de simulation plus importantes à cutoff 14 qu'à cutoff 7.5 Å. Nos simulations s'effectuant à des durées identiques et relativement courtes, il est donc possible que les pics n'aient pas le temps d'émerger dans les simulations à cutoff 14 Å. Il conviendrait donc d'effectuer des simulations plus longues pour tester cette hypothèse.

- ii) un autre type d'explication : cela pourrait venir de l'emploi d'une fonction shift pour le lissage des interactions non liées.
- iii) Enfin, il est possible que le traitement des frontières de notre système soit à l'origine du phénomène. En effet, les molécules aux limites du système étant partiellement ou totalement figées, on peut imaginer que les molécules d'eau libres les plus proches des frontières, qui sont donc au contact de molécules d'eau beaucoup moins mobiles, ont tendance à se structurer davantage que celles situées plus au centre. Cette structuration peut alors être exacerbée par un petit cutoff. Car contrairement au grand cutoff, ce dernier limite en grande partie le paysage vu par ces molécules. Dans le cadre de cette explication, la présence de ces deux pics, serait un signe d'une structuration plus grande du solvant.

Afin de tester l'hypothèse (i) nous avons effectué des simulations « témoins » de solvant pur à cutoff 14 et 7.5 Å. Chacune de ces simulations a une durée totale de 70 ps, donc plus importante que les précédentes. Les RDF o-o ont été calculées sur les 20 dernières picosecondes . Ces dernières sont représentées ci-dessous :



courbes I-3 : rdf o-o réalisées à partir des 20 dernières ps de simulations de 70 ps sans cation et avec alternativement un cutoff 7.5 et 14Å .

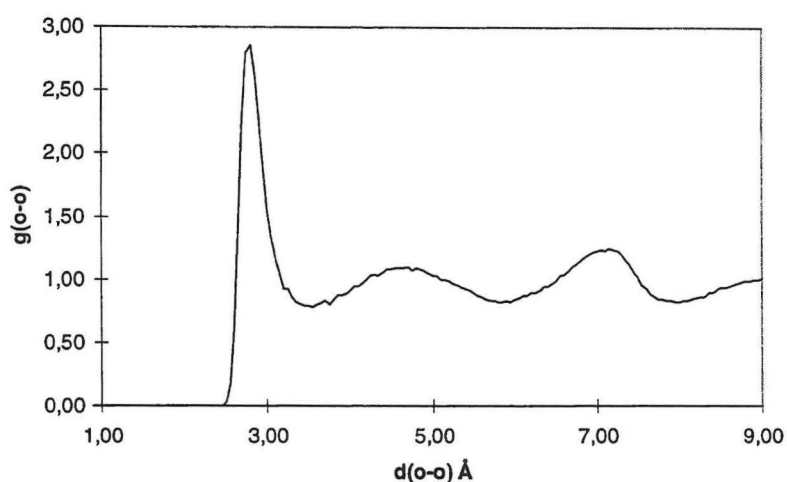
Comme on peut le constater sur les courbes I-3, les 2^{ème} et 3^{ème} pics sont présents uniquement dans la courbe relative à la simulation réalisée à cutoff 7.5 Å. On peut donc conclure que l'émergence de ces pics n'est pas imputable à la présence des cations au centre du système. Par ailleurs la durée plus importante de ces nouvelles simulations indique que l'absence de ces pics ne doit pas non plus être due à une augmentation du temps de relaxation lorsque le cutoff augmente .

De fait, ces pics ne peuvent être dus qu'à d'autres éléments du protocole. Il reste donc à les imputer au traitement de nos frontières ou à l'emploi de la fonction de shift.

Les fonctions de Shift semblent hors de cause et ce sur la base des faits suivants: une courbe rdf (o-o) réalisée à partir d'une simulation d'une goutte d'eau TIP3P avec un cutoff 7.5 Å ne présente pas de deuxième pic (15). Cette goutte d'eau ayant environ la même taille que notre système, soit 463 molécules d'eau, on peut donc considérer que cette simulation met en œuvre un système homologue au nôtre au traitement des frontières près (toute forme de contrainte est alors absente des frontières). Il subsistait cependant une ambiguïté dans la mesure ou les fonctions de lissages utilisées lors de cette simulation étaient de type shift pour l'électrostatique et switch pour le van der waals.

Nous avons donc réalisé une ultime simulation afin de vérifier que le deuxième pic, dans nos simulations, n'est pas imputable au lissage du van der Waals avec une fonction shift. Nous avons effectué 70 ps avec notre sphère d'eau pure, un cutoff 7.5 et un switch du van der waals entre 6.5 et 7.5 Å. La courbe rdf o-o tirée de cette simulation figure ci dessous.





courbes I-4 : rdf O-O relative à une simulation cutoff 7.5 Å. L'électrostatique est lissée avec une fonction shift et le van der Waals est lissé avec une fonction switch.

Dans cette courbe les deuxièmes et troisième pics subsistent encore. On peut donc en conclure que c'est la présence de frontières semi-rigides, conjuguée à l'utilisation d'un cutoff 7.5 qui est à l'origine du phénomène observé. Néanmoins, ces mêmes frontières altèrent peu le comportement du solvant lors de l'utilisation d'un cutoff de 14 Å, puisque comme nous l'avons vu, les courbes rdf o-o ont alors un profil très similaire à celui obtenu en conditions périodiques⁵.

2. Etude structurale des couches de solvation des cations

a) bilan des données expérimentales

(1) Première couche de solvation

La solvation des cations dans l'eau a fait l'objet de nombreuses études expérimentales. Ces travaux ont porté à la fois sur la distance cation oxygène, le nombre de coordination et la géométrie de la première couche. Des revues de Marcus(18), ainsi

⁵ La simulation Monte Carlo de Jorgensen utilisait un cutoff de 7.5 Å (3). Prevost et al utilisent un cutoff de 9 Å lors d'une trajectoire de dynamique moléculaire avec lissage des interactions non liées par une fonction shift(17).

que de Othaki et Radnai (19), rassemblent ces données issues de techniques très diverses (diffraction de rayons X ou de neutrons, spectroscopie Raman, RMN ...).

Ainsi, pour le calcium, on donne une distance cation premier oxygène⁶ de $2,43 \pm 0,09$ Å (19). Bien que cette distance moyenne soit assez bien résolue, le nombre de plus proches voisins auquel elle se réfère est relativement flou. En effet, on relève selon la référence (19) un nombre de coordination expérimental variant entre 6 et 10 molécules d'eau. Compte tenu de ces résultats, il s'en suit que la géométrie de la première couche de solvation du calcium est plutôt mal connue.

Les données relatives au magnésium sont plus précises. En effet, on trouve un nombre de coordination unique égal à 6 avec une distance moyenne Mg-O de $2,09 \pm 0,05$ Å, le tout formant un édifice à symétrie octaédrique.

Hormis la distance et le nombre de voisins, les sources expérimentales nous fournissent une autre information relative à l'orientation des molécules d'eau dans la première couche de solvation des cations. Cette orientation est déterminée à partir de la courbe de distribution radiale ion-atome de solvant. En effet, connaissant la géométrie de la molécule d'eau⁷, ainsi que les distances cation-O et cation-H issues des premiers pics de cette distribution, par un raisonnement trigonométrique simple, on peut déterminer la distance d_1 , puis l'angle θ représentés dans la Figure I-3.

⁶ Cette distance est la moyenne sur toutes les sources expérimentales. Il s'en suit que son incertitude reflète la dispersion des données publiées dans la référence (19) et non pas l'incertitude relative à chaque technique expérimentale.

⁷ Dans notre cas, il s'agit des angles HOH = $104,5^\circ$ et de la longueur de liaison OH = $0,9572$ Å du modèle TIP3P

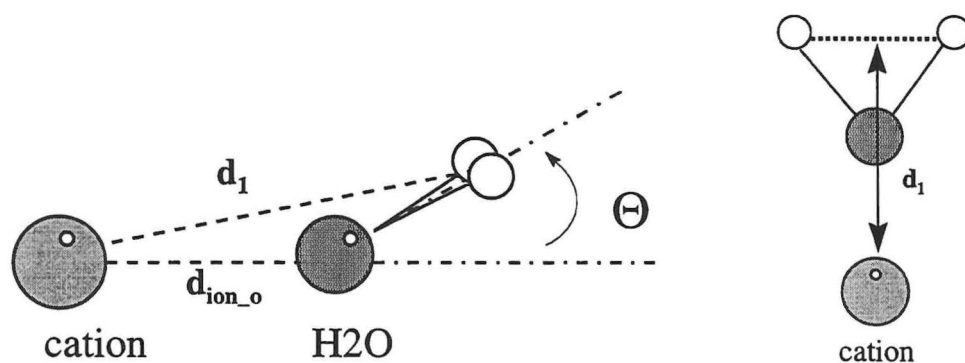


Figure I-3

Cet angle correspond à l'inclinaison des plans des molécules d'eau de la première couche par rapport à l'angle que fait le plan de la molécule d'eau avec la droite joignant le cation à l'oxygène. Il a été initialement introduit par Enderby et Neilson (20) pour l'interprétation de données neutroniques. La référence (19) donne des valeurs expérimentales pour cet angle de 14° pour Mg^{2+} et comprise, en fonction de la concentration, entre 26 et 43° pour Ca^{2+} .

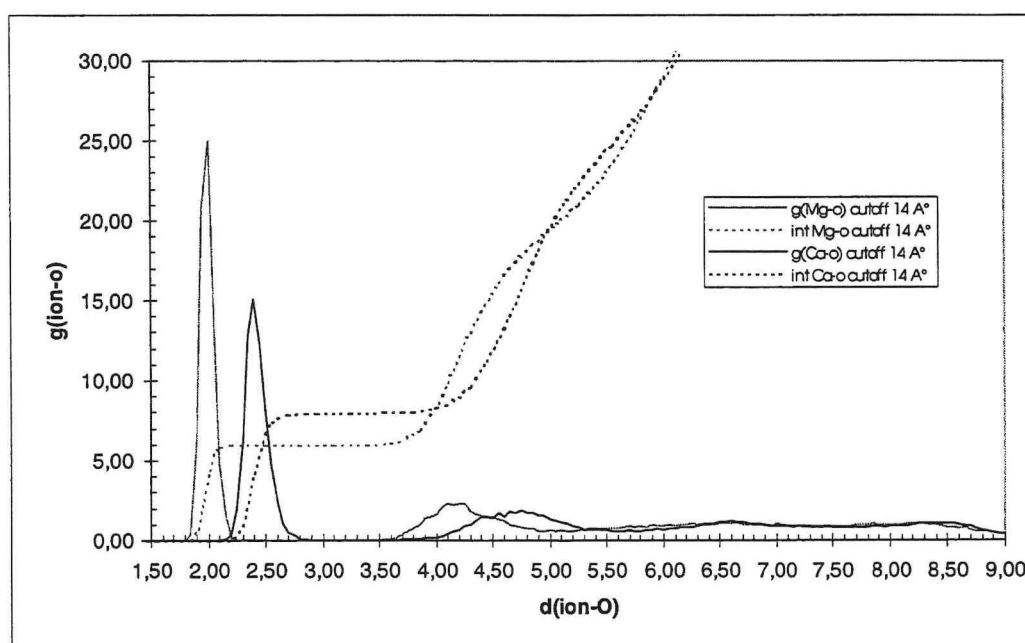
(2) Deuxième couche de solvation

Comme le soulignent Othaki et Radnai (19) dans leur revue, les informations disponibles pour la 2^{ème} couche de solvation sont plus pauvres que celles disponibles pour la 1^{ère}, car la deuxième couche est beaucoup moins structurée du fait que ses molécules d'eau constitutives sont plus nombreuses et surtout plus mobiles. De manière générale, sur le plan expérimental, la résolution des données est d'autant meilleure que le cation est petit et/ou chargé. Il s'en suit que les résultats relatifs au Mg^{2+} sont là encore mieux résolus que ceux du Ca^{2+} . En effet, on donne, en moyenne, 12 molécules d'eau à une distance de $4.21 \pm 0.06 \text{ \AA}$ pour Mg^{2+} et une à 6 molécules d'eau à une distance de $4.55 \pm 0.05 \text{ \AA}$ pour Ca^{2+} (19). Il va de soi que, comme le soulignent Othaki et Radnai, le nombre de molécules d'eau constitutives de la 2^{ième} couche du Ca^{2+} est rationnellement difficile à expliquer, si ce n'est par l'imprécision des données expérimentales.

b) rdf cation-atomes de solvant obtenues à cutoff 14 Å

Afin de privilégier la clarté nous avons fait le choix de traiter indépendamment les atomes d'oxygène et les atomes d'hydrogène constitutifs du solvant.

(1) rdf cation oxygène



courbes I-5 : rdf cation-oxygène déterminées à partir des simulations réalisées avec un cutoff de 14 Å

On peut observer sur les courbes I-5 un premier pic bien résolu, donnant une distance moyenne cation-oxygène de 2.00 ± 0.05 Å et 2.40 ± 0.05 Å respectivement pour Mg^{2+} et Ca^{2+} . Comme l'on pouvait s'y attendre compte tenu du mode de détermination des paramètres Åqvist(4), les positions de ces pics sont en bon accord avec l'expérience. En effet, alors que nos résultats sont identiques à ceux publiés par Åqvist, on note respectivement pour Ca^{2+} et Mg^{2+} des différences de 1 et 5% avec les distances cation-O expérimentales.

Ces résultats confirment notamment que le mode de solvation et le modèle d'eau que nous avons utilisés donnent des résultats comparables au modèle d'eau SPC avec les conditions de solvation SCAAS utilisées par Åqvist.

Nous avons approfondi l'analyse structurale des couches de solvation du calcium et du magnésium. Comme l'objectif d'Åqvist était la détermination des paramètres Lennard Jones, il n'a pas fait d'analyse structurale autre que celle relative aux distances des premiers voisins. Ces investigations sont d'autant plus intéressantes que l'aspect structurale de la première couche est plutôt mal connu pour Ca^{2+} . Ainsi à partir des rdf cation-O (cf. courbes I-5), on peut constater les éléments suivant :

- i) De manière générale, les molécules d'eau s'organisent dans le champ électrostatique du cation pour former des couches radiales autour de celui-ci. Cette organisation se traduit par la présence de 2 pics dans nos courbes rdf cation-o.
- ii) Les nombres de coordination obtenus par intégration du premier pic rdf ion-o, sont respectivement de 8 et 6 molécules d'eau pour Ca^{2+} et Mg^{2+} .
- iii) Pour les deux cations, la courbe passe par zéro après le premier pic, ce qui traduit le fait que pendant la durée de nos simulations il n'y a pas d'échange entre 1^{ère} et 2^{ème} couches.
- iv) De manière générale les positions des deuxièmes couches sont en assez bon accord avec l'expérience : on note des différence de l'ordre de 2 à 3%. On constate, pour les deux cations, que la deuxième couche est plus diffuse. Elle est située plus près, entre 4.10 et 4.15 Å, avec près de 12 molécules d'eau pour Mg^{2+} alors que pour Ca^{2+} , elle est plus étalée et donc plus difficile à délimiter précisément. Néanmoins, elle atteint un maximum à $4.70 \pm 0.05 \text{Å}$ et décroît lentement jusque 5.50 Å. Son profil étant moins marqué, il s'en suit aussi que le nombre de molécules d'eau relatif à cette couche est plus difficile à déterminer que dans le cas du Mg^{2+} . On mesure toutefois une aire de la courbe Rdf Ca-O

entre 0 et 5.5 Å égale à 26, ce qui après retranchement du nombre de premiers voisin, nous donne 18 molécules d'eau constitutives de la deuxième couche.

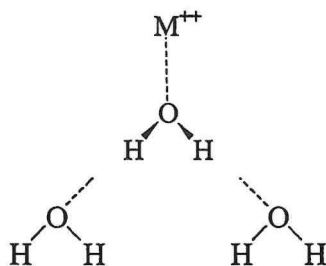


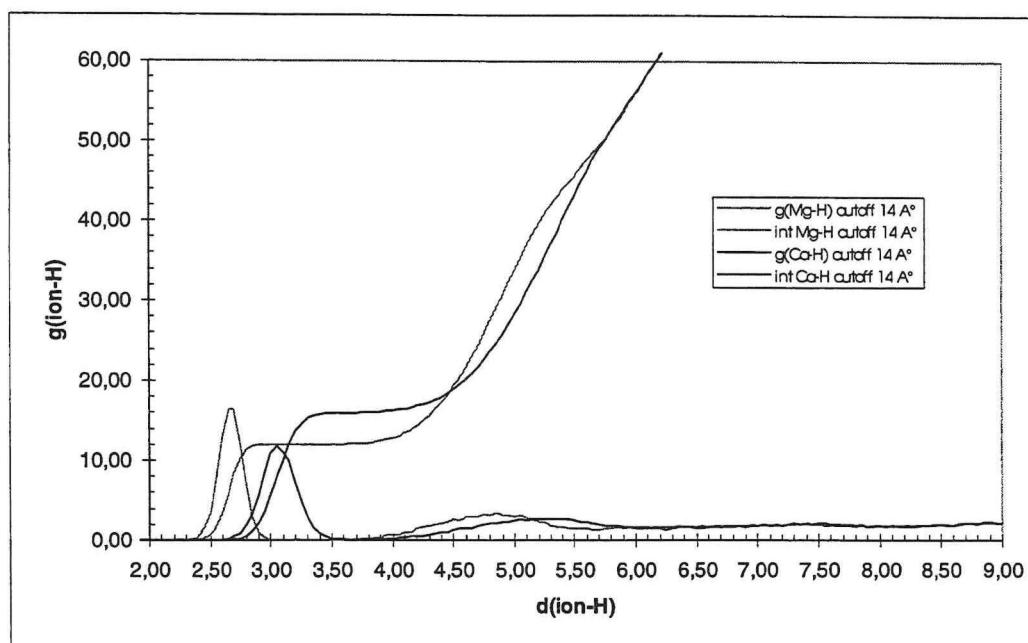
Figure I-4

Ce nombre de secondes voisines, est le double de celui des molécules d'eau présentes en première couche. Cette constatation est aussi valable pour Mg^{2+} , ce qui suggère l'organisation représentée Figure I-4, dans laquelle chaque molécule d'eau de la première couche forme une liaison hydrogène avec deux molécules d'eau appartenant à la seconde couche.

(2) rdf cation-hydrogène

Ce type de courbe nous donnera des informations complémentaires de celles obtenues avec les courbes rdf cation-O. En particulier, elles permettent la détermination de la distance moyenne cation-hydrogène de la première couche, ce qui permet la détermination de l'angle θ relatif à l'orientation des molécules d'eau dans cette première couche. Par ailleurs, pour plusieurs aspects, elles donnent des informations redondantes avec les résultats obtenus via les rdf cation-O.





courbes I-6 : rdf cation - hydrogène déterminées à partir des simulations réalisées à cutoff 14 Å.

Concernant les premières couches de solvation, comme on peut le voir sur les courbes ci-dessus, nous obtenons des premiers pics à respectivement $2,65 \pm 0,05$ et $3,05 \pm 0,05$ Å pour Mg^{2+} et Ca^{2+} . De manière générale, à un pic dans la courbe rdf cation-O on peut faire correspondre un pic dans la rdf cation-H. Les pics rdf cation-H sont décalés vers de plus grandes distances par rapport à leurs homologues des rdf cation-O. Ce qui traduit l'orientation spécifique de la molécule d'eau dans le champ du cation. De plus, l'intégration des premiers pics donne un nombre d'hydrogènes double de celui du nombre d'oxygènes observés précédemment dans les courbes rdf cation-o. Par ailleurs, la base des pics rdf cation-H est plus large que celle de leur homologue dans les rdf cation-o, ce qui traduit une mobilité orientationnelle des hydrogènes .

c) orientation des molécules d'eau dans la première couche

A partir de la géométrie d'une molécule d'eau TIP3P et des distances cation-oxygène et cation-hydrogène de la première couche, nous avons déterminé les valeurs des angles θ relatifs à l'orientation des molécules d'eau dans cette première couche. Les valeurs de ces angles figurent dans le tableau ci-dessous :

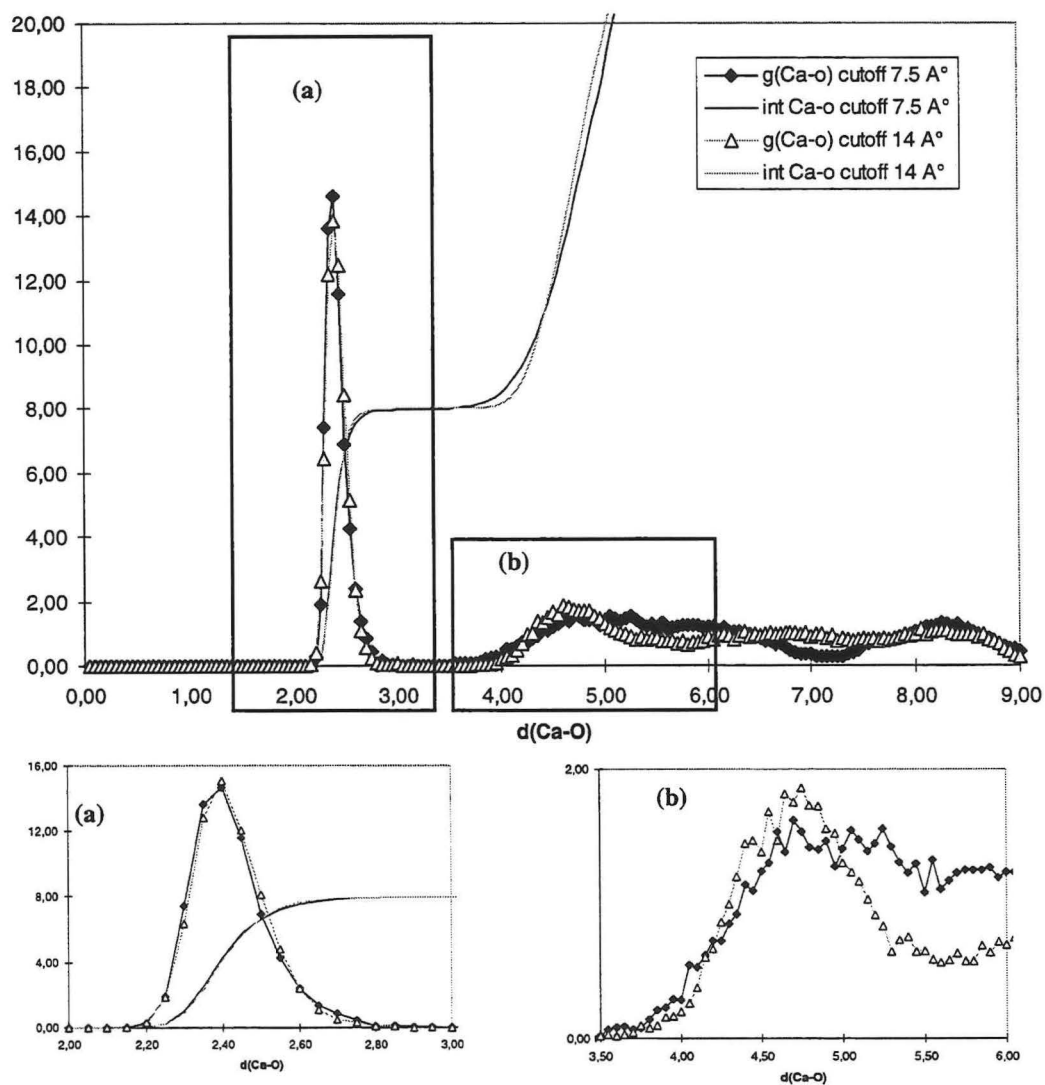
| | θ en degrés |
|------------------|--------------------|
| Ca ²⁺ | $21 \pm 10^\circ$ |
| Mg ²⁺ | $26 \pm 10^\circ$ |

Le calcul de l'angle θ est très sensible aux variations des distances cation-oxygènes et cation-hydrogènes. Comme nous avons déterminé ces grandeurs à 0.05 \AA près, il s'en suit que l'imprécision sur θ est de l'ordre de 10° . On peut donc considérer, à ce degré de précision, que les molécules adoptent une orientation à peu près identique dans les premières couches de solvation du calcium et du magnésium. Les angles θ , ont des valeurs du même ordre de grandeur que celles obtenues expérimentalement .

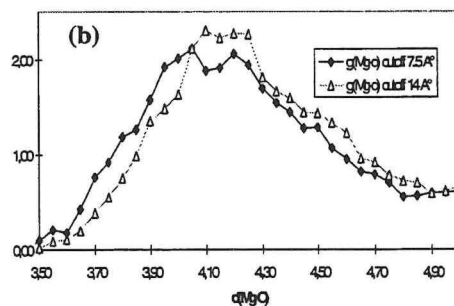
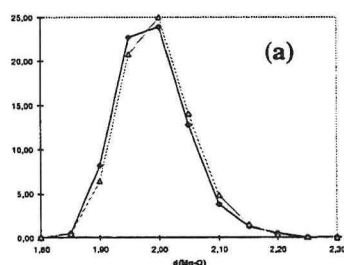
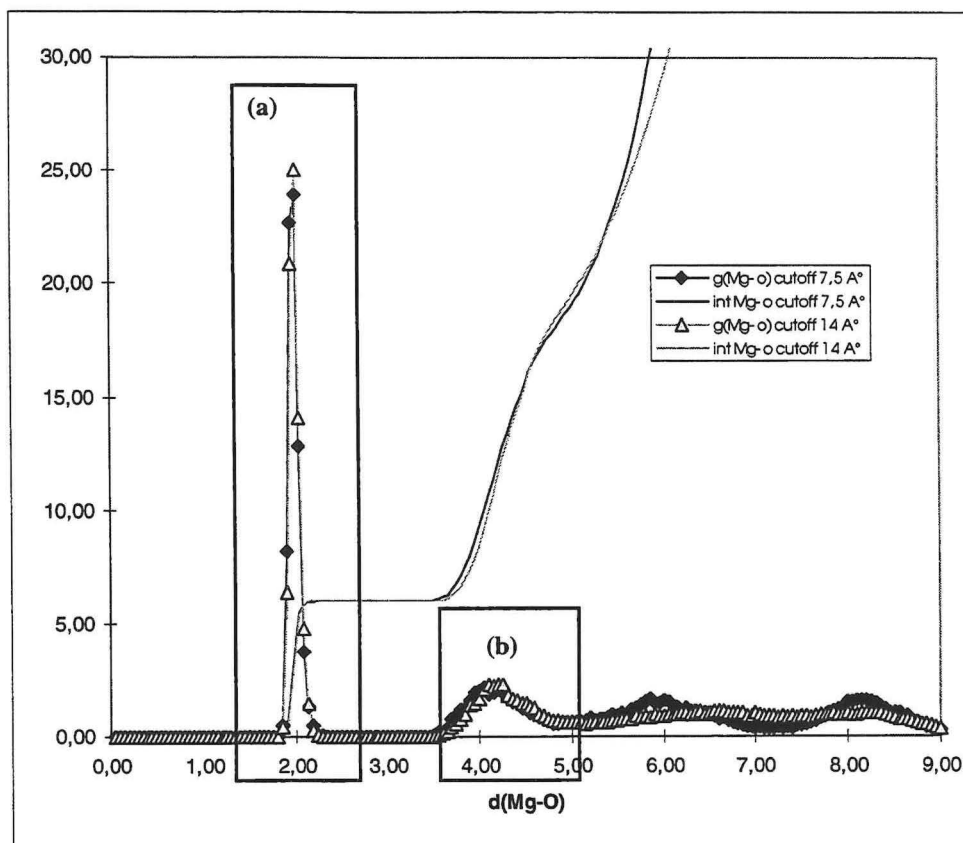


3. Influence du cutoff

Etudions maintenant l'influence du cutoff sur la structuration des molécules d'eau autour des cations. Pour ce faire, de manière analogue à notre étude du solvant, nous sommes attachés à relever les différences entre les courbes RDF ion-O aux deux valeurs de cutoff 7.5 et 14 Å. Ces courbes sont représentées ci-dessous.



courbes I-7 : comparaison des rdf Calcium-o à cutoff 7.5 et 14 Å. Les fenêtres (a) et (b) sont des agrandissements relatifs aux premiers et deuxièmes pics de ces courbes.



courbes I-8 : comparaison des rdf magnésium-o à cutoff 7.5 et 14 Å. Les fenêtres (a) et (b) sont des agrandissements relatifs aux premiers et deuxième pics de ces courbes.

Regardons tout d'abord les deux premières couches de solvation des cations. Sur les courbes I-7 (a) et courbes I-8 (a), on observe, indépendamment du cutoff, une structuration quasi identique du solvant en première couche pour les deux ions. Une différence de structuration significative, intervient essentiellement à partir de la deuxième couche. Concernant cette deuxième couche, pour le calcium, on peut voir (cf. courbes I-7 (b)) un profil plus marqué de la deuxième couche à cutoff 14 Å. Alors que la

distance au calcium reste quasi identique, il semble qu'il y ait une mobilité plus importante des molécules d'eau de seconde couche à cutoff 7.5Å.

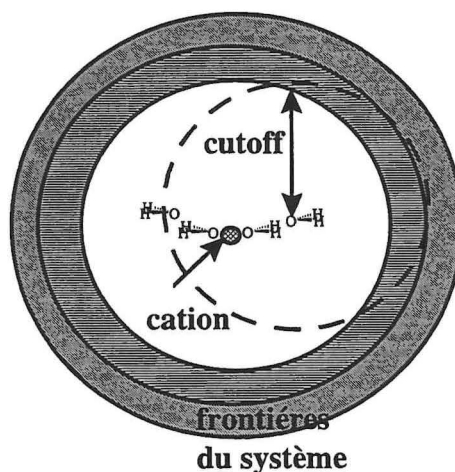
Concernant la deuxième couche du Mg²⁺ on peut remarquer que, en terme de structuration, l'influence du cutoff est moins perceptible que pour le calcium. Cependant on peut voir dans la fenêtre (b) des courbes I-8, que alors que l'intensité de cette couche reste à peu près constante, il y a une légère translation de 0.1 Å de la couche vers Mg²⁺. Ce phénomène semble traduire une légère structuration de la deuxième couche par rapport à la simulation effectuée à cutoff 14 Å. Nous avons donc le phénomène inverse de celui observé pour le calcium. Cette deuxième couche implique 12 molécules d'eau, soit le même nombre pour les deux valeurs du cutoff.

Regardons maintenant au delà des deux premières couches. Les rdf cation-oxygène obtenues à partir des simulations réalisées avec un cutoff 7.5 présentent des pics supplémentaires. Dans le cas du magnésium, on note la présence d'un 3^{ième} et d'un 4^{ième} pic à respectivement 6 et 8.2 Å du cation. Ces derniers sont certes beaucoup plus diffus que le deuxième pic, mais tout à fait discernables. Dans le cas du calcium, au delà des deux premiers pics, on ne discerne plus qu'un seul pic à environ 8.2 Å. Ce dernier est donc à la même position que le quatrième pic du magnésium. Il est à proximité immédiate des frontières du système. Contrairement à la courbe du magnésium on ne distingue pas de pic à 6.0 Å. Il semble que suite au décalage des deux premiers pics⁸, le deuxième pic proprement dit et ce qui était le troisième pic dans la simulation du Mg²⁺ aient fusionné pour donner la forme extrêmement diffuse du deuxième pic de la rdf Ca-O.

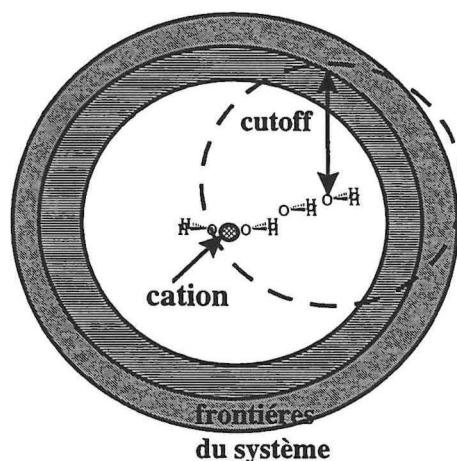
⁸ Compte tenu du rayon plus important du calcium, la position des deux premiers pics est décalée de 0.4 Å et 0.5 Å pour le premier et le second par rapport à leurs homologues dans la courbe relative au magnésium.

Il semble donc que l'origine de ces pics soit imputable aux effets conjugués du faible cutoff (7.5 Å) et de la proximité des frontières du système. Nous allons discuter cette hypothèse. Notons pour le moment que pour les deux cations, les pics artificiels disparaissent dans les courbes réalisées à grand cutoff.

Essayons maintenant d'interpréter ces résultats. Pour ce faire considérons la position des 1^{ère} et 2^{ème} couches suivant la taille du cutoff et celle du système. Comme on peut le voir sur les figures ci-après,



lors des simulations effectuées à cutoff 7.5, les molécules d'eau en deuxième couche voient partiellement les molécules de leur propre couche situées à l'opposé. Ce fait est accentué dans le cas du calcium. Par contre, elles interagissent pleinement avec les molécules d'eau de leur hémisphère coincées entre elles et leur congénères appartenant aux zones frontières.



Enfin les molécules d'eau situées au-delà de 7.5 Å du cation n'interagissent pas avec tous les atomes situés dans l'hémisphère opposé. Par contre elles voient pleinement les molécules appartenant aux frontières, qui doivent avoir une part importante dans le nombre des interactions prises en compte par ces molécules.

De manière générale, sont présentes dans notre système deux « entités structurantes », d'une part le cation qui grâce au champ électrostatique qu'il génère, oriente les molécules d'eau à relativement grande distance et, d'autre part les frontières constituées d'eau statique ou peu mobile. A la lumière des éléments exposés ci-dessus, il semble que nous soyons en mesure de clarifier les raisons pour lesquelles l'emploi d'un cutoff 7.5 pourrait être à l'origine d'une modification du comportement dynamique des molécules d'eau situées à moyennes et grandes distances des cations.

i) Le comportement de la première couche des cations est peu sensible au changement de cutoff sans doute de par la proximité immédiate de celui-ci. Mais il n'en est pas de même pour les autres couches.

ii) Les molécules à proximité des frontières ne ressentent pas l'influence du cation à cutoff 7.5, alors qu'à cutoff 14, elles subissent l'influence contradictoire des deux éléments structurants. On comprend ainsi la présence de pics à grande distance lors des

simulations à petit cutoff, qui semble traduire une structuration accrue à proximité des frontières du système.

La situation est plus ambiguë pour les molécules d'eau de deuxième couche qui sont placées dans une zone intermédiaire entre les deux agents structurants. Ainsi dans le cas du calcium, les molécules d'eau de deuxième couche sont situées quasi à mi-distance des deux entités structurantes. Cependant, à cutoff 7.5, les molécules d'eau situées plus à la périphérie ne voient pas le cation alors qu'elles ressentent pleinement l'influence des frontières. Dans ces conditions, on peut comprendre le caractère accidenté de la deuxième couche du cation qui semble être composée de molécules qui tantôt s'orientent vers le cation tantôt vers les frontières.

Dans le cas de la deuxième couche du magnésium la situation est différente. Ce cation étant plus petit cette couche est située plus près de celui-ci, et les molécules constitutives de la deuxième couche sont, de manière générale, plus structurées. Ainsi la seconde couche est moins affectée par les frontières que dans le cas du Ca^{2+} , mais comme nous l'avons vu, on assiste à un décalage du deuxième pic de 0.1 Å vers le cation. Ceci pourrait traduire le fait qu'avec un cutoff 7.5 les molécules de deuxième couche ressentent moins la contribution répulsive des molécules d'eau diamétralement opposés constitutives de cette même couche.

4. Géométrie de la première couche de solvation

A ce stade de notre étude les données « macroscopiques » issues des rdf nous ont fourni des informations relatives aux nombres de coordination et distances cation-o. Néanmoins, il est difficile à partir de ces seules informations de se représenter, par

exemple, la géométrie globale de la première couche. Ayant à notre disposition les positions des atomes durant les vingt picosecondes de nos simulations nous avons pu entreprendre l'étude de la géométrie de ces premières couches de solvation.

Dans un premier temps, pour chacun des cations, à partir des ensembles de coordonnées enregistrées lors des trajectoires réalisées à cutoff 14 Å, nous avons déterminé la géométrie moyenne afin d'avoir une idée générale de la structure de l'ensemble. Puis, dans un deuxième temps, nous nous sommes attachés au comportement dynamique de cet ensemble autour de sa géométrie moyenne durant nos simulations.

a) première couche du Magnésium

(1) géométrie moyenne

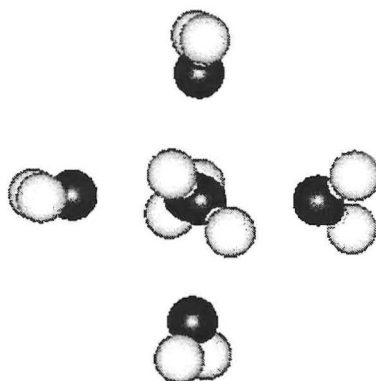


Figure I-5 : géométrie moyenne de la première couche de solvation du magnésium, réalisée sur l'ensemble des coordonnées de la trajectoire à cutoff 14 Å.

Comme on peut le voir dans la Figure I-5, la géométrie moyenne de la première couche de solvation est celle d'un arrangement octaédrique des oxygènes. Par ailleurs, on peut voir que les hydrogènes des molécules d'eau diamétralement opposées adoptent une disposition décalée. L'interaction inter hydrogènes étant défavorable sur le plan électrostatique, cet arrangement permet de maximiser les distances entre hydrogènes.



(2) Structure Dynamique de la première couche

Sur un plan dynamique, cet arrangement s'avère être très rigide.

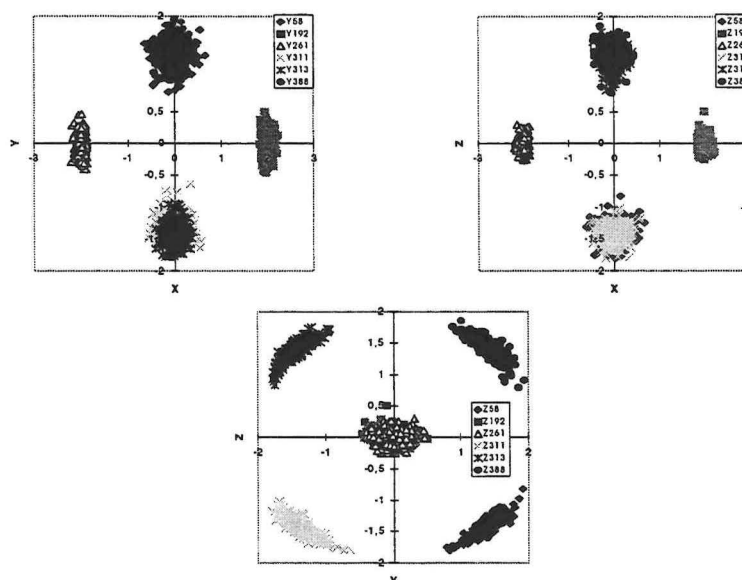


Figure I-6

Projection sur les plans XY, YZ et XZ des coordonnées des molécules d'eau de la première couche du magnésium enregistrées durant la trajectoire à cutoff 14 Å. Le cation est au centre du repère. Pour chaque ensemble de coordonnées, avant projection toute la première couche a été orientée par rapport à un référentiel formé de 4 molécules d'eau .

En effet, si l'on considère les figures ci-dessus on peut constater une faible dispersion des nuages de points relatifs à chacune des molécules d'eau. De plus, l'ensemble de la structure s'écarte peu de l'arrangement octaédrique durant toute la trajectoire. Ainsi, bien que la nature des interactions soit de type non covalent, il semble que l'ensemble magnésium première couche de solvation se comporte comme une structure rigide unique assimilable à une molécule.

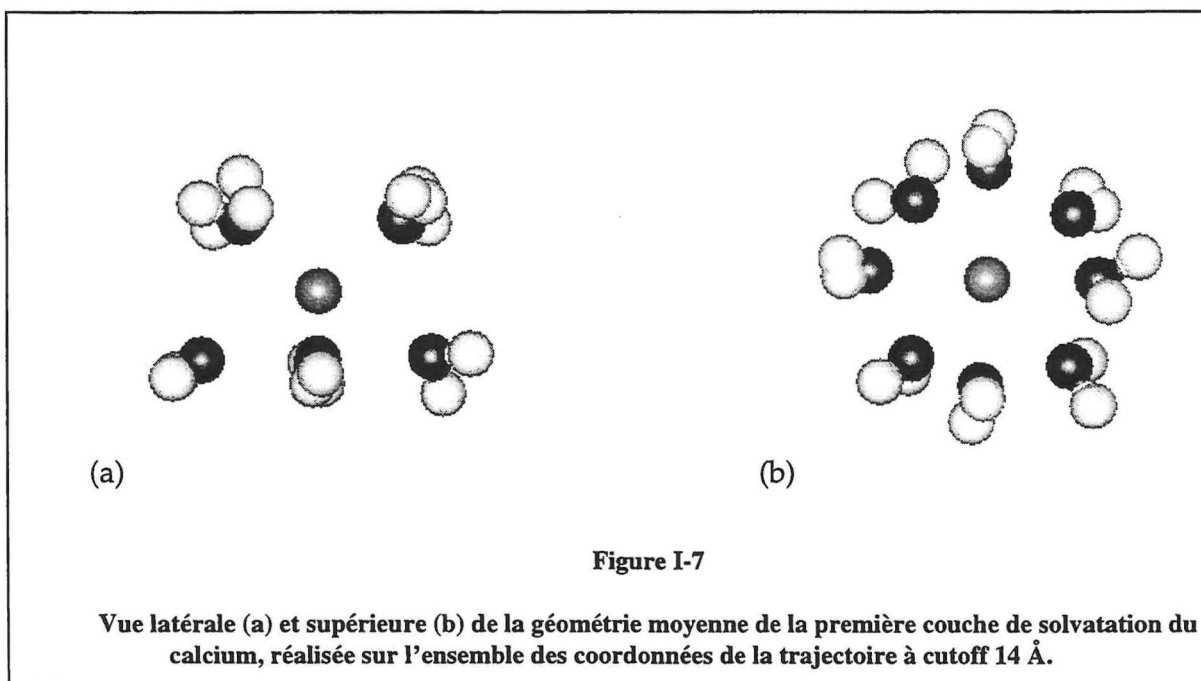
Si l'on s'attache maintenant à l'orientation moyenne des molécules d'eau, nous pouvons à partir des coordonnées moyennes de cette première couche déterminer les angles θ avec une précision beaucoup plus importante. Comme on peut le deviner sur la

Figure I-5 les molécules d'eau et le magnésium semblent être dans un même plan. Cette orientation est confirmée puisque l'angle θ a une valeur moyenne de $1 \pm 0.5^\circ$.

b) première couche du calcium

La structure adoptée par la première couche de solvation du calcium est comme nous l'avons vu expérimentalement bien moins connue que celle du Magnésium. Nous avons mis en évidence que lors de nos simulations il y a 8 molécules d'eau dans cette couche. Considérons maintenant, la géométrie de la première couche de solvation du calcium lors de nos simulations.

(1) géométrie moyenne



La géométrie moyenne de la première couche du calcium est présentée dans la Figure I-7. Sur ces deux vues, on peut constater que les 8 voisins constitutifs de cette couche adoptent une structure de type anti-prisme à base carrée. Il est à noter que de

même que pour le magnésium, les hydrogènes ne sont pas dans des plans parallèles, mais au contraire semblent décalés. Enfin, l'orientation moyenne des molécules d'eau est caractérisée, de même que pour Mg^{2+} , par un angle θ moyen quasi nul de $0.7 \pm 0.3^\circ$. Cet angle est ici déterminé de manière beaucoup plus précise que précédemment. Il est cependant nettement plus petit que les valeurs expérimentales qui sont de l'ordre de 30° .

(2) Structure Dynamique de la première couche

Le nombre de voisins étant plus important, nous n'avons pas utilisé le même mode de présentation que pour Mg^{2+} , la projection sur les plans donnant alors des graphes peu clairs et difficiles à interpréter.

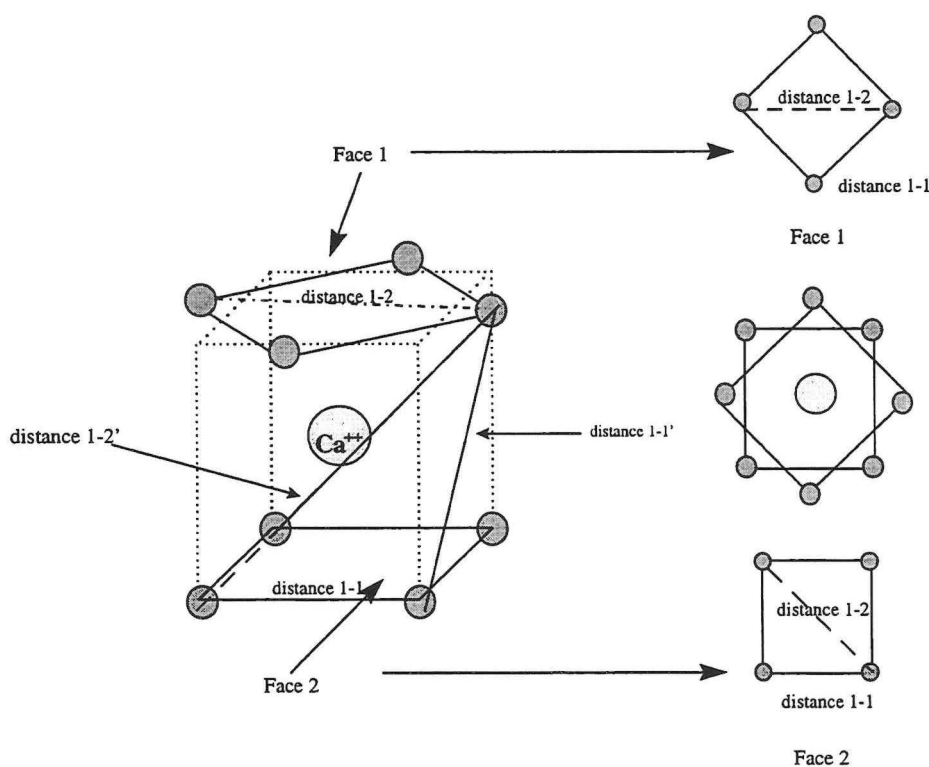
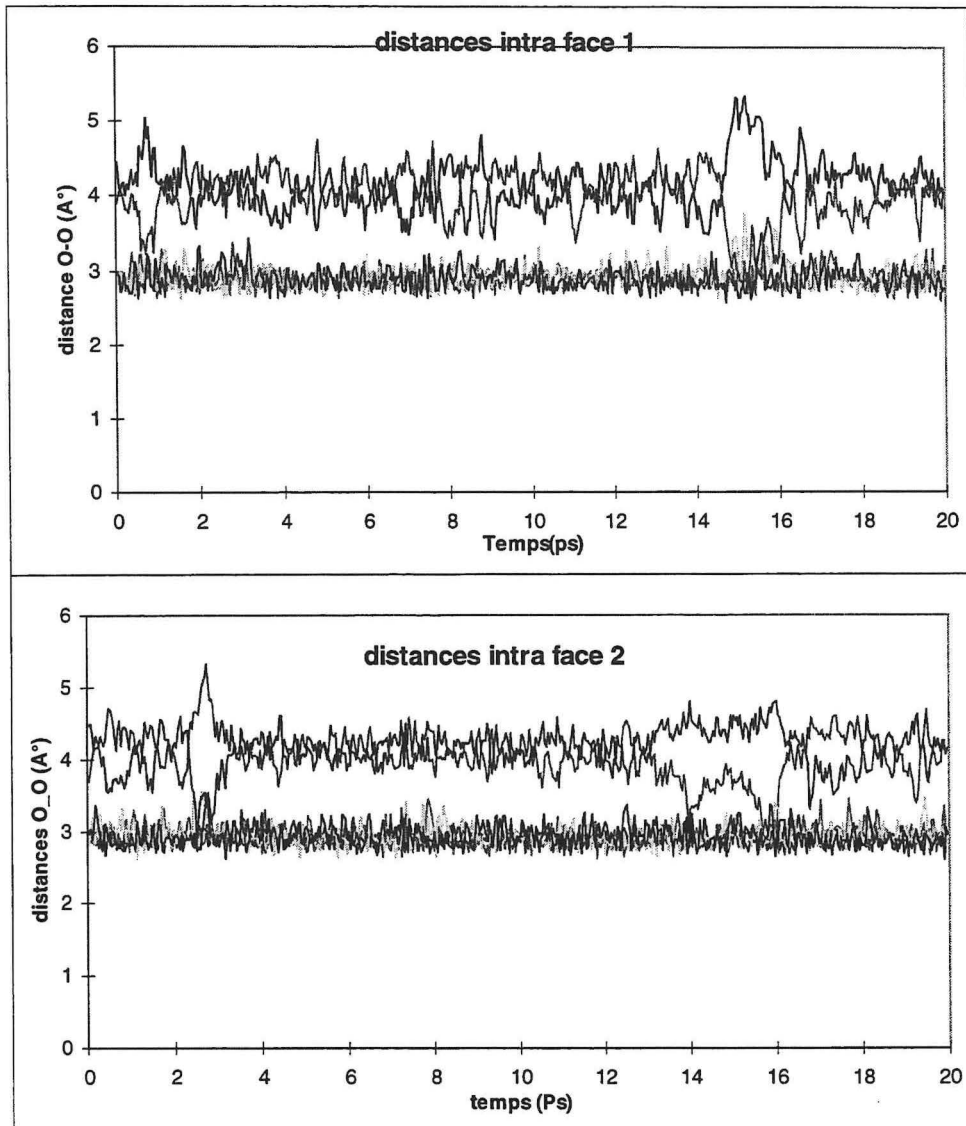


Figure I-8 : représentation de la structure de la première couche de solvation du calcium dans l'eau. Partant de l'hypothèse que durant la dynamique la géométrie de la première couche reste voisine de la géométrie moyenne, nous avons fait le choix de suivre les distances entre oxygènes de cette couche durant la trajectoire. Dans un souci de clarté nous avons séparé les distances relatives à chaque face de celles entre faces de la structure. Dans une

face, à priori, les distances entre oxygènes doivent se distribuer en deux classes. Il en est de même pour les distances entre faces (cf. Figure I-8).

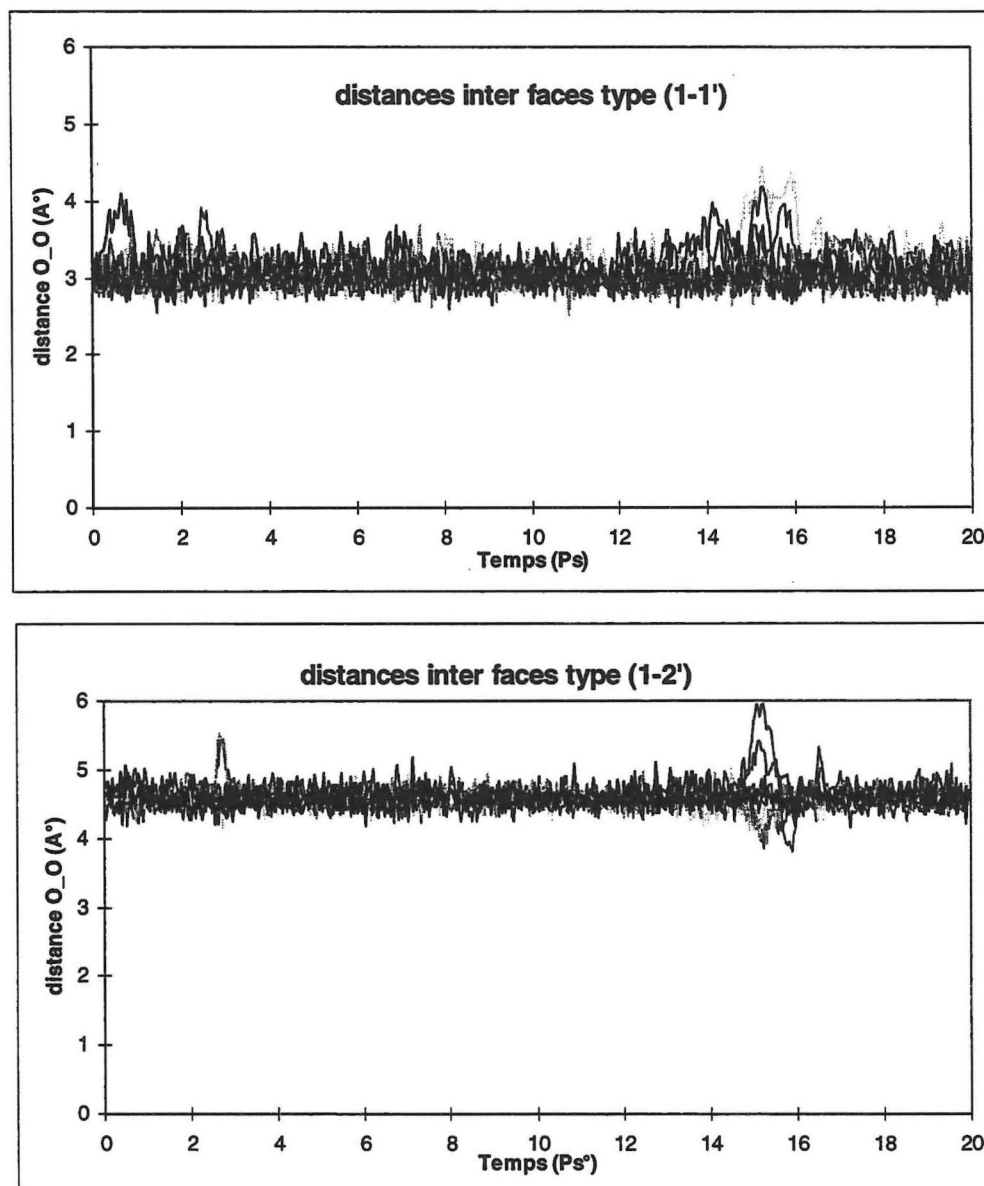


courbes I-9 : distances O-O des oxygènes constitutifs de chacune des faces.

Sur les courbes I-9, on distingue facilement les distances relatives aux cotés de celles, plus longues, relatives aux diagonales des faces . Durant la simulation, ces dernières fluctuent de manière plus importante. En effet, de manière isolée (cf. face1 a environ 1ps et face2 à 3 ps) ou de manière simultanée (vers 15ps), les distances diagonales s'allongent et se raccourcissent de manière concertée au sein d'une face. Les distances relatives aux cotés des faces restent quant à elles à peu près constantes. Ces



phénomènes suggèrent donc un passage de la face d'une forme carrée à une forme losange. Regardons maintenant les mouvements des faces l'une par rapport à l'autre. Le nombre de ces distances inter faces étant plus important nous avons fait le choix de les représenter en deux classes, les distances type (1-1') et les distances (1-2') (cf. Figure I-8).



courbes I-10 : distances O-O des oxygènes appartenant à différentes faces (cf. Figure I-8).

Ces distances de même que les distances 1-1 (intra face), ont un comportement relativement stable. Seules sont à constater quelques elongations sans doute relatives aux transitions observées au sein des faces. Il semble donc que les deux faces, à l'échelle de temps de nos simulations, restent à distance constante et/ou n'opèrent que des



déplacements corrélés. On ne peut notamment pas observer de permutations entre les faces que l'on pourrait qualifier d'échanges type « Rubik's cube ». L'étude de ces derniers nécessiteraient des simulations de durées plus importantes.

C. conclusion

Les premières couches de solvation des cations ont des distances moyennes cation - oxygène en bon accord avec les données expérimentales. Ce fait est peu surprenant si l'on considère le mode d'obtention des paramètres que nous avons utilisés : il ne s'agissait pour nous que de tester nos choix concernant le mode de solvation des cations ainsi que le protocole de nos simulations.

Concernant ce protocole, nous avons mis en évidence la présence d'artefacts liés aux effets conjugués d'un petit cutoff (7.5 Å) et des frontières peu mobiles de notre système. Ces dernières induisent une structuration accrue du solvant à moyenne et grande distance du cation, qui se traduit par l'apparition de pics supplémentaires dans les courbes rdf cation atomes de solvant, ainsi que par un rapprochement (0.1 Å) de la deuxième couche de solvation du Magnésium. L'emploi d'un grand cutoff permet d'éliminer ces phénomènes. Il semble donc, qu'avec les conditions de solvation utilisées, nos simulations donnent sur le plan structural des résultats identiques à ceux obtenus lors de la détermination des paramètres(4).

Nos investigations nous ont permis d'obtenir des informations en ce qui concerne la première couche de solvation du calcium mal connue expérimentalement. On peut noter par exemple qu'en comparant leur géométrie dynamique et la largeur des premiers pics de rdf ion-o, la première couche de solvation du calcium semble moins rigide que celle du magnésium.

Toutefois, l'évaluation de la qualité d'une simulation ne doit pas être mesurée sur un seul critère. Nous allons donc aussi considérer les aspect énergétiques, via des calculs d'énergie libre.

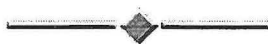
II. Calculs d'énergie libre : détermination de la différence d'énergie libre de solvation Ca²⁺ Mg²⁺

A priori ce type de calcul est principalement sensible à deux éléments : d'une part la description énergétique du système et d'autre part la qualité de l'exploration de l'espace des phases(21,22). Dans le cas qui nous intéresse, compte tenu de l'origine des paramètres relatifs à l'interaction cation-O, les divergences entre théorie et expérience doivent venir principalement du protocole d'échantillonnage. Ce système est donc un bon système modèle pour ajuster notre protocole.

Ainsi après avoir décrit le protocole de manière générale, nous aborderons les résultats proprement dits, obtenus lors de simulations à cutoff 14 et 7,5 Å.

A. Protocole de calcul

Cette étude utilise une approche alchimique couplée à la méthode de perturbation thermodynamique. Comme nous l'avons vu, cette approche fait intervenir la « mutation » de la nature du ligand, en l'occurrence, c'est le calcium qui est transformé en magnésium , et réciproquement, au sein de la sphère de solvation précédemment décrite. Dans sa forme la plus générale, le protocole de calcul d'énergie libre est illustré Figure II-1.



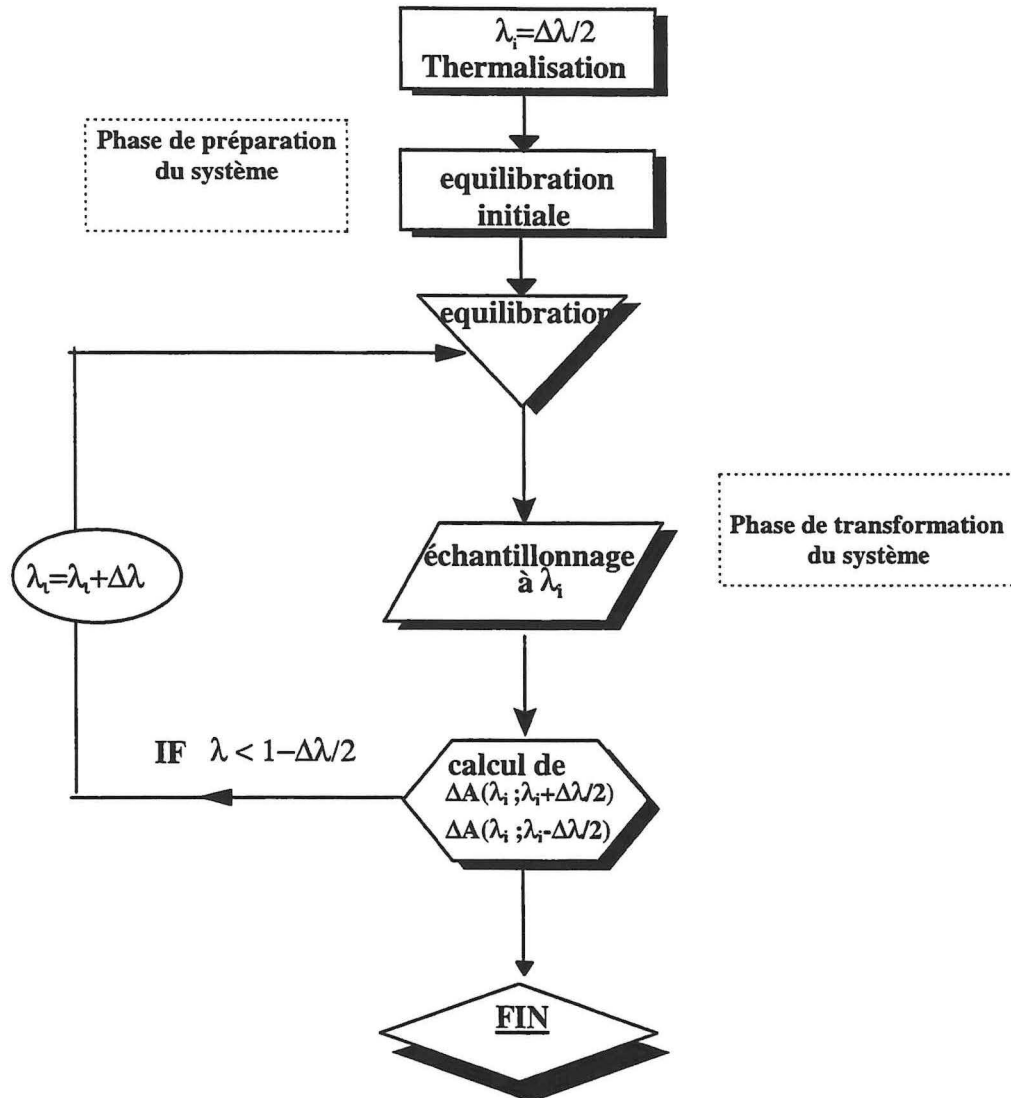


Figure II-1 : schéma d'un protocole de calcul d'énergie libre, via la méthode alchimique.

Ce protocole de calcul d'énergie libre peut être décomposé en deux phases. Tout d'abord, une phase de préparation durant laquelle, après initialisation de λ , le système est thermalisé puis équilibré. Cette phase de préparation est suivie d'une phase de transformation proprement dite, durant laquelle l'état initial est, cycle après cycle, progressivement transformé en état final, λ étant incrémenté de $\Delta\lambda$ entre deux cycles.

Le cycle de transformation élémentaire se décompose quant à lui en trois étapes :

i) Dans un premier temps, le système doit être rééquilibré, afin qu'il perde toute "mémoire" du λ précédent.

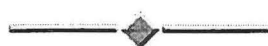
ii) Puis, un ensemble de coordonnées est collecté lors d'une dynamique de fluctuation effectuée à une valeur de λ donnée.

iii) Enfin, conformément à la technique du double fenêtrage, l'ensemble précédemment recueilli permet de calculer $\Delta A(\lambda; \lambda - \Delta\lambda)$ et $\Delta A(\lambda; \lambda + \Delta\lambda)$ selon la méthode de perturbation thermodynamique (cf. chapitre outils et méthodes). Il n'y a pas de discontinuité dans la transformation, c'est-à-dire que les dynamiques d'équilibration et de fluctuation se font les unes à la suite des autres. Pour ce faire, après chaque calcul de $\Delta A(\lambda; \lambda - \Delta\lambda/2)$ et $\Delta A(\lambda; \lambda + \Delta\lambda/2)$ et après incrémentation de λ , le premier pas de la trajectoire suivante est effectué à partir des dernières coordonnées et des dernières vitesses enregistrées lors de la dynamique précédente.

Une des difficultés majeures de ce type de calcul concerne l'évaluation de la précision du résultat. En effet, la comparaison aux données expérimentales est un critère qui ne permet en rien d'évaluer la précision du calcul. En première approximation, cette dernière peut être évaluée de manière globale à travers l'importance (en valeur absolue) de la variation $\Delta(\Delta A)$. Cette quantité est relative à la différence entre les variations d'énergie libre obtenues lors de transformations que l'on peut qualifier « d'aller et retour ». En effet, l'énergie libre étant une fonction d'état, nous avons :

$$|\Delta A(\text{Ca} \rightarrow \text{Mg})| = |\Delta A(\text{Mg} \rightarrow \text{Ca})|$$

La différence $\Delta(\Delta A) = \Delta A(\text{Mg} \rightarrow \text{Ca}) - \Delta A(\text{Ca} \rightarrow \text{Mg})$ devant être nulle lors d'un calcul "idéal", elle constitue donc un bon indicateur de la précision d'un calcul. C'est pourquoi, afin d'accéder à la quantité $\Delta(\Delta A)$, pour toutes nos simulations, nous avons réalisé un "cycle", à savoir qu'après avoir transformé le calcium en magnésium, nous avons, dans la continuité, effectué la transformation inverse.



De manière générale, les simulations que nous avons réalisées ont toutes en commun les éléments suivants:

- L'atome qui apparaît et disparaît est décrit selon une topologie simple, à savoir que le calcium et le magnésium occupent la même position, au centre de la sphère de solvation et ce durant toute la simulation.
- l'algorithme SHAKE nous a permis d'utiliser un pas temporel de 2fs.
- le période de thermalisation a une durée de 3ps. Elle est suivie d'une période d'équilibration de 5 à 20 ps durant laquelle les vitesses sont corrigées, le cas échéant, tous les 500 cycles d'intégration, de sorte que la température du système reste proche de 300K
- durant la période de fluctuation les coordonnées du système sont sauvegardées toutes les 20 fs.
- des fonctions de lissages type shift sont appliquées aux interactions non liantes.

Dans notre recherche de conditions optimales, nous avons principalement travaillé sur le nombre de cycles de transformation, ce nombre conditionnant par la même la durée totale de la simulation. Par ailleurs nous avons poursuivi l'étude de l'influence du cutoff, pour ce faire nous avons réalisé des simulations à cutoff 14 puis à cutoff 7,5 Å.

B. Simulation à cutoff 14 Å

1. Influence du nombre d'étapes de la transformation

Pour la mise au point du protocole d'échantillonnage, nous nous sommes principalement intéressés à l'impact du nombre d'étapes sur le calcul de la différence d'énergie libre de solvation de nos cations. Ainsi, nous avons réalisé trois simulations en faisant varier $\Delta\lambda$. Toutes ces simulations ont été réalisées à cutoff 14 Å. Les conditions ainsi que les résultats obtenus sont présentés dans le tableau ci-dessous :

| n° | $\Delta\lambda$ | pré | trans | $\Delta A_{\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}}$ kcal/mol | $\Delta A_{\text{Mg}^{2+} \rightarrow \text{Ca}^{2+}}$ kcal/mol | $\Delta(\Delta A)$ kcal/mol | ΔA_{moy} kcal/mol |
|----|-----------------|---------|----------|--|--|--------------------------------|-------------------------------------|
| A | 0.05 (20) | 3+5 ps | 5 +10 ps | -77.9 | 78.0 | 0.1 | -77.95 |
| B | 0.1 (10) | 3+5 ps | 5 +10 ps | -77.9 | 77.5 | -0.4 | -77.7 |
| C | 0.2 (5) | 3+20 ps | 10+10 ps | -73.6 | 77.7 | 4.1 | -75.6 |

Tableau II:1 :

Dans ce tableau, la colonne $\Delta\lambda$ contient le pas d'incrément de λ entre deux cycles ainsi que, entre parenthèses, le nombre de cycles nécessaires pour effectuer la totalité de la transformation. La colonne « pré » (pour préparation) contient les durées des périodes de thermalisation et de pré-équilibre. La colonne « trans » (pour transformation) contient la durée de la période d'équilibre relative à chaque étape de la transformation suivie de la durée de la période d'échantillonnage. Les colonnes $\Delta A(\text{Ca}^{2+} \rightarrow \text{Mg}^{2+})$ et $\Delta A(\text{Mg}^{2+} \rightarrow \text{Ca}^{2+})$ contiennent les résultats des transformations « aller et retour », $\Delta(\Delta A)$ étant la différence entre ces deux valeurs. Enfin ΔA_{moy} est la valeur moyenne du ΔA calculée à partir des transformations aller et retour. Ainsi, compte tenu du nombre de cycles des transformations et de la durée respective de ceux-

ci, pour calculer un ΔA , 300 , 150 et 100ps ont été nécessaires lors, respectivement, des calculs A, B et C (Tableau II:1).

Les deux premiers résultats moyens obtenus sont du même ordre de grandeur et le troisième est légèrement plus élevé ($\approx 2\text{kcal}$). Sur le plan de l'hystérésis ($\Delta(\Delta A)$), on note que cette dernière varie peu entre les calculs A et B, et qu'elle croît d'un ordre de grandeur entre les calculs B et C.

Comparons dans un premier temps nos résultats avec ceux initialement publiés par Åqvist(4), soit -75.3 ± 3.9 kcal/mol pour la transformation $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$ (et ce à partir d'un ajustement des paramètres pour une valeur expérimentale de -74.7 kcal/mol (23)). Ce résultat a été obtenu lors d'une simulation totale de 80 ps, sans cutoff. Les conditions de solvation ont déjà été décrites. Le nombre de cycles n'est pas précisé dans le texte. Cependant, la simulation réalisée avec $\Delta\lambda=0.2$ est en très bon accord avec cette étude. Lors de simulations plus précises ($\Delta\lambda=0.1$ et 0.05), on observe une différence de 3.6 et 3.9 kcal/mol par rapport à la valeur obtenue par Åqvist. Néanmoins cette différence est du même ordre de grandeur que l'incertitude associée à ce résultat. On peut considérer, par conséquent, que nos résultats sont en très bon accord avec ceux obtenus par Åqvist et ce malgré que nous ayons utilisé un protocole et des conditions de solvation différents.

Il est donc logique que nos résultats soient aussi en très bon accord avec l'expérience. De manière générale, on note dans la littérature des valeurs comprises selon les sources entre $-74.7(23)$, et $-80.3(24)$ kcal/mol, pour cette différence d'énergie libre de solvation, la valeur expérimentale la plus récente étant de $-77,7$ kcal/mol (25). Cette dernière valeur est très proche des résultats relatif aux simulations que nous avons réalisées à $\Delta\lambda=0.05$ et 0.1 .

Comme on peut le voir dans le Tableau II:1 :la précision la plus importante a été obtenue lors de la simulation réalisée avec $\Delta\lambda=0.05$, cette simulation ayant une hystérésis très faible, de 0.1 kcal/mol. Néanmoins, si on prend en compte l'aspect temps de calcul, il apparaît que la simulation effectuée avec $\Delta\lambda=0.1$ (Tableau II:1 :(B)), est un très bon compromis entre la précision du résultat obtenu et le temps de calcul nécessaire. C'est pourquoi pour la suite de notre travail nous avons adopté cette valeur.

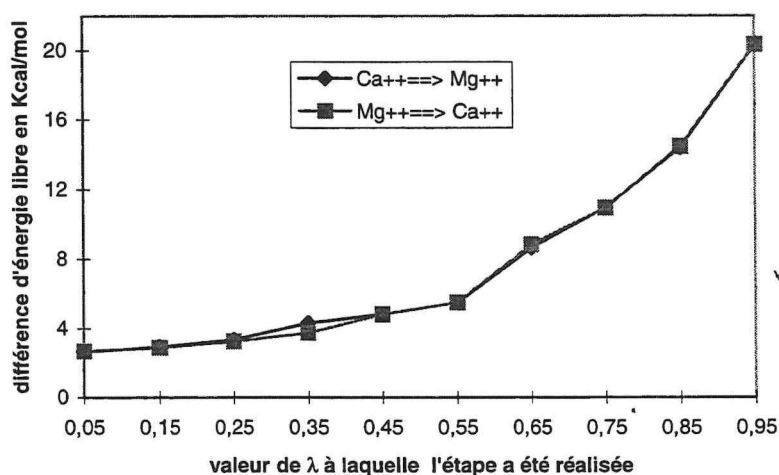
2. Analyses approfondies

Dans ce paragraphe, nous aborderons successivement, i) l'évolution de l'énergie libre au cours de la transformation. ii) l'évolution de l'environnement du système durant la « mutation » du cation. iii) la corrélation entre d'une part la différence d'énergie libre relevée entre deux étapes réalisées à une même valeur de λ et d'autre part les différences observées au niveau de l'échantillonnage de ces mêmes étapes.

a) L'évolution de l'énergie libre au cours de la transformation.

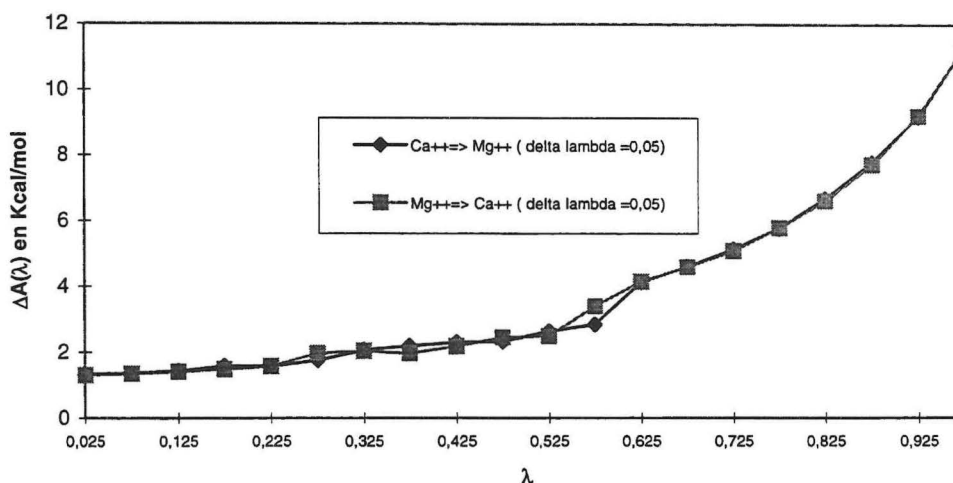
Les courbes ci-après représentent les quantités $\Delta A(\lambda)$ calculées à chaque étape de la transformation⁹ en fonction du paramètre λ et ce pour les simulations à $\Delta\lambda=0.1$ et $\Delta\lambda=0.05$. Du fait de la technique de double fenêtrage employée, on a :

$$\Delta A(\lambda_i) = \Delta A(\lambda_i + \Delta\lambda/2) - \Delta A(\lambda_i - \Delta\lambda/2)$$



courbes II-1 : (A)

⁹ Le cumul de ces quantités donne la quantité : $-\Delta A(Ca^{2+} \rightarrow Mg^{2+})$ et $\Delta A(Mg^{2+} \rightarrow Ca^{2+})$ figurant dans le Tableau II:1 : (A) et (B).



courbes II-1 : (B)

Représentation pour la transformation aller ($\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$) et retour ($\text{Mg}^{2+} \rightarrow \text{Ca}^{2+}$) de la quantité $\Delta A(\lambda)$ calculée à partir des échantillonnages réalisés en fonction du paramètre λ . Les courbes (A) et (B) sont respectivement relatives aux simulations réalisées avec un $\Delta\lambda=0.1$ et $\Delta\lambda=0.05$.

Ainsi, bien que la transformation du cation fasse intervenir une combinaison linéaire des énergies d'interaction du Ca^{2+} et du Mg^{2+} avec leur environnement, l'énergie libre calculée n'évolue pas linéairement en fonction de λ . Comme on peut le voir sur les courbes II-1 (A) et (B), elle semble croître presque exponentiellement au fur et à mesure que l'on s'approche de « l'état Magnésium » ($\lambda=1$).

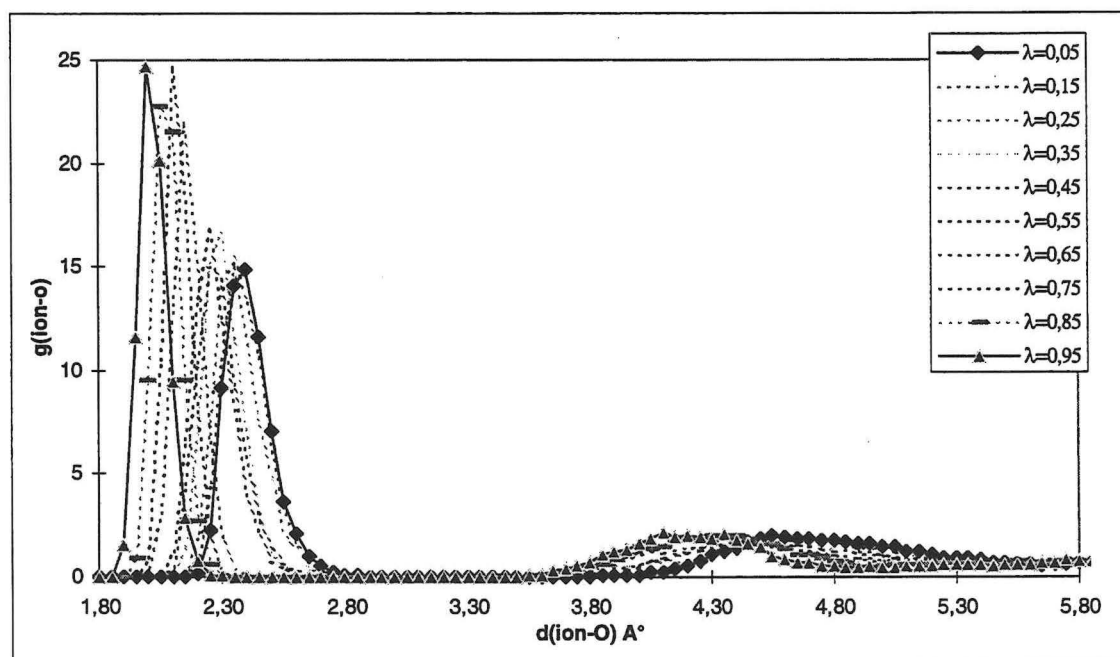
De plus, on peut voir que l'hystérésis observée lors de ces simulations (0.4 et 0.1 kcal/mol respectivement pour les simulations $\Delta\lambda=0.1$ et à $\Delta\lambda=0.05$), est imputable, dans le premier cas, principalement à l'étape réalisée à $\lambda=0.35$, alors que pour la simulation à $\Delta\lambda=0.05$ elle semble provenir de l'étape $\lambda=0.575$. Néanmoins, pour la majorité des étapes les $\Delta A(\lambda)$ obtenus sont quasi identiques lors de l'aller et du retour. Ce fait suggère fortement que la durée d'échantillonnage utilisée doit être suffisante. Mieux encore, que le résultat obtenu à un λ donné est peu sensible aux conditions initiales de la trajectoire effectuée. Notons que la comparaison des $\Delta A(\lambda)$ relatifs aux transformations aller et retour est un critère beaucoup plus précis que l'hystérésis totale. En effet, nous aurions

très bien pu observer une hystérésis nulle, qui aurait été le fruit de la compensation des $\Delta A(\lambda_i)$ obtenus à des λ_i différents.

b) Etude de l'échantillonnage

Nous avons étudié de manière plus précise l'échantillonnage à partir duquel l'énergie libre a été calculée. Comme nous l'avons vu, il est en effet possible d'avoir une idée globale de la structuration du système, à partir des courbes de distribution radiale cation-oxygène. Nous avons donc suivi l'évolution de la structuration du solvant autour du cation au cours de la transformation. Puis, nous avons cherché à mettre en évidence une corrélation entre la variation observée dans les différences d'énergie libre relatives aux transformations aller et retour avec celles observées dans les échantillonnages.

Les courbes ci-dessous illustrent le changement de structuration s'opérant au fur et à mesure de la transformation du cation. Ces courbes sont représentatives de ce que nous avons observé lors de tous nos calculs d'énergie libre.

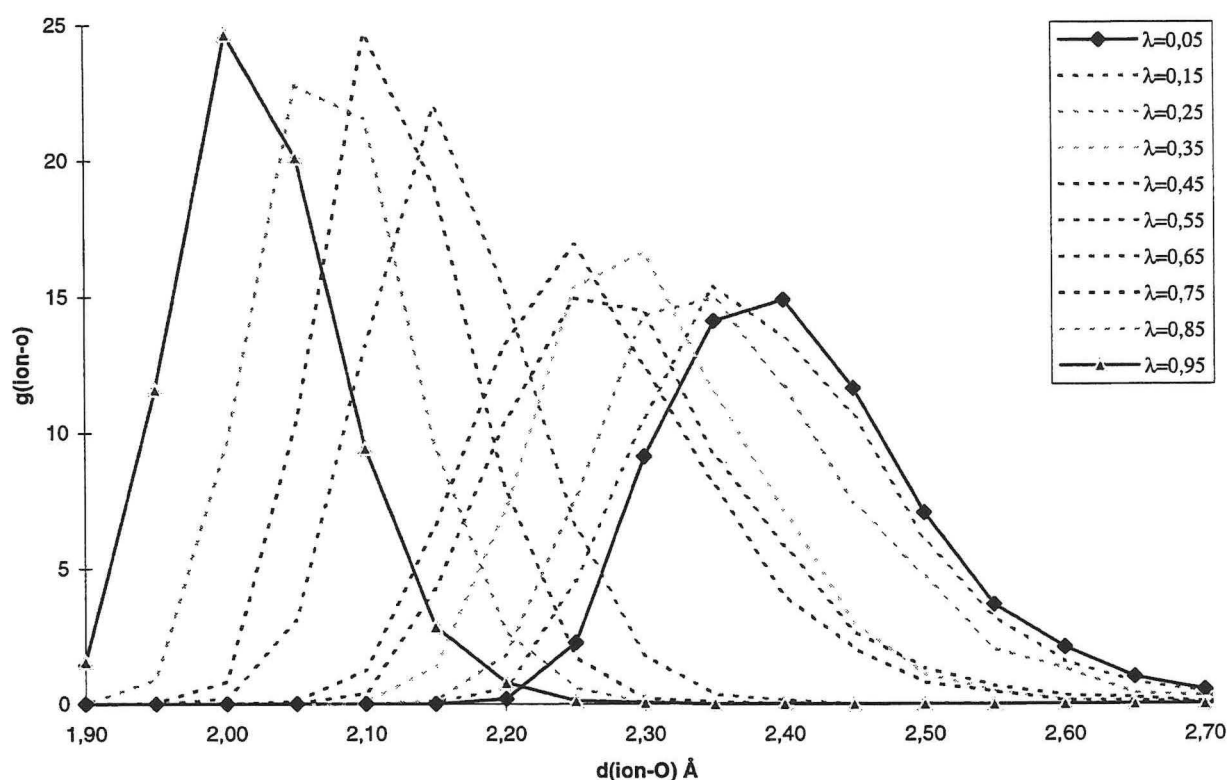


courbes II-2 : rdf cation-O , réalisées à partir des échantillonnages recueilli à chaque valeur de λ , lors de la transformation $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$, avec $\Delta\lambda=0.1$.

On note une translation progressive des deux premières couches de solvation au cours de la transformation. Nous avons étudié plus particulièrement l'évolution de ces deux couches.

(1) Première couche de solvation

Les courbes ci-dessous, sont un agrandissement de la partie relative au premier pic rdf des courbes II-2.

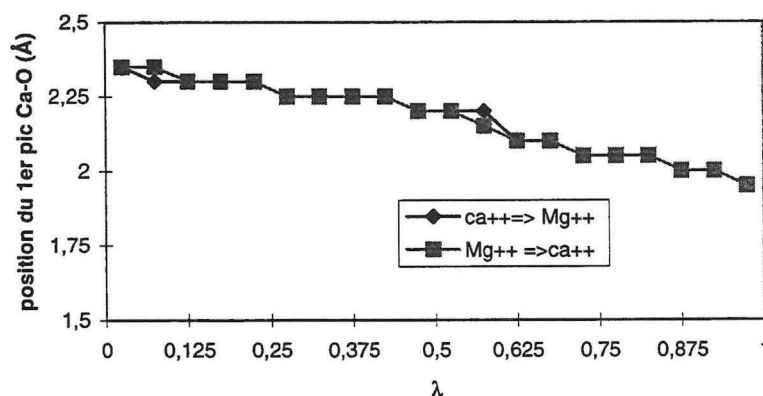
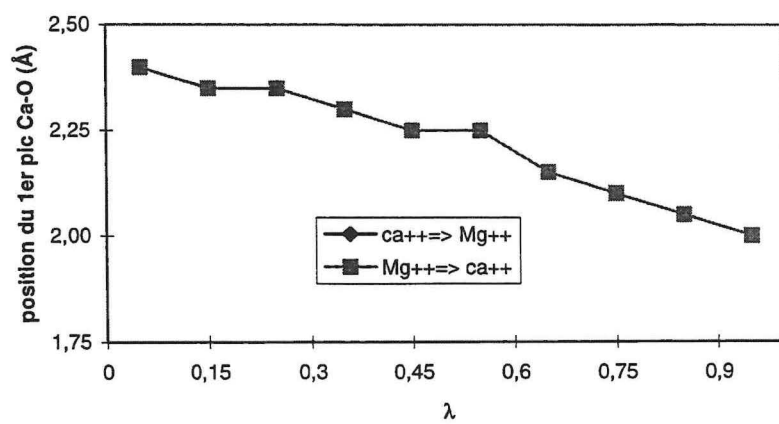


courbes II-3 :

premier pic rdf cation-O aux différents stades de la transformation

On peut voir de la droite vers la gauche, au fur et à mesure de la transformation, que le premier pic évolue de la position correspondant au Ca²⁺, soit 2.40 Å, vers celle correspondant au Mg²⁺, soit 2.00 Å. Ainsi, la distance d'équilibre cation-oxygène passe progressivement de l'état « calcium » à l'état « Magnésium ».





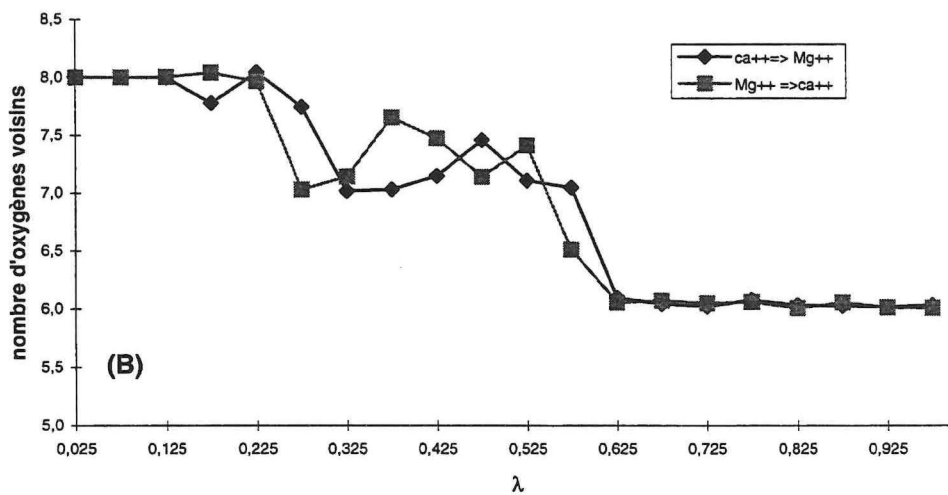
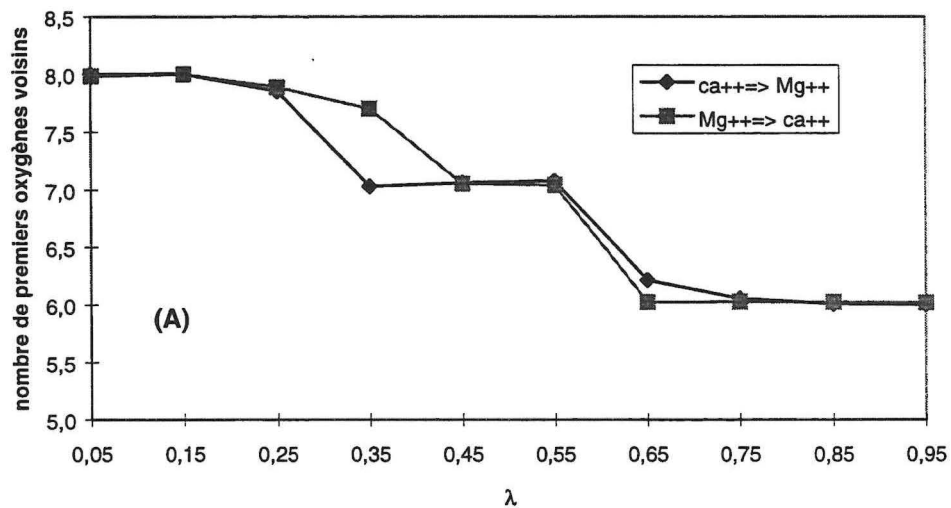
courbes II-4 : évolution de la position du maximum du premier pic de rdf cation-O en fonction de λ
 ((A) \Rightarrow simulation à $\Delta\lambda=0.1$, (B) \Rightarrow simulation à $\Delta\lambda=0.05$)

Comme on peut le voir sur la courbe ci-dessus, cette position évolue presque linéairement en fonction de λ , et ce de manière quasi identique entre les transformations « aller » et « retour ». (les petites différences sont dues à la précision de notre détermination de la position du maximum de ce pic)



(a) Nombre de voisins

Qu'en est-il du nombre de molécules d'eau impliquées dans cette première couche ? L'étude de l'évolution de l'aire du premier pic en fonction de λ est présentée dans les courbes ci-après.



courbes II-5 : nombre de molécules d'eau constitutives de la première couche de solvation du cation en fonction de λ . ((A) \Rightarrow simulation à $\Delta\lambda=0.1$, (B) \Rightarrow simulation à $\Delta\lambda=0.05$)

Comme on peut le voir ci-dessus, le nombre de molécules d'eau constitutives de la première couche évolue de 8 à 6. Néanmoins, contrairement à la distance d'équilibre, ce



nombre évolue par paliers définissant alors trois régimes . Le premier et le dernier régime font intervenir le même nombre de molécules d'eau que Ca^{2+} et Mg^{2+} , respectivement. Le premier régime se situe entre $\lambda = 0.0$ et $\lambda \approx 0.225$, le nombre de voisins en première couche suggérant une structure proche de celle décrite lors des trajectoires du Ca^{2+} en solution. Le troisième régime est relatif aux étapes comprises entre $\lambda \approx 0.625$ et $\lambda = 1$. Lors de ces étapes la première couche de solvatation du cation doit être proche de celle observée lors des trajectoires du Mg^{2+} en solution.

Dans le régime central, la première couche est constituée d'environ 7 voisins. Nous n'avons pas encore étudié en détail la géométrie de cette couche. Par contre, nous avons regardé quelles étaient les molécules d'eau qui sont présentes autour du cation durant la transformation à $\Delta\lambda=0.05$. Il est apparu que dans le régime à 7 molécules d'eau dans la première couche il y a fréquemment échange de molécule entre cette couche et la deuxième couche. Trois des sept molécules sont concernées. Aucun échange n'est à noter pendant les régimes à 6 ou à 8 voisins.

Dans l'état actuel des choses, deux hypothèses semblent plausibles. La couche à 7 voisins peut correspondre soit à une structure « métastable » dans laquelle les molécules d'eau auraient un temps de résidence plus court que dans les deux autres régimes, soit être le fruit d'un équilibre rapide entre des structures à 6 et à 8 voisins. L'étude de la structure dynamique de la couche de solvatation pourra nous permettre de trancher de manière définitive.

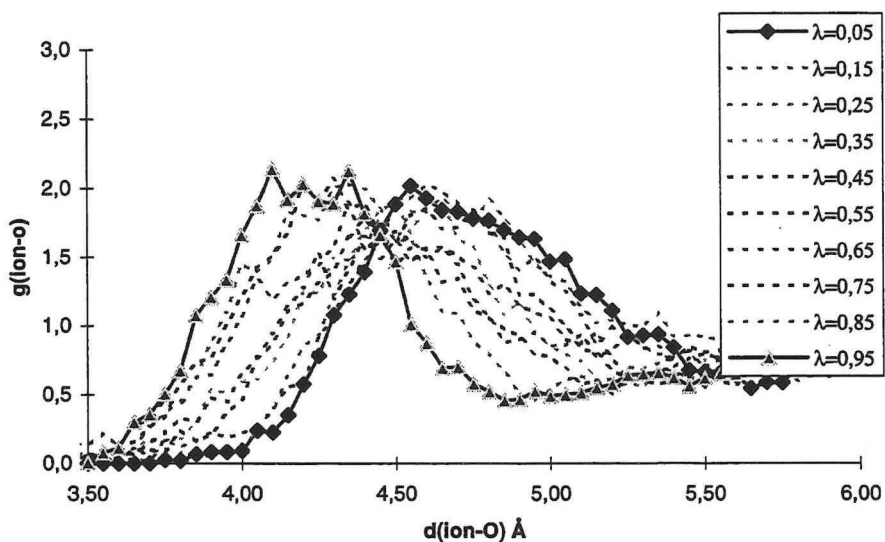
On peut noter par ailleurs que, lors de la comparaison des nombres de voisins au cours des transformations « aller » et « retour », les différences observées sont particulièrement importantes lors des étapes précédant (simulation $\Delta\lambda=0.1$) et/ou constituant (simulation $\Delta\lambda=0.05$) le « régime à 7 voisins ». Ce qui suggère que les



événements se produisant dans ce régime sont responsables en grande partie de l'hystérésis d'énergie libre.

(2) Deuxième couche de solvation

Comme nous l'avons vu lors de l'étude des trajectoires de cations en solution, les seconde couches du calcium et du magnésium sont beaucoup plus diffuses que la première et elles font intervenir un plus grand nombre de molécules d'eau. L'étude de cette couche est par conséquent beaucoup plus difficile. Comme on peut le voir sur les courbes ci-dessous, de même que pour la première couche, on assiste à un déplacement progressif de la deuxième couche de la position « Ca²⁺ » vers la position « Mg²⁺ » au cours de la transformation.



courbes II-6 : évolution de la position de la deuxième couche au cours de la transformation à $\Delta\lambda=0.1$.

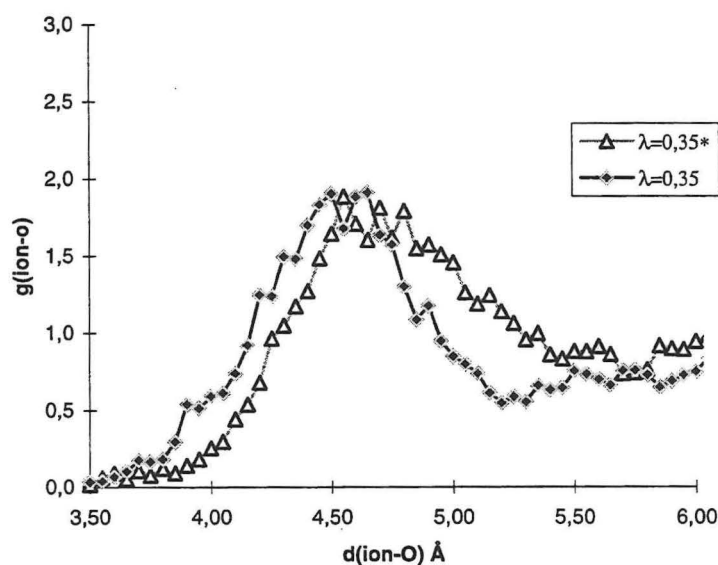
Le nombre de seconds voisins évolue, lui, de 17 à $\lambda=0.05$, à 12 pour $\lambda=0.95$ (simulation $\Delta\lambda=0.1$ (aller)). De manière générale, ces nombres sont proches de ceux obtenus lors des trajectoires relatives aux cations en solutions (pour les étapes extrêmes donc, à $\lambda=0$ et $\lambda=1$). Comme l'étude détaillée du nombre de voisins est pour les



deuxièmes couches beaucoup plus difficile, nous nous sommes contentés de comparer de manière qualitative des étapes illustrant les cas de figure suivant : i) étape à forte hystérésis d'énergie libre, ii) étape à faible hystérésis d'énergie libre.

Si l'on considère la simulation à $\Delta\lambda=0.1$, l'étape effectuée à $\lambda=0.35$ est responsable, en grande partie, de la différence d'énergie libre observée lors de cette simulation.

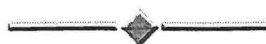
Les courbes II-7 ci-dessous représentent les pics rdf cation-o relatifs aux deuxièmes couches de solvation observées à cette valeur de λ .

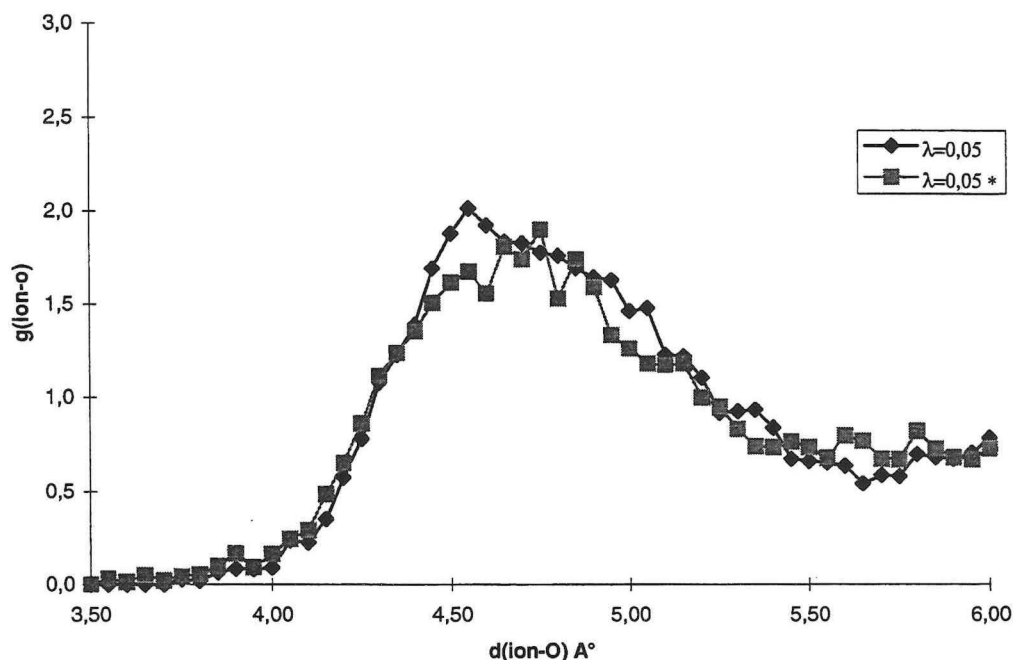


courbes II-7 : représentation des 2^{ème} pics rdf cation-O pour une valeur de λ le symbole * signifiant que l'étape est relative à la transformation retour

On constate, dans les courbes II-7, des différences importantes en ce qui concerne l'allure des deuxièmes couches, ce qui n'est pas surprenant puisque il y a aussi d'importantes différences en première couche.

A l'inverse les étapes aller et retour réalisés à $\lambda=0.05$ sont responsables d'une faible hystérésis.





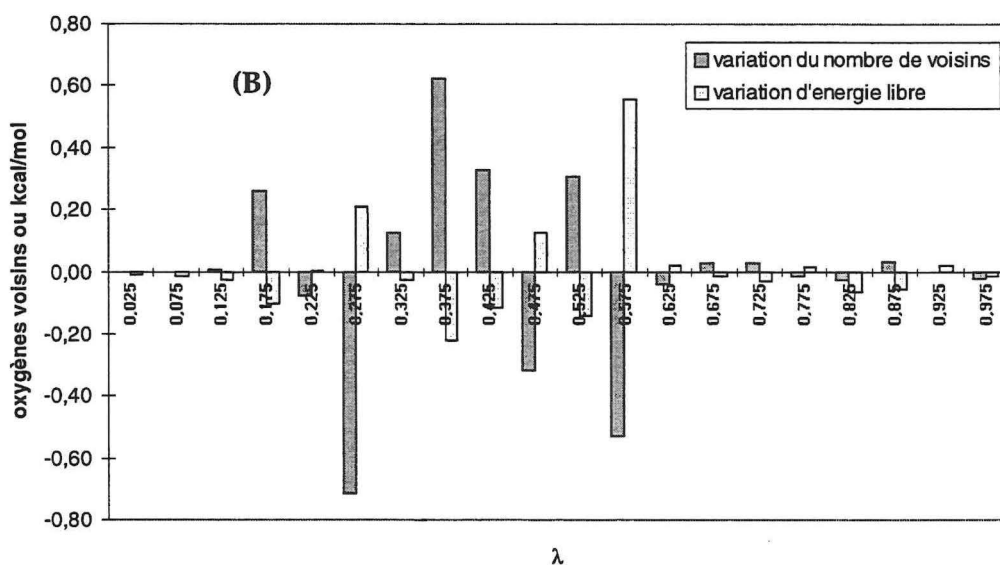
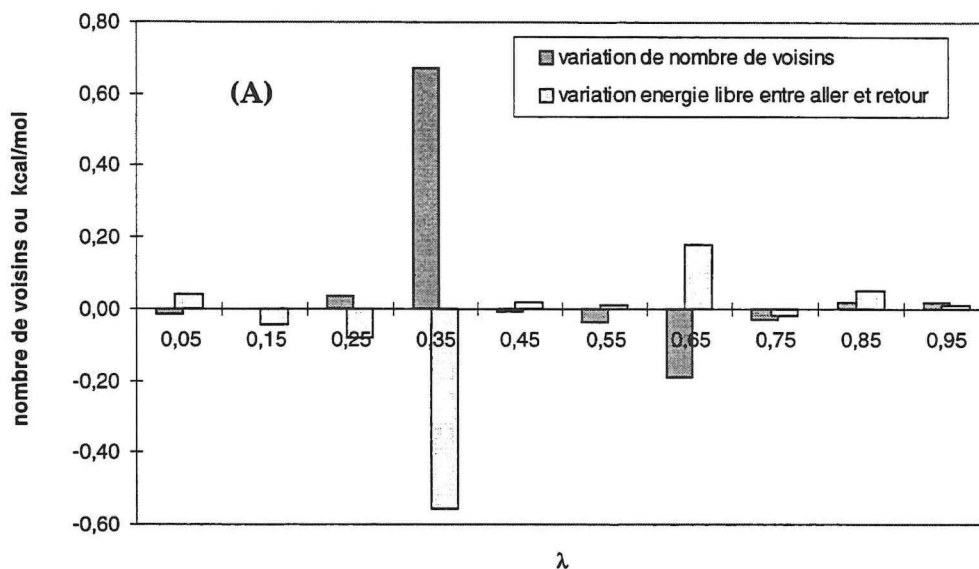
courbes II-8 : représentation des 2^{ème} pics rdf cation-O pour une valeur de λ le symbole * signifiant que l'étape est relative à la transformation retour

Comme on peut le voir sur les courbes ci-dessus, la deuxième couche semble avoir dans ce cas une structuration très comparable lors des transformations aller et retour.

c) Corrélation structure énergie libre

Nous avons tenté de faire le lien entre, d'une part, les différences observées entre aller et retour au niveau du calcul de l'énergie libre d'une étape, et d'autre part, les différences observées au niveau de l'échantillonnage réalisé lors de ces mêmes étapes. Pour ce faire nous avons réalisé les histogrammes ci-dessous. Ceux-ci représentent les variations d'énergie libre ainsi que les variations en terme de nombre de voisins de la première couche observées entre aller et retour et ce pour chaque étape d'un calcul.





histogrammes II-1 : représentation en fonction de λ de la différence d'énergie libre entre aller et retour ainsi que celle concernant le nombre de molécules d'eau constitutives de la première couche de solvation du cation « muté ». (A) simulation à $\Delta\lambda=0.1$ et (B) simulation à $\Delta\lambda=0.05$.

On peut constater qu'il y a une bonne corrélation entre la variation d'énergie libre et la différence des nombres de premiers voisins. De manière générale, une importante variation du nombre de premiers voisins se traduit par une importante variation en terme d'énergie libre.

Nous avons vu précédemment que les différences observées en ce qui concerne la structuration de la première couche se traduisent aussi par des différences de structuration au niveau de la deuxième couche. C'est sans doute pourquoi, dans les histogrammes II-1 (A) et (B), les variations d'énergie libre observées entre deux étapes réalisées à λ identique sont fortement corrélées avec les différences des nombres de molécules d'eau constitutives de la première couche. Ces histogrammes ne reflètent donc pas uniquement la part d'imprécision imputable à la structuration de la première couche. Néanmoins, la réorganisation de cette dernière semble être le phénomène majeur responsable des différences d'énergie libre observées.

Nous avons évalué la quantité d'énergie libre imputable à la réorganisation de cette couche¹⁰, et nous avons pu constater que le ΔA alors calculé, -62 kcal/mol, représente environ 70% de la différence d'énergie libre calculé pour l'ensemble du système. La contribution de la seconde couche dans le ΔA total ne peut donc représenter plus de 30% soit \approx -16 kcal/mol.

¹⁰ Pour ce faire, à partir de l'échantillonnage recueilli lors de la simulation réalisée à $\Delta\lambda=0.05$, nous avons calculé la différence l'énergie libre en ne tenant compte que des molécules d'eau situées dans la première couche, à moins de 3.5 Å du cation.

C. Simulation à cutoff 7,5 Å

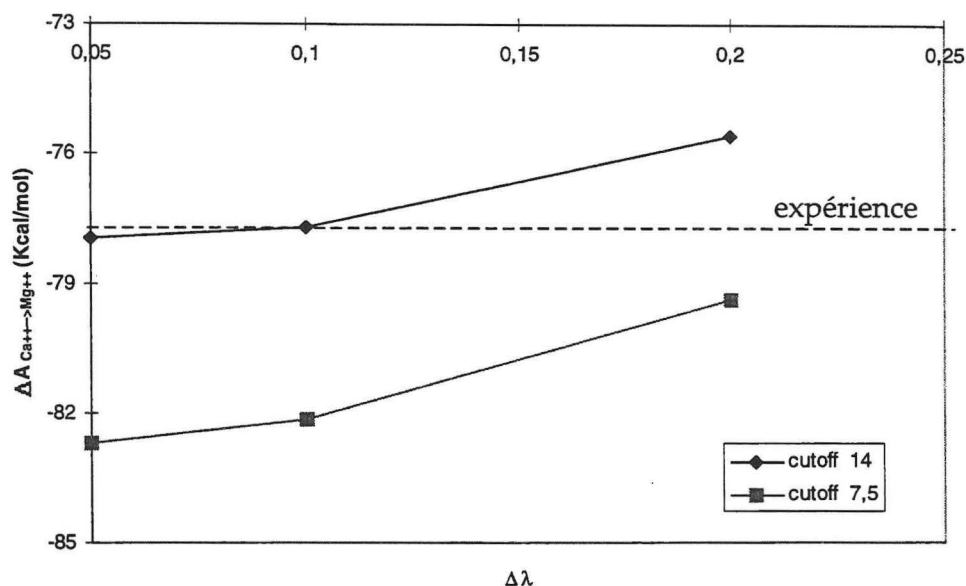
1. Influence de $\Delta\lambda$

Les résultats et conditions de ce nouveau calcul sont résumés dans le tableau ci-dessous.

| n° | $\Delta\lambda$ | pré | trans | $\Delta A_{Ca^{2+} \rightarrow Mg^{2+}}$ kcal/mol | $\Delta A_{Mg^{2+} \rightarrow Ca^{2+}}$ kcal/mol | $\Delta(\Delta A)$ kcal/mol | ΔA_{moy} kcal/mol |
|----|-----------------|---------|----------|--|--|--------------------------------|------------------------------|
| A | 0.05 (20) | 3+5 ps | 5 +10 ps | -82.9 | 82.5 | -0.4 | -82.7 |
| B | 0.1 (10) | 3+5 ps | 5 +10 ps | -78.9 | 85.0 | 6.1 | -81,95 |
| C | 0.2 (5) | 3+20 ps | 5+10 ps | -75.7 | 83.2 | -7.5 | -79.4 |

Tableau II:2 : La colonne $\Delta\lambda$ contient le pas d'incrémentation entre deux cycles de la transformation. La colonne « pré » contient les durées de thermalisation et pré-équilibration. La colonne « trans » contient les durées d'équilibration et de fluctuation de chaque cycle de la transformation. Les quatre autres colonnes sont relatives aux résultats proprement dits.

De manière générale, de même que lors des simulations à cutoff 14, nous pouvons constater une diminution de l'hystérésis au fur et à mesure que $\Delta\lambda$ diminue, néanmoins ces hystérésis sont nettement plus importantes que celles observées précédemment. Comparons, maintenant, les ΔA moyens obtenus lors des simulations réalisées à cutoff 14 Å avec ceux obtenus à cutoff 7.5.



courbes II-9 : Valeurs obtenues pour les deux conditions de cutoff. La ligne pointillée marque la valeur expérimentale.

Ainsi, il apparaît que les ΔA obtenus à cutoff 7,5 Å sont toujours inférieurs à ceux obtenus à cutoff 14 Å. Cette différence est en moyenne de 5,7 kcal/mol. Il en est de même en ce qui concerne l'accord aux données expérimentales, puisque, comme nous l'avons vu, celles-ci sont très bien reproduites à cutoff 14. Tout ceci n'est pas vraiment surprenant puisqu'Åqvist avait développé les paramètres que nous utilisons à partir de simulations sans cutoff.

2. Analyses approfondies

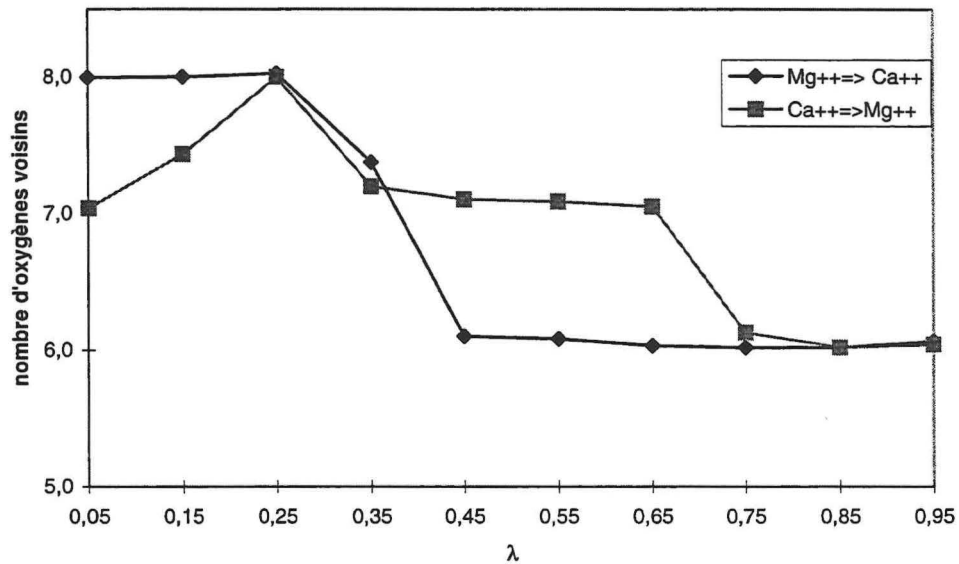
Etudions de manière plus précise l'échantillonnage de ces simulations. La simulation réalisée à $\Delta\lambda=0.1$ semble un bon choix dans la mesure où, lorsque on la compare à son homologue réalisée à cutoff 14, on note, d'une part, une différence d'énergie libre de -4.3 kcal/mol et d'autre part, une hystérésis entre aller et retour 15 fois supérieure. Ce phénomène est assez surprenant dans la mesure où les temps de



relaxation et d'échantillonnage n'ayant pas changé, on aurait pu s'attendre, à des résultats similaires en ce qui concerne le calcul de $\Delta(\Delta A)$.

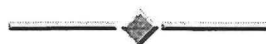
a) Etude de l'échantillonnage

Regardons de plus près ce qui se passe au niveau des échantillonnages au travers de l'évolution du nombre de voisins en première couche.



courbes II-10 : évolution du nombre d'oxygènes voisins en fonction du paramètre λ pour les transformations aller ($\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$) et retour ($\text{Mg}^{2+} \rightarrow \text{Ca}^{2+}$). Ce nombre est issu de l'intégration des premiers pics rdf cation O.

On peut noter que, cette fois, l'environnement du cation n'évolue pas correctement de l'état Ca^{2+} à l'état Mg^{2+} . En effet, l'environnement immédiat du cation n'est constitué que de 7 voisins en début de transformation ($\lambda=0,05$ et $0,15$) alors qu'il devrait s'agir en toute logique de 8 voisins. C'est d'ailleurs ce qui est observé à la fin de la transformation retour. Le chemin suivi est par ailleurs différent (cf. courbes II-10) d'une part de celui suivi lors de la simulation réalisée à cutoff 14 (cf. courbes II-10) et d'autre part de celui suivi au sein de la même simulation lors de la transformation $\text{Mg}^{2+} \rightarrow \text{Ca}^{2+}$.



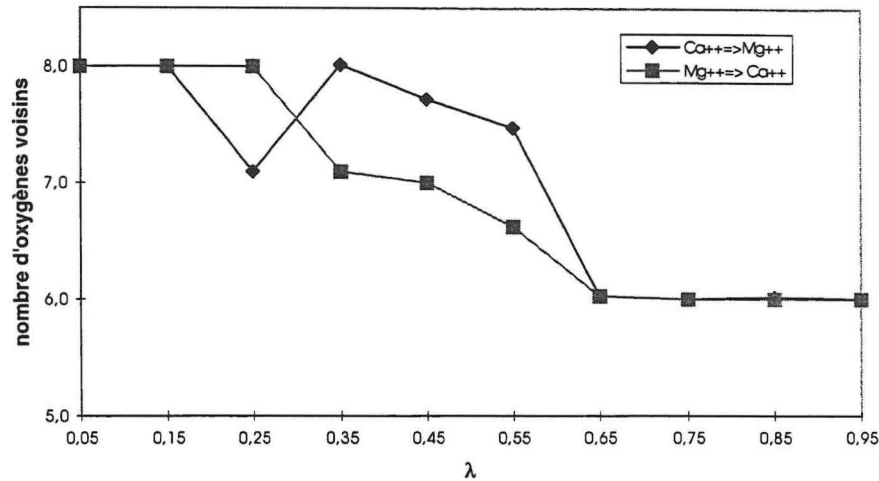
Comme à $\lambda=0.05$, c'est à dire proche de l'état Ca^{2+} , le cation n'est entouré que de 7 molécules d'eau voisines, on peut penser qu'un manque de relaxation du système peut être la cause de ce phénomène ainsi que des divergences observées entre les étapes aller et retour. Afin de confirmer cette hypothèse, nous avons réalisé une nouvelle simulation dans laquelle nous avons augmenté la durée de la pré-équilibration ainsi que les durées d'équilibration. Les résultats et conditions de ce calcul figurent dans le tableau ci-dessous.

| $\Delta\lambda$ | Pré | équilibration | fluctuation | $\Delta A \text{ Ca}^{2+} \rightarrow \text{Mg}^{2+}$ | $\Delta A \text{ Mg}^{2+} \rightarrow \text{Ca}^{2+}$ | $\Delta(\Delta A)$ | ΔA_{moy} |
|-----------------|---------|---------------|-------------|---|---|--------------------|-------------------------|
| 0.1 | 3+20 ps | 20 ps | 10 ps | -79.7 | -81.8 | -2.1 | -80.7 |

Tableau II:3

Cette simulation réalisée à cutoff 7.5 montre que l'augmentation de la durée de relaxation à bien pour conséquence une diminution de la valeur absolue de $\Delta(\Delta A)$. Le ΔA moyen est cependant toujours supérieur, de 3 kcal/mol, au ΔA moyen obtenu lors de la simulation à cutoff 14. A priori, la diminution de l'hystérésis d'énergie libre devrait ce traduire au niveau de l'évolution structurale de la première couche du cation par un meilleur recouvrement entre aller et retour en ce qui concerne l'évolution du nombre de voisins en fonction de λ . Les courbes ci-dessous représentent cette évolution lors de cette nouvelle simulation.





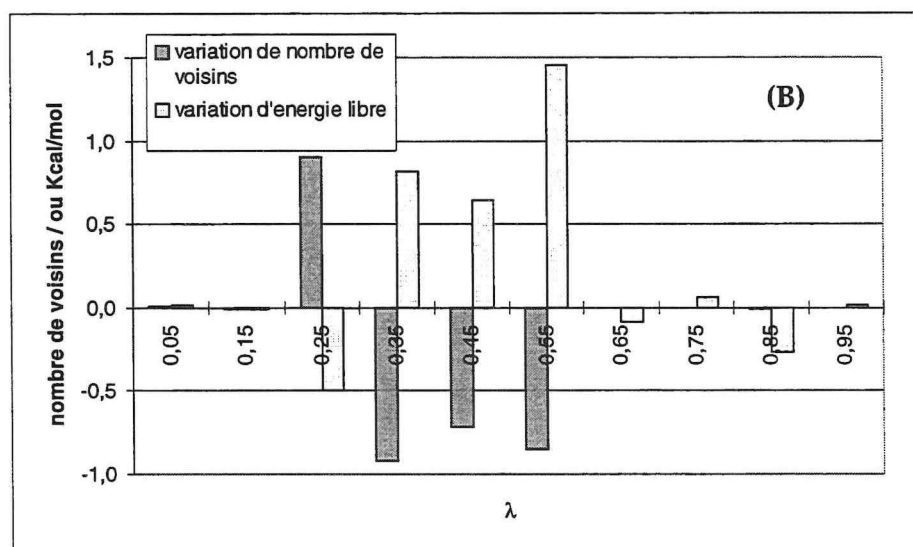
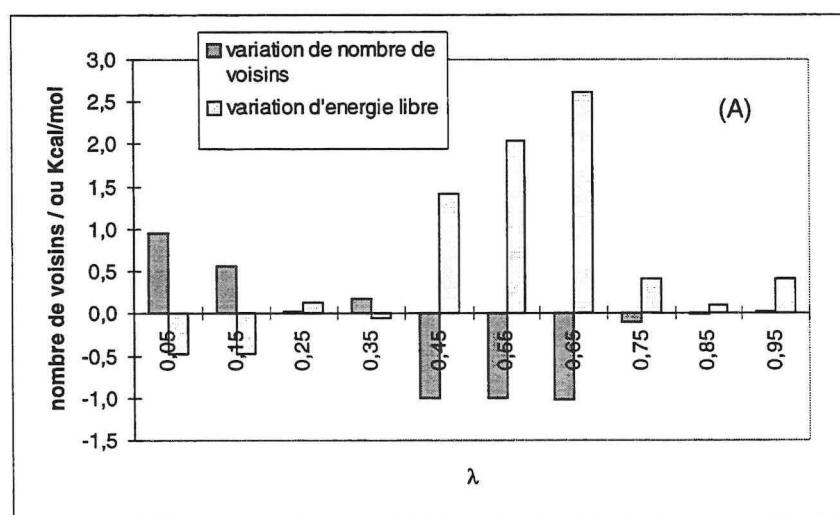
courbes II-11 : évolution du nombre de molécules d'eau de la 1^{ère} couche du cation en fonction de λ lors du calcul (Tableau II:3)

Dans cette simulation, l'environnement du cation est constitué cette fois de 8 voisins au début de la transformation ($\lambda=0.05$). L'augmentation de la durée de pré-équilibration a donc rempli son rôle. Toutefois, au total, entre thermalisation, pré-équilibration et première période d'équilibration, 28 ps de simulation auront été nécessaire pour arriver à ce résultat à cutoff 7.5 contre seulement 8 ps à cutoff 14 Å°.

Si l'on considère les courbes dans leur ensemble, l'augmentation des périodes d'équilibration a bien eu pour conséquence une amélioration du recouvrement entre les transformations « aller » et « retour ». Cependant, ce dernier est moins bon que celui obtenu à cutoff 14 et ce, malgré les 20 ps d'équilibration réalisée à chaque étape du calcul. D'importantes différences sont cependant encore observées, lors du régime à 7 voisins.

b) Corrélation entre différence de nombre de voisins et hystérésis d'énergie libre

Nous avons vu que lors des simulations réalisées à cutoff 14 nous avons une bonne corrélation entre la variation d'énergie libre et les différences observées en ce qui concerne le nombre de premiers voisins à une valeur de λ donnée. Effectuons la même comparaison pour les simulations réalisées à cutoff 7,5 Å.



histogrammes II-2: représentation en fonction de λ de la différence d'énergie libre entre aller et retour ainsi que celle concernant le nombre de molécules d'eau constitutives de la première couche de solvation du cation « muté ». (A) est relatif à la simulation (A) du Tableau II:2 et (B) à celle présentée dans le Tableau II:3 .



Les variations observées en ce qui concerne le nombre de voisins en première couche se traduisent par des variations importantes de la différence d'énergie libre, de même que lors des simulations réalisées précédemment à cutoff 14. Quelques exceptions sont néanmoins observées, en particulier dans la deuxième moitié de la transformation, c'est-à-dire au fur et à mesure que l'on se rapproche de l'état Mg^{2+} ¹¹. Comme ces histogrammes représentent seulement les différences observées en première couche, les hystérésis d'énergie libre de ces dernières étapes semblent donc être imputables au reste du système, et probablement en grande partie à des différences de structuration en deuxième couche.

Nous avons vu qu'à cutoff 14, une différence observée en deuxième couche était souvent due à une différence en première couche. Cela semble donc être un peu moins le cas à cutoff 7.5. Il pourrait s'agir d'un artefact supplémentaire dû au petit cutoff, à ajouter à ceux précédemment relevés lors de l'étude des trajectoires de cations en solution. On peut néanmoins penser que son origine est la même, à savoir que les molécules d'eau constitutives des deux première couches ne peuvent pas se voir les unes les autres alors qu'elles subissent toutes l'influence du cation. Cette hypothèse pourrait d'ailleurs aussi justifier l'augmentation du temps de relaxation que nous avons constaté.

¹¹ Nous avons vu que c'est lors de ces étapes que la variation d'énergie libre calculée est la plus importante (cf. courbes II-1).

D. conclusion

(a) Influence du cutoff

Un petit cutoff induit des artefacts au niveau des simulations, ce qui se traduit par une surévaluation de l'énergie libre calculée ainsi que par une hystérésis plus importante, imputable probablement à une augmentation du temps de relaxation du système. La cause de tous ces phénomènes est sans doute principalement la compartimentation inhérente à l'utilisation d'un petit cutoff. En effet, comme nous l'avons vu, les molécules d'eau constitutives des deux premières couches ont dans ce cas une vue partielle de l'ensemble de leurs congénères. Ainsi, alors qu'elles subissent pleinement l'influence structurante du champ électrostatique du cation, elle ne subissent celles des molécules d'eau plus éloignées que de manière indirecte, le tout aboutissant à une sur-structuration¹² de la deuxième couche de solvation et à une augmentation de l'énergie libre d'environ 6 kcal/mol. Cette quantité peu paraître faible comparée aux -77,7 kcal/mol escomptées mais il faut se rappeler que la différence d'énergie libre que l'on cherchera à calculer par la suite est du même ordre de grandeur.

On pourrait s'étonner du fait que l'implication éventuelle des frontières de notre système n'ait pas été évoquée. Nous pensons qu'elles sont cette fois hors de cause dans la mesure où des phénomènes similaires ont déjà été observés lors du calcul de l'énergie libre de solvation du Na^+ avec des conditions de solvation de type périodique et sphère SCAAS (7). En effet Åqvist a montré que l'énergie libre de solvation du Na^+ calculée avec un cutoff de 10 Å est surévaluée de 5 kcal/mol par rapport à celle

¹² Sur structuration, qui se traduit par un rapprochement de la deuxième couche de solvation d'environ 0.1 Å (cf. comparaison courbe rdf Mg-o de cation en solution aux deux valeurs de cutoff)

déterminée sans cutoff. De plus, par l'introduction d'un double cutoff¹³, il a montré que cette surévaluation disparaît lorsqu'on permet aux molécules d'eau présentes dans la sphère du cation d'interagir toutes entre elles.

Compte tenu de tous ces éléments, il est donc clair que l'emploi d'un petit cutoff dans les simulations de cation en solution est à éviter. Le cutoff doit être adapté de façon à ce que l'ensemble des molécules d'eau constitutives des couches de solvation du cation puissent interagir entre elles. C'est le cas lors des simulations réalisées avec un cutoff de 14 Å. Les résultats obtenus avec ce dernier sont d'ailleurs à peu près identiques à ceux initialement obtenus par Åqvist et ce indépendamment des conditions de solvation utilisées.

(b) *Protocole de simulation d'énergie libre*

Sur le plan du protocole nous avons vu que la décomposition de la transformation $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$ en dix étapes de 15 ps (5 d'équilibration suivies de 10 ps de fluctuation) avec $\Delta\lambda=0.1$ permet d'avoir des résultats avec une précision respectable et ce pour un coût en temps CPU raisonnable (300ps \approx 100 heures Cpu sur IBM SP2).

Bien que nous ayons obtenu de bons résultats, il reste difficile de définir des règles générales pour établir un protocole de calcul de différences d'énergie libre. On doit autant que faire se peut suivre les recommandations données en préambule, à savoir, d'une part avoir une description énergétique adaptée au problème étudié et d'autre part adapter les temps de relaxation et d'échantillonnage spécifiquement au système. Il semble que ce soit chose faite pour la transformation $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$ dans l'eau, reste à savoir ce qu'il en est pour les transformations au sein des sites de fixation de

¹³ C'est-à-dire qu'il a différencié le cutoff du cation de celui des molécules d'eau.

calciprotéines. De fait nous allons chercher à utiliser le protocole mis au point pour les simulations dans l'eau.

(c) *Calcium et Magnésium dans l'eau*

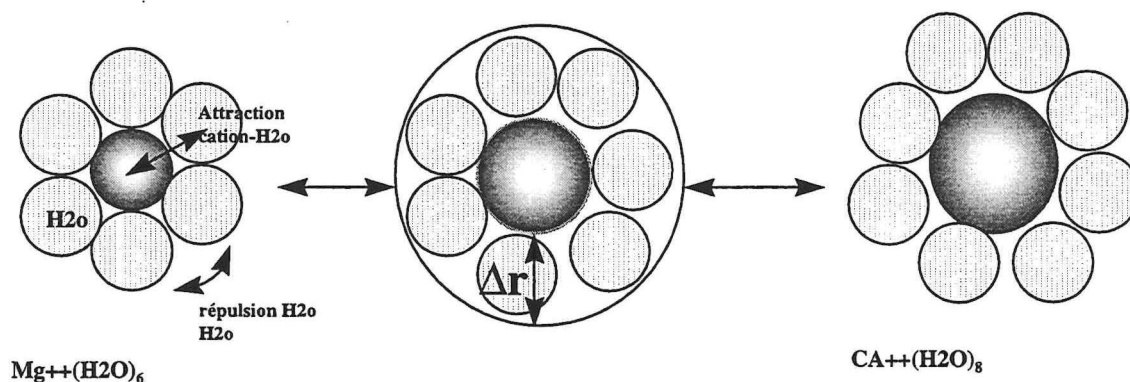
Lors des calculs d'énergie libre nous avons vu que l'eau se comporte comme une matrice qui s'adapte en fonction de la taille de l'objet solvato. Cette constatation n'est pas nouvelle, néanmoins ce travail en donne une belle illustration. Lors des trajectoires de l'eau autour du cation nous avons montré que l'environnement du calcium est constitué de 8 molécules d'eau adopte une géométrie type anti-prisme à base carrée alors que le magnésium est entouré de six molécules d'eau disposées en octaèdre. Nous avons suivi un chemin entre ces deux structures en transformant de manière alchimique le cation central. Le chemin n'a pas de sens physique mais le départ et l'arrivée sont eux parfaitement réels.

Nous avons vu que le nombre de molécules d'eau directement coordonnées au calcium est plutôt mal connu expérimentalement, il en est de même sur le plan théorique. En effet, avec 8 voisins autour du calcium, nos résultats sont certes en accord avec ceux obtenus récemment lors de trajectoires de dynamique moléculaire mixte quantique/classique (QMM) (26), ou classique (13,27), mais on trouve aussi, par ailleurs, des études dans lesquelles un nombre de voisins en première couche,¹⁴ varie selon les cas de 7(28) (29) à 9(30). Dans la plus-part de ces travaux seul l'aspect structural était considéré. Nous pourrions donc argumenter que les paramètres que nous avons utilisés donnent d'excellents résultats en matière de différences d'énergie libre, cependant on peut considérer que, compte tenu de leur origine, cet argument est pour le moins biaisé.

¹⁴ Ces deux études font intervenir des trajectoires de dynamiques moléculaires ou Monté Carlo, avec des fonctions énergie potentielle cation-o originale. Les modèles d'eau utilisés sont aussi différents.

Il semble donc que la nature de la géométrie et le nombre de voisins de la couche de solvation du calcium soit pour quelques temps encore une question ouverte ! Cependant, S'il était confirmé qu'il y a bien 8 voisins autour d'un calcium en solution, on peut croire que les informations géométriques que nous avons présentées s'avéreront alors aussi pertinentes.

Par ailleurs, ces simulations ont permis de souligner le lien, entre composition (nombre et géométrie) de la couche d'hydratation et énergie libre de solvation. C'est ce lien qui est d'ailleurs utilisé dans des modèles empiriques, tel celui de Marcus¹⁵ (25). Notre étude, en accord avec ce modèle, montre que la différence d'énergie libre est le fruit de la réorganisation de l'environnement du cation suite à la modification de la distance d'équilibre cation-o.



Ainsi le nombre de molécules en première couche est conditionné par la distance d'équilibre, qui est due à la taille du cation et à un équilibre subtil entre les forces attractives générées par la présence du cation et les forces répulsives entre molécules

¹⁵ Il permet de rendre compte de grandeurs thermodynamiques (modèles de solvation d'ions (charge -1 à +4)) développés empiriquement à partir de la charge et du rayon de l'ion r . Le modèle spécifie l'épaisseur de la première couche (Δr) ainsi que le nombre de molécules d'eau constitutives de la première couche de l'ion, proportionnellement à $|z|/r$ et aux propriétés diélectriques à haut champ de l'eau pure.

Calculs d'énergie libre : détermination de la différence d'énergie libre de solvation Ca^{2+} Mg^{2+}
d'eau voisines (cf. figure ci-dessus). Tout se passe comme si le cation jouait le rôle de clé de voûte centrale par l'intermédiaire du champ électrostatique qu'il génère et qu'en fonction de sa taille un nombre plus ou moins important de molécules d'eau peuvent s'en rapprocher. Lors de nos simulations, en faisant varier cette taille, nous voyons augmenter ou diminuer le nombre de molécules d'eau directement coordonnées au cation.



III. Références Bibliographiques :

1. Gerschel, A. (1995), pp. 96-97, Interéditions CNR éditions, Paris
2. van Gusteren, W. F., and Berendsen, H. J. C. (1985) *In Molecular Dynamics and protein Structure*, Hermans, J. Ed. Poly cristal book Serv (Hermans, J., Ed.), III. III vols., Springer, Western
3. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) *J. Chem. Phys.* **79**, 926
4. Aqvist, J. (1990) *J. Phys. Chem.* Vol **94**(Iss 21), pp 8021-8024
5. Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., and Hermans, J. (1981) *J. In intermolecular forces*, Reidel: Dordrecht Ed. (B., P., Ed.), the Netherland
6. King, G., and Warshel, A. (1989) *J. Chem. Phys.* **91**, 3647
7. Aqvist, J. (1994) *J. Chem. Phys.* **98**, 8253-8255
8. Warshel, A., and Creighton, S. (1989) (van Gunsteren, W. F., and Weiner, P. K., eds), pp. 120, ESCOM, Leiden
9. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) *J. Comput. Chem* **4**, 187
10. Marrone, T. J., and Merz, K. M. (1993) *J Phys Chem* **97**(24), 6524-6529
11. Marrone, T. J., and Merz, K. M. (1994) *J Phys Chem* **98**(33), 8256-8257
12. Shiratori, Y., and Nakagawa, S. (1991) *J. Comput. Chem.* **12**, 717-730
13. Obst, S., and Bradaczek, H. (1996) *J. Phys. Chem.* **100**, 15677-15687
14. Soper, A. K., and Philips, M. G. (1987) *J. Chem. Phys.* **107**, 47
15. Alary, F., and Durup, J. (1993) *J. Phys. Chem.* **97**, 13864-13876
16. Sanejouand, Y. H., and Tapia, O. (1995) *J Phys Chem* **99**(15), 5698-5704
17. Prevost, M., Van Belle, D., and Wodak, S. (1990) *Molecular Physics* **71**, 587-603
18. Marcus, Y. (1988) *Chem. Rev.* **88**, 1475

19. Othaki, H., and Radnai, T. (1993) *Chem. Rev.* **93**, 1157-1204
20. Enderby, J. E., and Neilson, G. W. (1980) *Adv.Phys.* **29**, 323
21. Van Gusteren, W. F., and Berensen, H. J. C. (1987) *Journal of Computer-Aided Molecular Design* **1**, 171-176
22. Kollman, P. A. (1993) *Chemical Reviews* **93**, 2395-2417
23. Burgess, M. A. (1978) *Metal Ions un Solution* (Horwood, E., Ed.), Chichester
24. Gomer, R., and Tryson, G. (1977) *J.Chem.Phys.* **66**, 4413
25. Marcus, Y. (1994) *Biophys Chem* **51**(2-3), 111-127
26. Tongraar, A., Liedl, K. R., and Bernd, M. R. (1997) *J.Phys.Chem.* **101**, 6299-6309
27. KalKo, S. G., Sesé, G., and Padro, J. A. (1996) *J.Chem.Phys* **104**, 9578
28. Probst, M. M., Radnai, T., Heinzinger, K., Bopp, P., and Rode, B. M. (1985) *J.phys.Chem.* **89**, 753
29. Bernaluruchurtu, M. I., and Ortegablake, I. (1995) *J Chem Phys* **103**(4), 1588-1598
30. Bounds, D. G. (1985) *Molecular Physics* **54**, 1335-1355

B) CATIONS DANS LES CALCIPROTEINES

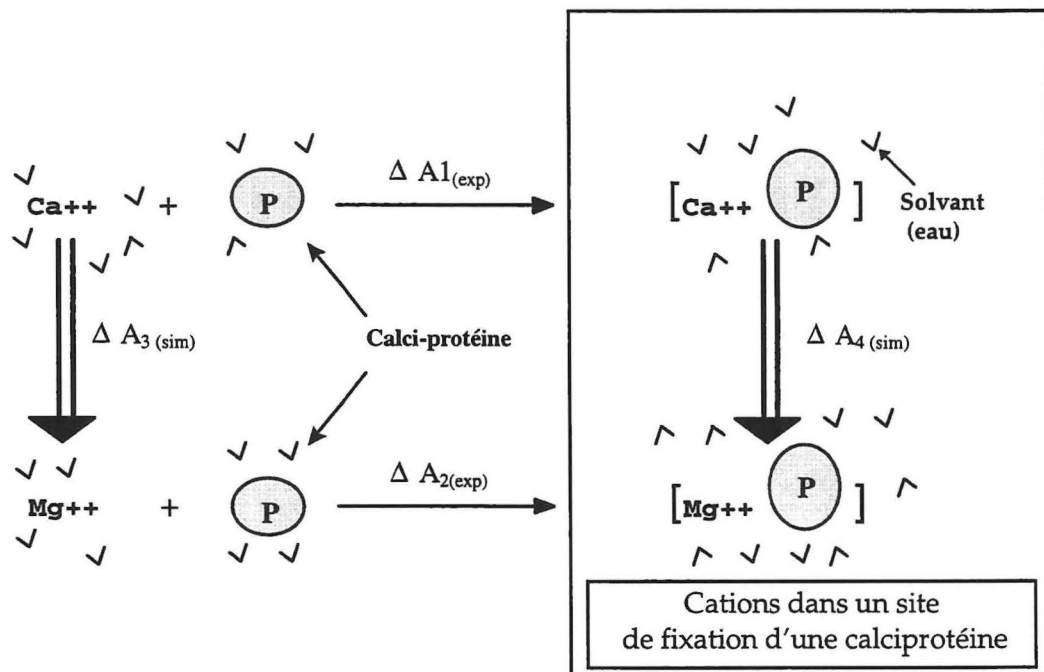




Table des matières

| | |
|--|------------|
| I. cations dans les calciprotéines | 193 |
| A. Protocole de calcul | 199 |
| 1. Modèle | 199 |
| 2. Champ de force. | 200 |
| 3. Minimisation et simulation de Dynamique Moléculaire . | 200 |
| 4. Calcul de différences d'énergie libre. | 201 |
| B. Etude de la parvalbumine | 202 |
| 1. L'échange $\text{Ca}^{2+}/\text{Mg}^{2+}$ dans le site EF - 3 de la parvalbumine. | 202 |
| 2. Echange $\text{Mg}^{2+}/\text{Ca}^{2+}$ dans le site EF-4 de parvalbumine. | 212 |
| C. Echange $\text{Ca}^{2+}/\text{Mg}^{2+}$ dans d'autres protéines à Main EF . | 221 |
| 1. Troponine C | 221 |
| 2. Calmoduline | 222 |
| 3. protéine sarcoplasmique | 225 |
| D. DISCUSSION | 226 |
| 1. Echange $\text{Ca}^{2+}/\text{Mg}^{2+}$ dans le site EF-4 de la parvalbumine. | 226 |
| 2. Le site EF-3 occupé par Mg^{2+} . | 229 |
| 3. Conformation de la chaîne latérale de Glu 12 | 230 |
| 4. Sites main EF d'autres calciprotéines à main EF. | 232 |
| E. Conclusion | 233 |
| F. Tentative avec une nouvelle paramétrisation de l'interaction cation - oxygène | 234 |
| G. Références Bibliographiques | 239 |



Introduction

L'ensemble de paramètres utilisés lors de la transformation $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$ (1), a permis de bien reproduire la différence d'énergie libre de solvation de ces cations dans l'eau. Ce chapitre est consacré aux résultats obtenus lors des simulations mettant en œuvre le calcium et le magnésium au sein des sites de fixation de calciprotéines.

Nous avons vu dans le premier chapitre que, en comparant deux structures à haute résolution de la parvalbumine (2), PaCaCa.NH4 (1pal) et PaCaMg.Mg (4pal), dans lesquelles le site EF-4 (ou EF) est respectivement occupé par un Ca^{2+} ou un Mg^{2+} Declercq et col, ont observé que le résidu Glu 101 effectue un changement de conformation. En effet, il passe du rotamère χ_1 gauche(+) ($\chi_1 = -75^\circ$) dans la forme chargée en calcium au rotamère χ_1 gauche(-) ($\chi_1 = +60^\circ$) dans la forme chargée en magnésium, et ce alors que l'angle dièdre χ_2 reste inchangé et que l'angle dièdre χ_3 adopte des formes quasi miroirs l'une de l'autre dans les deux structures. Comme le montre la Figure II-1A, la transformation gauche(+) \leftrightarrow gauche(-) de χ_1 satisfait à un critère bien précis, i.e., la position du cation reste pratiquement invariante dans la structure tertiaire de la protéine. Cette adaptation conformationnelle de la chaîne latérale du Glu 101 permet un changement du nombre de coordination du cation, Glu 101 interagissant alors comme un ligand bidentate en présence de Ca^{2+} et comme un ligand monodentate en présence de Mg^{2+} .

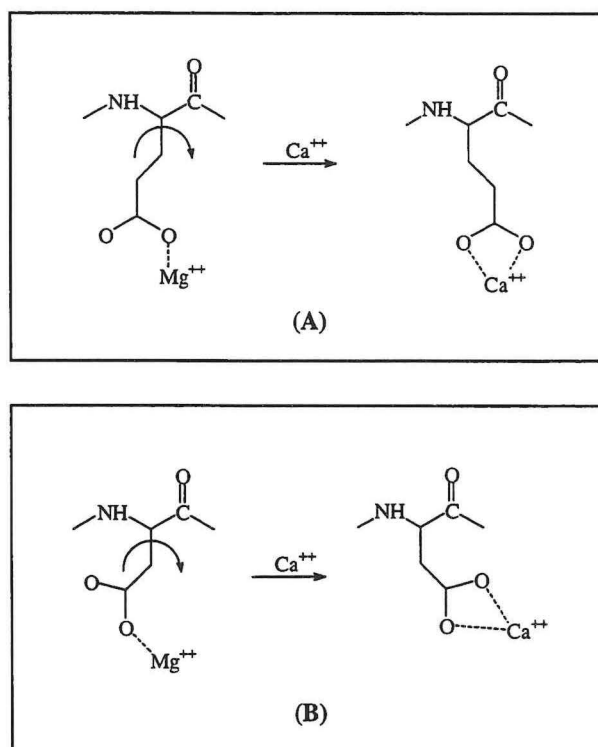


Figure II-1 : Transition Conformationnelle du a) Glu , et du b) Asp en position relative 12 d'un site de fixation main EF, quand le nombre de coordination du cation fixé change d'une unité. La première transition est décrite dans le cas de la parvalbumine. L'autre est hypothétique et nécessiterait une adaptation du site de fixation tout entier (voir le texte).

A partir de données RMN, il a aussi été montré que dans le site EF-3 (ou CD) de la parvalbumine l'homologue du Glu 101, le résidu Glu 62 , a un comportement similaire lors de l'échange $\text{Ca}^{2+}/\text{Mg}^{2+}$, au moins en ce qui concerne χ_1 (3). Les deux résidus glutamique Glu 62 et Glu 101 occupent la position relative 12 dans le motif main EF des sites EF-3 et EF-4. Ce Glu 12 est un résidu hautement conservé dans ce motif. Il est remplacé par un Asp dans seulement 8% des séquences connues(4). Comme on peut le voir sur la Figure II-1.B cette substitution pourrait compromettre alors la capacité de la protéine à faire l'échange $\text{Ca}^{2+}/\text{Mg}^{2+}$ (cf. Figure II-1B), car le passage bidentate \leftrightarrow monodentate implique alors un déplacement du cation, ainsi que celui de tous les autres oxygènes de la sphère de coordination.



Dans le but de tester l'hypothèse que l'adaptabilité conformationnelle de la chaîne latérale du Glu en position relative 12 joue aussi un rôle dans d'autres motifs main EF, nous avons étendu notre étude à d'autres calciprotéines à main EF pour lesquelles aucune structure « Magnésium »¹ n'est connue. Ainsi, nous avons transformé selon l'approche alchimique le cation initialement présent dans les sites main EF et nous avons suivi l'évolution structurale de ces sites durant la transformation, tout en déterminant, via un cycle thermodynamique la stabilité relative des différents complexes cation protéine.

A. Protocole de calcul

1. Modèle

Comme nous l'avons vu, les calculs d'énergie libre requièrent un échantillonnage efficace de l'espace configurationnel. En pratique, avec la puissance informatique communément disponible, seule l'étude d'un site EF à la fois est envisageable. Nous avons utilisé d'autre part un mode de solvation comparable à celui des simulations de cations en milieu aqueux. Ainsi, les systèmes que nous avons étudiés ont été construits de la manière suivante.

A partir de la structure cristallographique de la calciprotéine considérée (incluant l'eau cristallographique), une sphère de molécules d'eau de 15 Å de rayon extraite d'une boîte cubique équilibrée est placée autour du cation. Toutes les molécules d'eau ayant leur oxygène à moins de 2.3 Å d'un atome lourd (i.e., non - hydrogène) de la protéine sont éliminés, de même que tout résidu situé à plus de 15 Å du cation . Par la suite, les

¹ Géométrie tridimensionnelle de la protéine avec un ou plusieurs site occupé par Mg²⁺ .

atomes distants de plus de 11 Å du cation sont maintenus fixes , ainsi que les atomes d'oxygène des molécules d'eau situés à plus de 9 Å . Ainsi, les atomes situés dans une sphère de 9 Å de rayon autour du cation sont libres de bouger. Ils sont entourés par une barrière molle dans laquelle les molécules d'eau sont seulement libres de tourner.

2. Champ de force.

Pour la description des atomes de la protéine nous avons utilisé le champ de force de CHARMM22 dans lequel les groupements CH₃, CH₂, et CH sont considérés comme des atomes "étendus" (5), c'est-à-dire que les atomes d'hydrogène de ces groupements sont traités de manière implicite. Les molécules d'eau du système sont décrites selon le même modèle que précédemment, à savoir le modèle TIP3P (6) développé spécifiquement pour traiter les interactions protéine - eau ; les paramètres correspondant sont donnés, par exemple, dans Alary et al.(7). Enfin, les interactions cation - oxygène sont décrites de la même manière que lors de nos simulations dans l'eau, à savoir que nous avons utilisé les paramètres issus de la référence(1).

Les trajectoires de dynamique moléculaire ont été réalisées avec un cutoff de 14Å et les interactions électrostatiques et Lennard-Jones ont été lissées avec une fonction de type SHIFT- (5).

3. Minimisation et simulation de Dynamique Moléculaire .

Dans le but de libérer l'excès d'énergie potentielle qui apparaît d'ordinaire lors de la construction d'un système initial, notamment du fait qu'une structure cristallographique est une moyenne de toutes les structures présentes dans le cristal, 500 étapes de minimisation ont été exécutées, avec un algorithme de gradient conjugué

(pendant ce traitement nous avons imposé des contraintes harmoniques sur les positions des atomes de la protéine ainsi que sur les cations, avec une constante de force de 50 kcal/mole.Å²).

Pendant le déroulement de la dynamique, la liste de voisins non liés (i.e., la liste des atomes situés à moins de 15 Å de chaque atome) est « mise à jour » tous les 25 cycles d'intégration. Lors des 3 premières ps, le système est thermalisé. Puis il est équilibré pendant les 15 ps suivantes, les vitesses atomiques étant périodiquement contrôlées (toutes les 500 étapes), et réassignées quand la température moyenne est en dehors d'une fenêtre de 5°K de part et d'autre de la température prévue de 300°K.

Les longueurs de liaisons de l'eau et de la protéine ont été contraintes à conserver leur valeur d'équilibre, avec l'algorithme SHAKE, permettant ainsi d'utiliser un pas temporel de 2 fsec.

4. Calcul de différences d'énergie libre.

Au vu des résultats obtenus lors des « mutations » dans l'eau, nous avons utilisé un protocole avec $\Delta\lambda=0.1$. Ainsi, à la suite d'une période d'équilibration de 10 ps pendant laquelle $\lambda=0.05$, une série de 10 trajectoires est réalisée. Dans chaque simulation à une valeur λ donnée, le système est tout d'abord équilibré pendant une durée de 5 ps, puis un échantillon de 500 configurations est extrait d'une trajectoire de 10 ps, enfin à partir de cet ensemble les différences d'énergie libre $\Delta A(\lambda+\Delta\lambda)$ sont déterminées, pour $\Delta\lambda=-0.05$ et $\Delta\lambda=+0.05$. Le paramètre λ est alors incrémenté de $\Delta\lambda=0.1$. Le dernier point de la simulation précédente est le point de départ de la suivante, etc...

De même que dans l'eau, en partant de la valeur $\lambda=0.95$, nous avons réalisé une transformation « retour », afin d'évaluer la précision du calcul. (La variation d'énergie

libre le long du chemin $\lambda=0 \rightarrow \lambda=1 \rightarrow \lambda=0$ doit être nulle, si l'échantillonnage est parfait, et si aucune dérive systématique n'est survenue au cours de l'ensemble du calcul.)

B. Étude de la parvalbumine

1. L'échange $\text{Ca}^{2+}/\text{Mg}^{2+}$ dans le site EF - 3 de la parvalbumine.

Cette simulation "alchimique" est particulièrement intéressante dans la mesure où aucune structure tridimensionnelle de calciprotéine à Main EF totalement chargée en magnésium n'a pour le moment été déterminée expérimentalement (à notre connaissance). Nous avons donc construit notre système en partant de la structure cristallographique Pa.Ca.Mg.Mg (4pal) déterminée à une résolution de 1.75Å . Comme nous l'avons vu le système étudié inclut seulement une partie de la structure de la protéine. Ce domaine sphérique est centré sur le cation du site EF-3. La sphère de 9Å dans laquelle tous les atomes sont libres de bouger englobe complètement les chaînes latérales des résidus coordonnant le cation (résidus 50 à 63) ainsi que 56 molécules d'eau et une partie des résidus du site adjacent (résidus 96-97) .



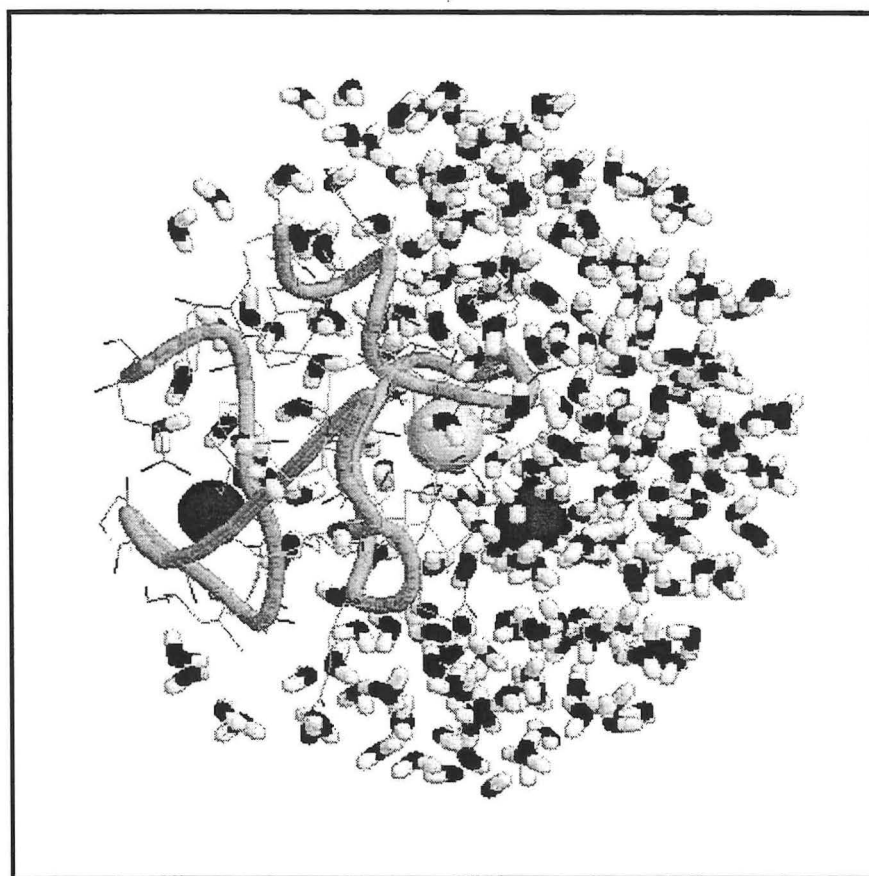
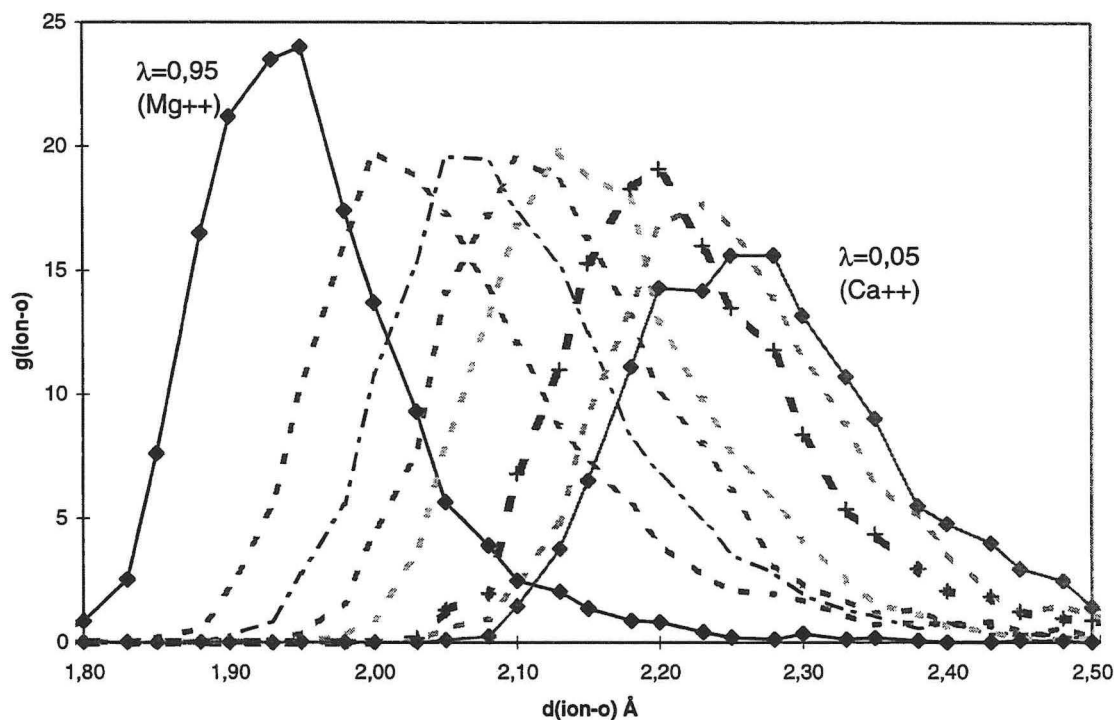


Figure II-2 : modèle utilisé lors de la simulation relative au site EF-3 (4pal) - une sphère de 15 Å Les cations sont représentés par de grosses boules, claire dans le cas du calcium. Le magnésium entouré de nombreuses molécules d'eau est celui du « troisième site ».

Compte tenu de la taille de la zone mobile, il est probable que la structure PaMgMg.Mg obtenue à la suite de la transformation du cation dans le site EF-3 ne sera pas totalement réaliste, du fait que tout le reste de la protéine est maintenu figé et ne pourra donc pas se relaxer.



courbes II-1 : évolution du premier pic rdf cation-O en fonction de λ , dans le site EF-3 de la parvalbumine.

Dans ce cas, les distributions radiales cation - oxygène (ci-dessus) sont peu intéressantes, car, elles ne permettent de déterminer que le nombre mais pas la nature des atomes coordonnant le cation.

a) Analyse structurale

(1) Evolution des distances cation - oxygène

Nous avons donc suivi a) l'évolution des distances cation - oxygène en fonction du paramètre de transformation λ et b) l'évolution des trois angles de torsion χ_1 - χ_2 χ_3 du Glu62.



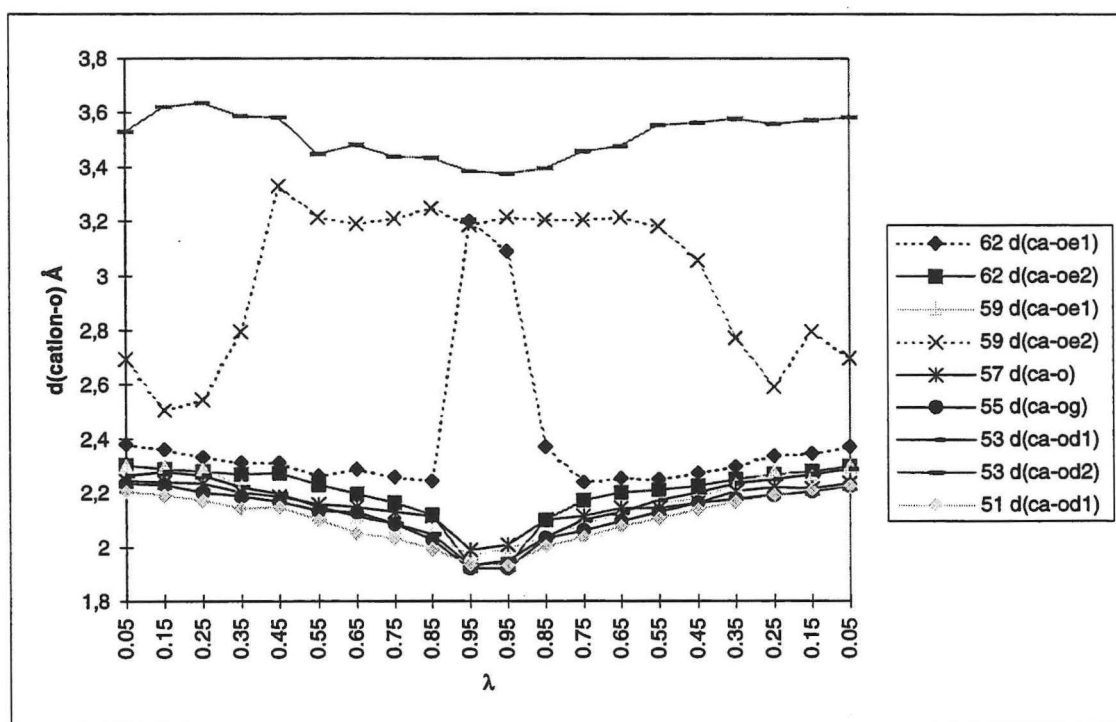


Figure II-3 : Transformation $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+} \rightarrow \text{Ca}^{2+}$ dans le site EF-3 de la parvalbumine : Distance moyenne cation- oxygène en fonction de λ ($\lambda=0$ correspond à un cation dans l'état 'Ca' , tandis que $\lambda=1$ correspond à l'état 'Mg'). Le Glu 62 est dans la position relative 12 du motif main EF.

A $\lambda=0,05$, sept oxygènes sont en moyenne à 2,30Å du cation. Cette valeur est nettement plus basse que la distance moyenne de 2,43Å observée dans la structure cristallographique. De plus, un huitième oxygène, à 2,7Å, est aussi nettement plus proche que dans la structure cristallographique (3,5Å). ceci traduit certainement l'inadéquation de la description de l'interaction cation-oxygène du site de la protéine. Une topologie plus « normale » à sept oxygènes autour du cation est clairement rétablie lorsque λ atteint 0,45. Cette topologie persiste jusque $\lambda=0,85$. De plus, une diminution générale des distances cation-oxygène est observée au fur et à mesure de



l'incrémentation de λ , avec la diminution du rayon ionique du cation², et ce, de manière analogue à ce qui a été observé lors d'études cristallographiques de parvalbumines substituées par des cations divalents de rayon ionique différent: Ca^{2+} , Mn^{2+} , Mg^{2+} (2). Pendant la dernière étape de la simulation ($\lambda=0.95$), L'oxygène OE1 du résidu Glu 62 sort de la sphère de coordination, se maintenant à une distance supérieure à 3Å du cation central, Mg^{2+} , alors que simultanément, l'oxygène OE2 de Glu 62 devient plus proche du cation. Cet événement correspond au fait que Glu 62 devient un ligand monodentate.

A ce stade, le cation est hexacoordiné, et ce en accord avec les données cristallographiques concernant l'autre site de fixation, le site EF4 de la structure Pa.Ca.Mg.Mg. Néanmoins, les distances moyennes Mg^{2+} - oxygène obtenues (1.90Å) sont plus courtes que celles observées expérimentalement dans EF-4 (2.11Å). Ce fait est aussi probablement imputable au traitement de l'interaction cation - oxygène dans notre champ de force.

Par ailleurs, si on en juge par le caractère plutôt symétrique de l'évolution des distances cation - oxygène (Figure II-3), au cours des transformations « aller » et « retour », l'adaptation de la protéine semble être similaire et on peut donc penser que les résultats obtenus sont relativement indépendants de la façon dont sont mises en place les conditions initiales de chacune des simulations, à λ donné.

² Nous avons vérifié (numériquement) que la distance d'équilibre d'un système constitué d'un cation et d'un atome d'oxygène évolue quasi linéairement en fonction de λ .

(2) Evolution des dièdres de Glu 62

Nous avons pu voir que le changement le plus conséquent lors de la transformation $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$ implique le résidu Glu 62. Ce dernier passe d'un état bidentate à monodentate. Cependant au simple vu de l'évolution des distances cation - oxygène, il ne nous est pas possible de préjuger de la façon dont s'effectue ce passage. C'est pourquoi nous avons aussi tracé les variations des trois angles de torsion χ_1 - χ_2 et χ_3 de Glu 62 en fonction de λ .

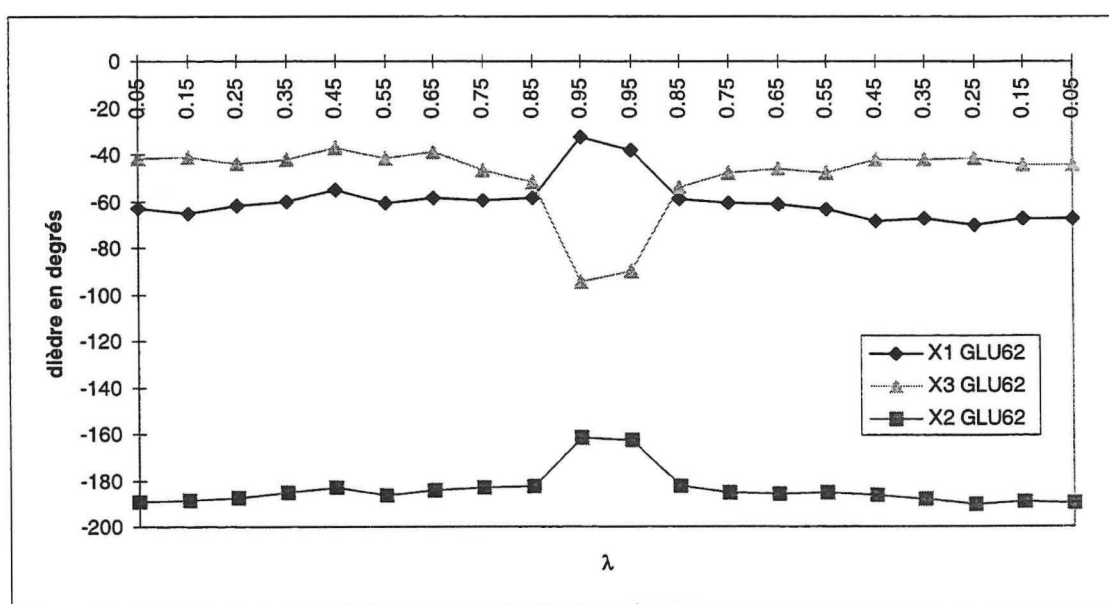


Figure II-4 : valeurs moyennes des trois angles de torsion de la chaîne latérale du Glu 62, pour chaque valeur de λ ($\lambda=0$ correspond à un cation dans l'état ' Ca^{2+} ', tandis que $\lambda=1$ correspond à son état ' Mg^{2+} ').

Comme on peut le voir ci-dessus, ces trois angles restent pratiquement inchangés pour des valeurs de λ comprises dans l'intervalle 0.05-0.85. Une variation marquée survient lorsque λ atteint 0.95, c'est-à-dire quand Glu 62 devient un ligand monodentate du cation : χ_1 varie d'environ -60° à -30° , χ_2 de -190° à près de -160° et χ_3 de -40° à -90° .

Bien que nous ayons montré que, dans le site EF-3, Glu 62 devient monodentate en présence de Mg^{2+} , le mécanisme de cette transition semble différent de celui déduit de la comparaison des structures PaCaCa.NH₄ (1pal) ou PaCaCa.Mg (3pal) avec Pa.Ca.Mg.Mg (4pal). En effet, la substitution de Ca^{2+} par Mg^{2+} dans le site EF-4 implique une variation de χ_1 plus importante, χ_2 restant quasiment invariant, et une variation significative de χ_3 . En fait, les valeurs χ_1 - χ_3 de Glu 62 que nous avons obtenu lors de l'étape $\lambda=0.95$ sont très proches de celle obtenue pour Glu 101, quand le site EF-4 est occupé par Mn^{2+} (2). (ce cation a un rayon ionique intermédiaire entre ceux du calcium et du magnésium)

Comme nous l'avons vu, chaque étape de la transformation est exécutée avec une période d'équilibration de 5 ps suivie par une trajectoire d'échantillonnage de 10 ps. Il est possible qu'à l'issue de ces périodes le système ne soit pas assez équilibré. De même, il est possible que le comportement du système à $\lambda = 0.95$ soit différent de celui à $\lambda=1.0$ (le cation central dans le site EF-3 est alors un "pur" Mg^{2+}). Pour tester ces éventualités, nous avons donc prolongé la simulation à $\lambda=0.95$ par une trajectoire de dynamique moléculaire de 100 ps à $\lambda=1.0$, et ce, en démarrant du point atteint à la fin de l'étape $\lambda=0.95$. Lors de cette trajectoire, les angles de torsion du Glu 62 ont un comportement particulièrement intéressant (cf. Figure II-5).

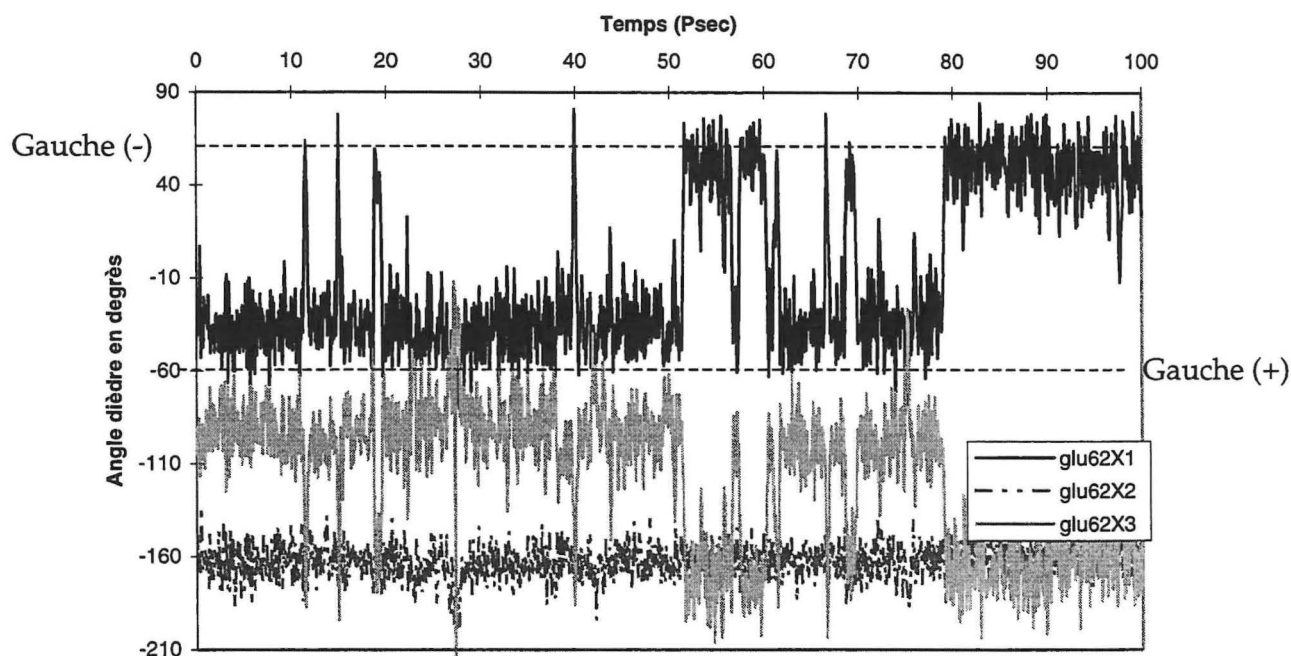


Figure II-5 : Angle dièdre de la chaîne latérale de glu62, durant une trajectoire de prolongation de 100 ps à $\lambda=1.0$ (état 'Mg²⁺').

L'angle χ_2 reste stable tout au long de la trajectoire (valeur moyenne : -170°), mais il n'en est pas de même pour les angles χ_1 et χ_3 . Dans la première moitié de la trajectoire, la tendance générale est très similaire à ce qui est observé lors de l'étape à $\lambda=0.95$. On note cependant quelques brèves échappées, concertées, des angles χ_1 et χ_3 . Dans la deuxième moitié de la trajectoire, les variations des angles χ_1 et χ_3 sont plus durables, χ_1 passant de -40° à $+50^\circ$, et χ_3 de -90° à -180° . Lors de cette trajectoire de prolongation, il semble donc que Glu 62 soit en équilibre entre deux états, l'un correspondant à celui observé durant l'étape à $\lambda=0.95$ et la quasi totalité des 50 premières ps de la trajectoire à $\lambda=1.0$, et l'autre étant très proche de celui observé pour le Glu 101, dans la forme chargée en magnésium du site EF-4. Comme notre système est limité à une sphère autour du site EF-3, et que l'on peut penser que la prise en compte de la totalité

de la protéine serait nécessaire pour permettre une meilleure adaptation au changement de cation dans le site, nos résultats suggèrent fortement que lors de l'échange $\text{Ca}^{2+} \leftrightarrow \text{Mg}^{2+}$, le site EF3 de la parvalbumine subit un réarrangement conformationnel très similaire à celui observé expérimentalement dans le site EF-4.

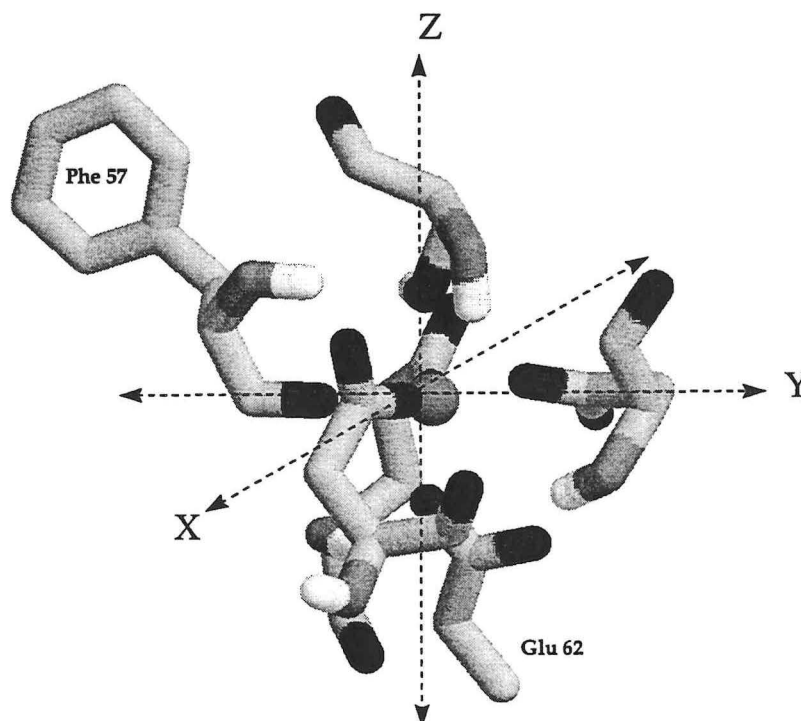


Figure II-6 : site EF-3 de 4pal après transformation du calcium. Coordonnées issues de la fin de la trajectoire de 100 Ps.

On peut avoir une idée de la géométrie du site EF-3 occupé par Mg^{2+} , via la Figure II-6 obtenue à partir de coordonnées extraites de la fin de la trajectoire de 100 ps exécutée avec $\lambda=1.0$. Pour les valeurs $\chi_1=+50^\circ$, $\chi_2=-170^\circ$ et $\chi_3=-180^\circ$ de la chaîne latérale du Glu 62, les oxygènes du site forment, autour du cation central un octaèdre assez régulier. Cette géométrie est comparable à celle du site EF-4, quand il est occupé par Mg^{2+} .

A ce stade, en ce qui concerne la flexibilité de la chaîne latérale du Glu 62 observée lors de la trajectoire à $\lambda=1.0$, nous ne savons pas si ce phénomène est spécifiquement lié à



la présence du Mg^{2+} . C'est pourquoi nous avons réalisé une autre trajectoire de 100 ps à partir, cette fois, de Pa.Ca.Mg.Mg, à $\lambda=0.0$, qui correspond donc à la forme chargée en calcium du site EF-3.

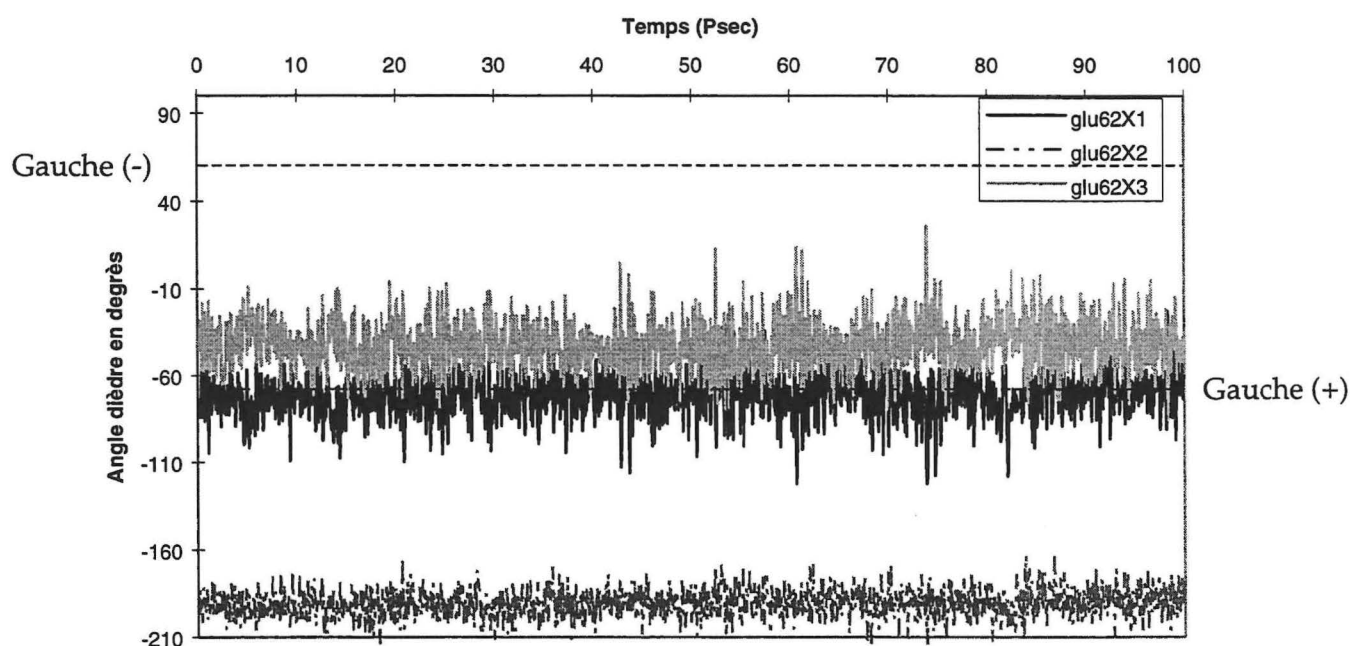


Figure II-7 : Angles dièdres de la chaîne latérale du Glu 62 lors d'une trajectoire de 100 ps à $\lambda=0.0$ (forme Ca^{2+})

Comme on peut le voir sur les courbes ci-dessus, durant toute la trajectoire, un seul état conformationnel est observé pour Glu 62. Ce dernier est similaire à celui observé lors de l'étape $\lambda=0.05$ du calcul d'énergie libre. Ainsi, il semble qu'un équilibre conformationnel de la chaîne latérale Glu 62 soit peu vraisemblable dans la forme occupée en calcium. Notons que ceci pourrait être contrôlé au niveau expérimental³.

³ Par l'étude, en fonction de la température, de la variation des constantes de couplage RMN relatives à l'angle χ_1 .

b) Rapport $K_{Ca^{2+}/KMg^{2+}}$

La transformation $Ca^{2+} \rightarrow Mg^{2+}$ réalisée dans le site EF3 de 4pal nous a permis d'obtenir les quantités d'énergie libre suivantes :

| | $\Delta A_{Ca^{2+} \rightarrow Mg^{2+}}$ | $\Delta A_{Mg^{2+} \rightarrow Ca^{2+}}$ | $\Delta A_{moyenne}$ | hysteresis |
|------------------|--|--|----------------------|------------|
| Site EF-3 (4pal) | -73.8 | 72.8 | -73.3 | 1 kcal/mol |

Tableau II-1 les deux première colonnes sont relatives aux calculs d'énergie libre aller et retour. Ensuite est reporté la moyenne entre les deux calculs et l'hystérésis(en kcal/mole).

Ainsi, entre la transformation dans l'eau, soit -77,7 kcal/mol et dans le site nous avons obtenu une différence, $\Delta(\Delta A)$ de -4.4 ± 0.5 kcal/mol, ce qui correspond, à température ambiante, à un rapport de constantes d'affinité $k_{Ca^{2+}}/k_{Mg^{2+}}$ d'environ 10^3 . Ce rapport est en bon accord avec les données expérimentales (entre 10^3 et 10^4). Nous discuterons de la pertinence de ce résultat plus loin. Néanmoins, il suggère que les paramètres et la fonction énergie potentielle utilisés dans notre étude sont capables de conduire, malgré leurs imperfections, à une assez bonne description des caractéristiques énergétiques d'une main EF.

2. Échange Mg^{2+}/Ca^{2+} dans le site EF-4 de parvalbumine.

Nous avons ensuite pris la structure cristallographique (4pal) Pa.Ca.Mg.Mg comme point de départ d'une simulation de l'échange $Mg^{2+} \rightarrow Ca^{2+}$ dans le site EF-4. Cet autre calcul d'énergie libre est destiné à valider les résultats obtenus avec le site EF-3 de la parvalbumine, puisque la structure qu'on obtiendra en fin de transformation est déjà connue expérimentalement. Il s'agit de la structure cristallographique PaCaCa.Mg, (3pal). De même que précédemment nous avons suivi l'évolution structurale du site.

101 devient aussi un ligand bidentate, si bien que le nombre de coordination du cation passe à 8.

La transformation retour $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$ est tout d'abord caractérisée par un échange d'atomes d'oxygène dans la sphère de coordination, qui survient durant l'équilibration de l'étape $\lambda=0.95$ (état Ca^{2+}) : tandis que Asp 90 devient bidentate, Asp 92 devient lui, monodentate. Ainsi, le nombre de coordination du cation reste égal à 8, mais la nature des atomes au contact du cation est maintenant différente. A $\lambda < 0.25$, les deux résidus Asp90 et Glu 101 basculent de la configuration bidentate à monodentate, si bien que, en accord avec les données expérimentales, le cation redevient hexacoordiné pour une valeur de λ proche de zéro, c'est-à-dire, lorsque le cation est dans un état Mg^{2+} .

On peut noter que Asp94, en position relative 5, reste un ligand monodentate tout au long de l'aller retour ($\text{Mg}^{2+} \rightarrow \text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$). De plus, comme on pouvait s'attendre, durant le passage $\text{Mg}^{2+} \rightarrow \text{Ca}^{2+}$, l'ensemble des distances cation-oxygène de la sphère de coordination augmente progressivement, passant d'une valeur moyenne de 1.9Å, à une valeur moyenne de 2.3Å. De même que dans le calcul précédent, ces distances sont plus courtes que celles déterminées expérimentalement.

Pour le moment, l'instabilité de la configuration des oxygènes autour du Ca^{2+} lors des étapes à $\lambda=0.95$ n'est pas claire, mais il est probable qu'elle reflète, de même que la présence d'un huitième oxygène dans la sphère de coordination de Ca^{2+} , l'inadéquation de la fonction énergie potentielle que nous avons utilisée pour traiter l'interaction cation - oxygène.

(1) Évolution des distances cation - oxygène

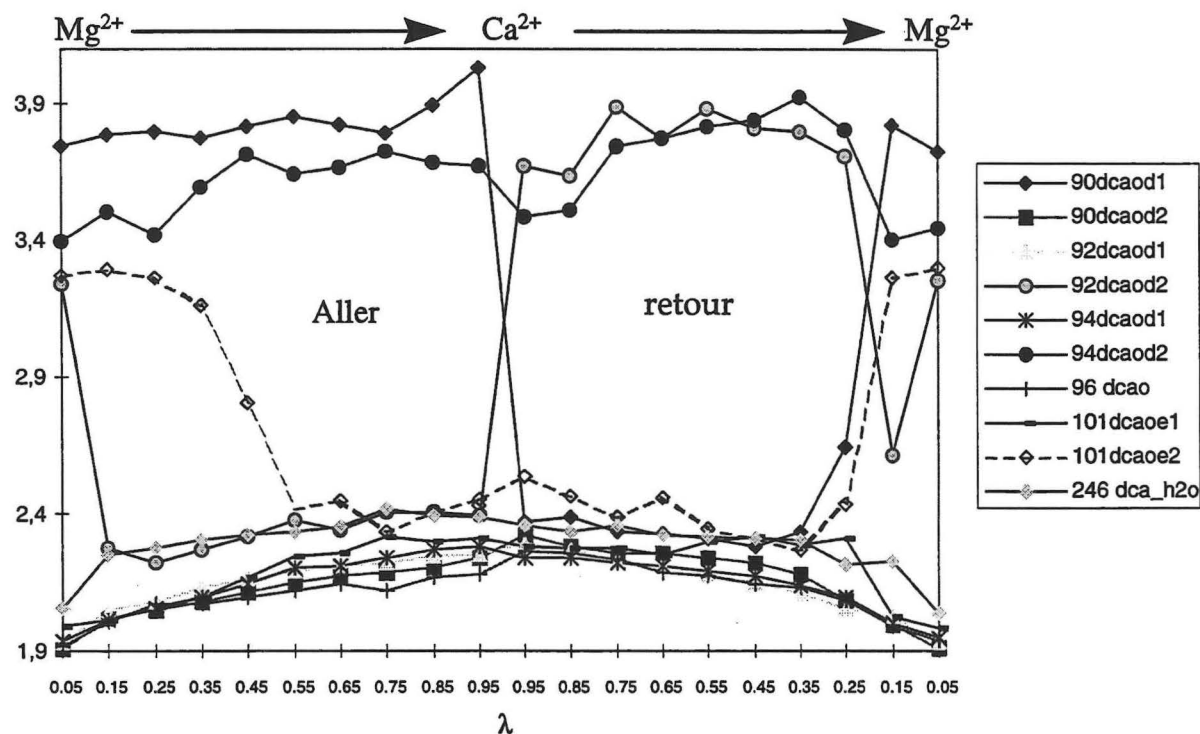


Figure II-8 : transformation $Mg^{2+} \rightarrow Ca^{2+}$ dans le site EF-4 de la parvalbumine. Distance Moyenne cation-oxygène.

Comme on peut le voir ci-dessus, à $\lambda=0.05$ («état Mg^{2+} ») le cation est hexacoordiné - ses ligands étant Asp 90 (monodentate), Asp 92 (monodentate), Asp 94 (monodentate), Met 96 (C=O chaîne principale), une molécule d'eau (N° 246), Glu 101 (monodentate) en position relative 12 du motif Main EF. Pendant la transformation $Mg^{2+} \rightarrow Ca^{2+}$, dès que $\lambda=0.15$, Asp92 devient un ligand bidentate. Ce comportement n'est probablement pas pertinent. En effet, dans toutes les structures connues de parvalbumines chargées en calcium, Asp 92 se comporte comme un ligand monodentate. Pourtant, lors de cette simulation, les deux atomes d'oxygène de Asp 92 restent proches du cation central jusqu'à la fin de la transformation, c'est-à-dire jusqu'à l'état Ca^{2+} ($\lambda=0.95$). A $\lambda=0.55$, Glu

(2) Évolution des dièdres de Glu 101

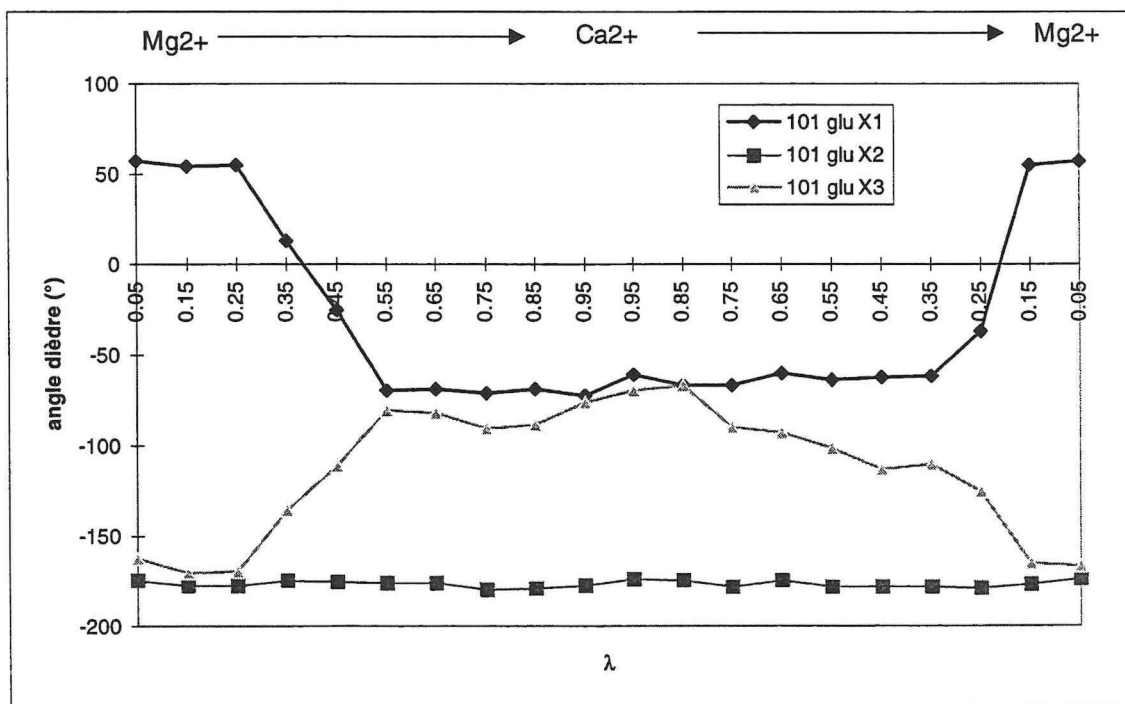


Figure II-9 : valeurs moyennes des trois angles de torsion de la chaîne latérale du Glu 101, pour chaque valeur de λ ($\lambda=0$ correspond à un cation dans l'état 'Mg', tandis que $\lambda=1$ correspond à son état 'Ca'). Le Glu 101 est dans la position relative 12 du motif main EF.

Comme on peut en juger dans la Figure II-9, le comportement de la chaîne latérale du Glu 101, lors des transformations « aller » et « retour », est assez symétrique et ce bien que les distances moyennes cation - oxygène de la Figure II-8 évoluent de manière moins symétriques que dans la Figure II-3, correspondant à l'échange $\text{Ca}^{2+}/\text{Mg}^{2+}$ dans le site EF-3. Ainsi, χ_2 reste à peu près constant tout au long de la simulation, à une valeur proche de -180° , qui correspond à un rotamère trans de la liaison $\text{C}\beta\text{-C}\gamma$. Dans la forme initiale chargée en Mg^{2+} , $\chi_1=57^\circ$ correspond au rotamère gauche(-) de la liaison $\text{C}\alpha\text{-C}\beta$, tandis qu'au bout de la transformation $\text{Mg}^{2+} \rightarrow \text{Ca}^{2+}$, $\chi_1=-65^\circ$, ce qui correspond à un rotamère gauche(+), χ_3 passant simultanément de -165° à -90° . Ces résultats sont très

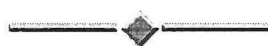
proches des valeurs que l'on trouve dans les structures cristallographiques des formes chargées en calcium PaCaCa.Mg (3pal) et PaCaCa.NH₄ (1pal) (cf. tableau ci-après).

| Valeurs expérimentales | χ_1 | | χ_2 | | χ_3 | |
|---|----------|--|----------|--|----------|--|
| 1pal→2pal Site (EF-3): Ca ²⁺ → Mn ²⁺ (Glu 62) | | | | | | |
| 1pal→2pal Site E.F-4 Ca ²⁺ → Mn ²⁺ (Glu 101) | | | | | | |
| 1pal→4pal Site (EF-4): Ca ²⁺ → Mg ²⁺ (Glu 101) | | | | | | |

Tableau II-2 : récapitulatif de la géométrie expérimentale de la chaîne latérale du résidu Glu en position 12 dans les motifs main EF de la parvalbumine. Tableau réalisé à partir de la référence (2).

| Valeurs calculées | χ_1 | | χ_2 | | χ_3 | | | |
|---|----------|--|----------|--|----------|--|--|--|
| 4pal Site EF-3: Ca ²⁺ → Mg ²⁺ (Glu 62) | | | | | | | | |
| 4pal Site EF-4 : Ca ²⁺ ← Mg ²⁺ (Glu 101) | | | | | | | | |

Tableau II-3 : récapitulatif de la géométrie de la chaîne latérale du résidu Glu en position 12 dans les motifs main EF de la parvalbumine lors de nos simulations.



(3) Rapport KCa^{2+}/KMg^{2+}

| | $\Delta A_{Mg^{2+} \rightarrow Ca^{2+}}$ | $\Delta A_{Ca^{2+} \rightarrow Mg^{2+}}$ | $\Delta A_{moyenne}$ | hysteresis |
|------------------|--|--|----------------------|---------------|
| Site EF-4 (4pal) | 81.3 | -77.1 | -79,2 | -4.2 kcal/mol |

Tableau II-4 les deux premières colonnes sont relatives aux résultat d'énergie libre aller et retour. Ensuite est reporté la moyenne entre aller retour et l'hystéresis.

Bien que les caractéristiques structurales observées pendant le calcul d'énergie libre soient plus proches des données expérimentales que dans le calcul précédent (cf. Tableau II-1), la valeur de la différence d'énergie libre obtenue semble en désaccord avec les données expérimentales puisqu'on trouve que l'affinité du site EF-4 pour Mg^{2+} est supérieure, voire au mieux égale, à son affinité pour Ca^{2+} . ($\Delta\Delta A = 1.5 \pm 2.1$ kcal/mol). Ce résultat peut sembler paradoxal. Néanmoins, la possibilité que ce résultat soit aussi pertinent ne doit pas être écartée car il peut permettre d'expliquer pourquoi la forme Pa.Ca.Mg.Mg a pu être purifiée et cristallisée. En effet, si le site EF-4 de la parvalbumine est occupé par Mg^{2+} dans toutes les mailles cristallines de cette forme - le site EF-3 étant occupé par Ca^{2+} - cela signifie que le rapport des constantes d'affinité kCa^{2+}/kMg^{2+} est plus grand pour le site EF-3 que pour le site EF-4, au moins dans les conditions de cristallisation de la forme Pa.Ca.Mg.Mg⁴. On peut notamment penser que cette différenciation des propriétés des deux sites Main EF de la parvalbumine survient en fonction de l'occupation du site secondaire ("le troisième site" de la parvalbumine est en l'occurrence occupé par Mg^{2+}). Afin de tester cette hypothèse et de vérifier que les résultats structuraux précédemment décrits ne sont pas trop dépendants du système choisi, nous avons réalisé de nouveaux calculs à partir de formes, dont les sites primaires

⁴ Les conditions de cristallisation font intervenir des concentrations de Mg^{2+} importantes.

sont tous deux occupés par Ca^{2+} , PaCaCa.Mg (3pal) et PaCaCa.NH₄ (1pal), le troisième site étant occupé respectivement par Mg^{2+} et NH_4^+ .

L'ensemble des résultats d'énergie libre obtenus figurent dans le tableau ci-dessous.

| | site EF-3 | site EF-4 |
|--|--|--|
| <i>Structure de départ</i> | <i>structure d'arrivée / type de transformation</i> | <i>structure d'arrivée / type de transformation</i> |
| Pa.Ca.Mg.Mg (4pal) $\Delta(\Delta A)$ en kcal/mole | Pa.Mg.Mg.Mg $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$ 4.4 ± 0.5 | Pa.Ca.Ca.Mg $\text{Mg}^{2+} \rightarrow \text{Ca}^{2+}$ -1.5 ± 2.1 |
| Pa.Ca.Ca.Mg (3pal) $\Delta(\Delta A)$ en kcal/mole | Pa.Mg.Ca.Mg $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$ 2.9 ± 0.0 | Pa.Ca.Mg.Mg $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$ -0.5 ± 0.0 |
| Pa.Ca.Ca.NH ₄ (1pal) $\Delta(\Delta A)$ en kcal/mole | Pa.Mg.Ca.NH ₄ $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$ 5.1 ± 0.3 | Pa.Ca.Mg.NH ₄ $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$ 6.6 ± 1.9 |

Tableau II-5 : Différences d'énergie libre ($\Delta(\Delta A)$) en kcal/mol obtenues pour les deux sites de la parvalbumine.

Comme on peut le voir ci-dessus, les différences d'énergie libre obtenues au bout de la transformation $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$ dans le site EF-3 sont relativement proches quelle que soit la structure de départ (PaCaMg.Mg, PaCaCa.Mg ou PaCaCa.NH₄). Quand aux transformations réalisées dans le site EF-4, on peut constater que la différence d'énergie libre obtenue à partir de la structure PaCaCa.NH₄ est différente des valeurs obtenues à partir des deux autres structures.

En effet, lors du calcul à partir de la structure 1pal, en accord avec les données expérimentales, les propriétés relatives à la fixation de Ca^{2+} et Mg^{2+} des deux sites Main EF de la parvalbumine sont alors similaires. Ceci est conforme à l'hypothèse mentionnée ci-dessus, à savoir que les propriétés de fixation du site EF-4 de la parvalbumine ne sont pas habituelles lorsque le troisième site est occupé par Mg^{2+} . Notons que dans ce cas deux molécules d'eau cristallographique relient le cation du troisième site à deux oxygènes des chaînes latérales du site EF-4, tandis qu'aucune molécule d'eau n'est observée quand le troisième site est occupé par NH_4^+ . Il est possible que ces deux molécules d'eau participent à la stabilisation d'une conformation légèrement différente du site EF-4, ou d'une orientation relative des deux sites Main EF légèrement différente.

Sur le plan structural, lors de l'échange $\text{Ca}^{2+}/\text{Mg}^{2+}$, le comportement dynamique du Glu en position relative 12 est affecté lors des deux simulations du site EF-3 effectuées à partir des formes totalement occupées par Ca^{2+} . En effet, il n'est pas observé de transition conformationnelle marquée, et ce malgré une prolongation de 100 ps effectuée lors de trajectoires à $\lambda=1.0$. Dans les deux simulations correspondantes du site EF-4, des modifications des dièdres sont observées seulement dans le cas de PaCaCa.NH₄, quand une prolongation de 100 ps est effectuée à $\lambda=1.0$. Elles sont cependant très différentes de celles que nous avons décrites ci-dessus.

Les angles dièdres χ_1 - χ_3 sont alors tous trois impliqués dans la mutation d'une manière analogue à celle observée lors de la comparaison du site EF-4 des structures cristallographiques PaCaNH₄⁺ et PaMnMnMn (cf. Tableau II-2). Il faut néanmoins souligner que ces quatre calculs prédisent correctement une hexacoordination pour Mg^{2+} , de même qu'un état monodentate du Glu en position relative 12 lors de l'échange $\text{Ca}^{2+}/\text{Mg}^{2+}$.



De manière générale, ces résultats suggèrent qu'il est plus difficile de prédire la topologie exacte d'un site main EF autour du Mg^{2+} , en partant d'un site occupé par un calcium, que le contraire. Ce fait est probablement lié à la légère contraction générale observée expérimentalement dans la structure de la parvalbumine lorsque son site EF-4 est occupé par Mg^{2+} (2). Dans le cadre de notre travail, ceci est regrettable puisque les structures de calciprotéines à Main EF connues actuellement sont toutes chargées en Ca^{2+} .

C. Échange $\text{Ca}^{2+}/\text{Mg}^{2+}$ dans d'autres protéines à Main EF.

1. Troponine C

Nous avons effectué la transformation du calcium présent dans le site EF-3 de la troponine C (Structure 1top). Initialement, Ca^{2+} est hepta-coordonné avec respectivement les oxygènes OD1 des résidus Asp106, Asn108, Asp110 ainsi que le $\text{C}=\text{O}$ de Phe112, une molécule d'eau ($\text{N}^{\circ}176$) dans la direction $-\text{X}$ et enfin les deux oxygènes de Glu 117 (bidentate) dans la direction $-\text{Z}$.

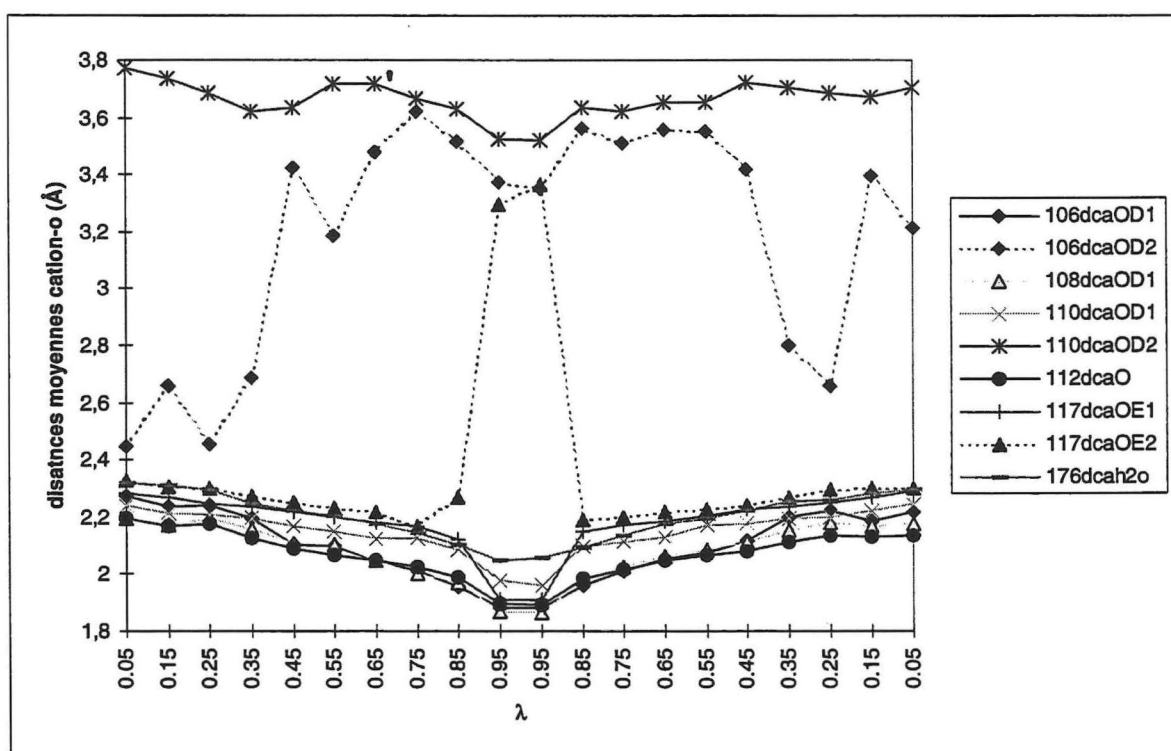


Figure II-10 : Transformation $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+} \rightarrow \text{Ca}^{2+}$ dans le site EF-3 de la troponine C : Distance moyenne cation- oxygène en fonction de λ ($\lambda=0$ correspond à un cation dans l'état 'Ca', tandis que $\lambda=1$ correspond à son état 'Mg'). Glu 117 est dans la position relative 12 du motif main EF-3.

Comme on peut le voir ci-dessus, les transformations aller et retour sont assez symétriques. Le profil général est très proche de celui observé pour le site EF-3 de la parvalbumine (cf. Figure II-3). L'homologue du Glu 62 de la parvalbumine est dans la

troponine C le Glu117. Ce dernier effectue une transition de l'état bidentate (lorsque le site est occupé par Ca^{2+}) à l'état monodentate (lorsque le site est occupé par Mg^{2+}).

Les valeurs moyennes des dièdres de la chaîne latérale de Glu117 à l'états calcium et magnésium figurent dans le tableau suivant :

| Valeurs calculées | χ_1 | | χ_2 | | χ_3 | |
|---|----------|------|----------|-------|----------|------|
| 1Top Site EF-3 : $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$ (GLU 117) | | | | | | |
| | -69° | -52° | 180° | -147° | -31° | -79° |

Tableau II-6

A la fin de la transformation $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$, l'état conformationnel du Glu 117 est très proche de celui obtenu au même point, pour le Glu 62 de la parvalbumine (cf. Figure II-4). Aucune transition gauche(+) \rightarrow gauche(-) n'est observée pour χ_1 .

En ce qui concerne l'énergie libre, nous avons obtenu une différence entre la transformation dans l'eau et dans la protéine de -4.5 ± 0.5 kcal/mol soit un rapport $K_{\text{Ca}^{2+}}/K_{\text{Mg}^{2+}}$ de l'ordre de 10^4 . Selon notre calcul, ce site serait donc plus affine pour Mg^{2+} que pour Ca^{2+} , ce qui n'est pas raisonnable. Nous avons vu qu'expérimentalement ce rapport est en fait de l'ordre de 10^3 .

2. Calmoduline

Nous avons, ensuite, effectué la transformation du calcium présent dans le site EF-1 de la calmoduline humaine (Structure 1cll). Initialement, Ca^{2+} est hepta-coordonné avec respectivement les oxygènes OD1 des résidus Asp20, Asp22, Asp24 ainsi que le $\text{C}=\text{O}$ de Trh26, une molécule d'eau ($\text{N}^{\circ}179$) dans la direction $-\text{X}$ et enfin les deux oxygènes de Glu 31 (bidentate) dans la direction $-\text{Z}$.

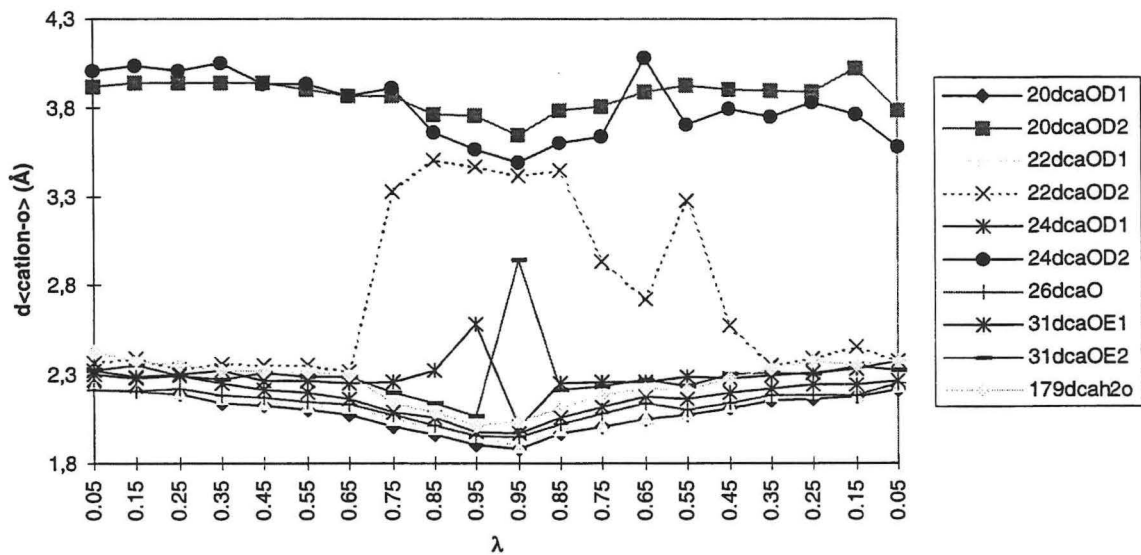


Figure II-11: Transformation $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+} \rightarrow \text{Ca}^{2+}$ dans le site EF-1 de la calmoduline humaine (1cll) : Distance Moyenne cation- oxygène en fonction de λ ($\lambda=0$ correspond à un cation dans l'état 'Ca', tandis que $\lambda=1$ correspond à son état 'Mg'). Glu 31 est dans la position relative 12 du motif main EF.

Comme on peut le voir ci-dessus, les transformations aller et retour sont assez dissymétriques et ce malgré le fait que nous avons doublé la durée des périodes d'équilibration de chaque étape de la transformation. L'homologue du Glu 62 du site EF-3 de la parvalbumine est dans la calmoduline le Glu 31. Ce dernier effectue une transition de l'état bidentate (lorsque le site est occupé par Ca^{2+}) à l'état monodentate (lorsque le site est occupé par Mg^{2+}).

Les valeurs moyennes des dièdres de la chaîne latérale de Glu117 à l'états calcium et magnésium figurent dans le tableau suivant :

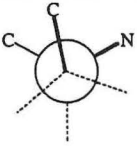
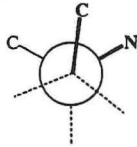
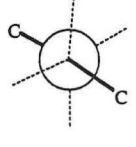
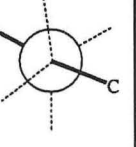
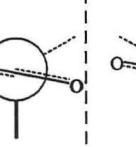
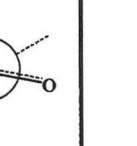
| Valeurs calculées | χ_1 | | χ_2 | | χ_3 | |
|--|---|---|---|---|---|---|
| 1cII Site EF-1 : $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$ (GLU 31) |  |  |  |  |  |  |
| | -61° | -47° | 177° | 174° | -72° | -87° |

Tableau II-7

A la fin de la transformation $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$, l'état conformationnel de Glu 31 est proche de celui obtenu pour le Glu 62 de la parvalbumine (cf. Figure II-4). Là aussi, aucune transition gauche(+) \rightarrow gauche(-) n'est observée pour χ_1 . Les résultats montrent cependant que la transition bidentate monodentate est dans ce cas plus difficile. En effet, contrairement aux autres tentatives, celle-ci ne survient que lors de l'étape $\lambda=0.95$ de la transformation retour, ou après prolongation de la simulation à $\lambda=1.0$.

Sur le plan de l'énergie libre, nous avons obtenu une différence d'énergie libre entre la transformation dans l'eau et dans la protéine de 1.6 ± 1.7 kcal/mol soit un rapport $K_{\text{Ca}^{2+}}/K_{\text{Mg}^{2+}}$ de l'ordre de 1. Compte tenu de l'imprécision de ce calcul, ce site serait donc aussi affine pour Mg^{2+} que pour Ca^{2+} , ce qui est la encore difficile à justifier.

3. Protéine sarcoplasmique

Dans tous nos calculs, le rôle de Glu dans la position relative 12 s'est avéré primordial lors de l'échange $\text{Ca}^{2+}/\text{Mg}^{2+}$. Nous nous sommes donc intéressé à un contre-exemple, le site EF-1 de la protéine sarcoplasmique, dont la position relative 12 est occupée par Asp (Asp27). Nous avons utilisé la structure cristallographique de la protéine du ver *neresis* (2scp) déterminée à une résolution de 2.0 Å (8). Dans cette structure, le calcium est hepta-coordiné par les résidus suivants : Asp16 (monodentate), Asp18 (monodentate), Asp20 (monodentate), le C=O de Ala22, une molécule d'eau dans la direction -X et Asp27 (bidentate).

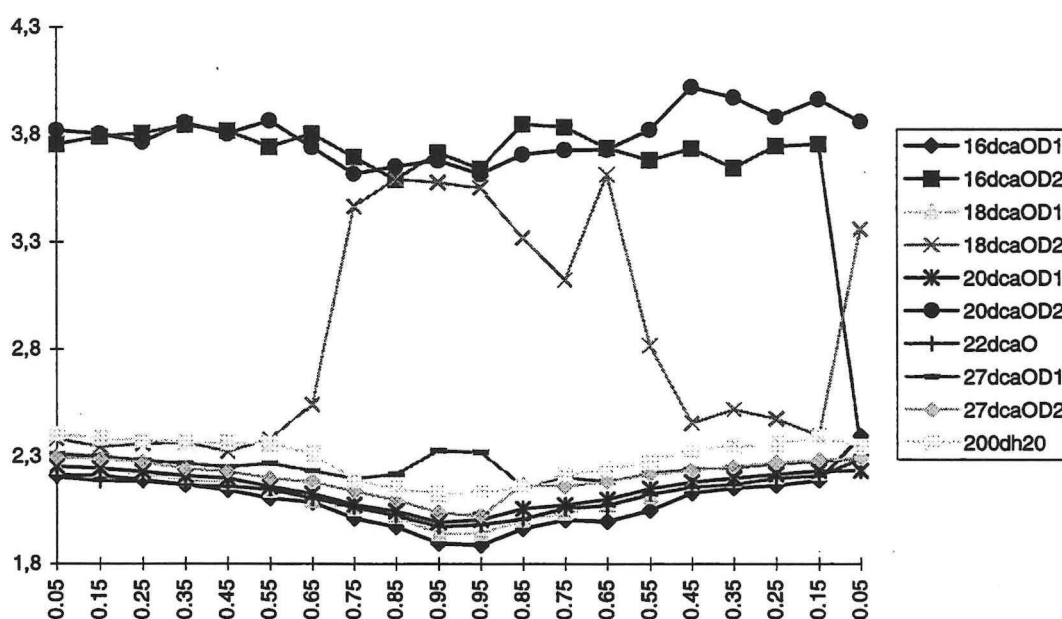


Figure II-12 : Transformation $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+} \rightarrow \text{Ca}^{2+}$ dans le site EF-1 de la protéine sarcoplasmique de ver *neresis* (2scp) : Distance Moyenne cation- oxygène en fonction de λ ($\lambda=0$ correspond à un cation dans l'état 'Ca', tandis que $\lambda=1$ correspond à son état 'Mg'). Asp 27 est en position relative 12.

Comme dans toutes les simulations précédentes, Ca^{2+} est octa-coordiné à l'aller comme au retour ($\lambda=0.05$). De plus, à $\lambda=0.95$, il n'y a pas d'hexacoordination nette du cation. Asp 27 en position relative 12 reste bidentate tout au long de l'échange

$\text{Ca}^{2+}/\text{Mg}^{2+}$, tandis que Asp 18 bascule de l'état bidentate à la configuration monodentate.

En ce qui concerne l'énergie libre, nous avons obtenu une différence d'énergie libre entre la transformation dans l'eau et dans la protéine de 2.0 ± 0.7 kcal/mol soit un rapport $K_{\text{Ca}^{2+}}/K_{\text{Mg}^{2+}}$ de l'ordre de 10. Selon notre résultat ce site serait donc, un peu plus affine pour Ca^{2+} que pour Mg^{2+} .

D. Discussion

1. Echange $\text{Ca}^{2+}/\text{Mg}^{2+}$ dans le site EF-4 de la parvalbumine.

Comme nous avons pu le voir, la substitution «alchimique» de Mg^{2+} par Ca^{2+} dans le site EF-4 nous a permis de tester la validité de notre approche, sachant que nous avons la possibilité de comparer les états Ca^{2+} ou Mg^{2+} du cation dans le site EF-4 à des données cristallographiques.

De même que l'eau, la protéine s'adapte en fonction de la taille du cation. Ainsi, comme on a pu le voir dans la Figure II-8, les distances cation - oxygène s'adaptent progressivement, les atomes se rapprochant progressivement tout en coordonnant l'atome central, au fur et à mesure que le rayon ionique diminue, à l'exception de Glu101. Pendant la transformation alchimique, celui-ci a un des atomes d'oxygène de son extrémité carboxylate qui entre (état Ca^{2+}) ou sort (état Mg^{2+}) de la sphère de coordination du cation. L'autre oxygène du carboxylate s'adapte de la même façon que les atomes d'oxygène restant constamment coordonnés au cation central. Dans le site EF-4 de la parvalbumine, la prédiction de la transition de Glu101 de ligand monodentate (état Mg^{2+}) en un ligand bidentate (état Ca^{2+}) valide apparemment notre approche. De plus, comme on peut le voir dans la Figure II-9, le réarrangement observé au niveau de

Glu101 est caractérisé par une variation des angles dièdres χ_1 et χ_3 , tandis que χ_2 reste constant, et ce en parfait accord avec les données cristallographiques. Cette « prédiction » réussie concernant la flexibilité conformationnelle de la chaîne latérale du Glu101 pendant l'échange $\text{Ca}^{2+}/\text{Mg}^{2+}$ valide notre tentative en parallèle dans le cas du site EF-3 de la parvalbumine, pour lequel seul l'état « Ca^{2+} » est connu cristallographiquement, ainsi que dans le cas d'autres protéines à main EF - (TnC et Cam), pour lesquelles aucune structure en présence de Mg^{2+} n'est connue pour le moment.

Néanmoins nous avons aussi pu mettre en évidence le fait que certaines tendances observées lors de l'analyse ne sont guère en accord avec les données expérimentales. Par exemple, Asp92 devient un ligand bidentate après une petite variation du paramètre λ lorsque l'on part de la forme occupée en Mg^{2+} . Il reste bidentate jusqu'à $\lambda=0.95$, c'est-à-dire longtemps après que Glu101 ait basculé d'une configuration monodentate à une configuration bidentate, à $\lambda=0.55$ (voir Figure II-8), si bien que dans notre simulation le calcium du site EF-4 est octa-coordiné et ce en désaccord avec toutes les structures connues, dans lesquelles Ca^{2+} est hepta-coordiné. Bien que, dans le site EF-4, huit atomes d'oxygène autour du Ca^{2+} aient été initialement décrits, dans la structure cristallographique de parvalbumine de carpe Pi4.25 (9), il est en effet maintenant parfaitement établi, sur la base de plusieurs études cristallographiques de parvalbumines à haute résolution dont les sites sont occupés par du Ca^{2+} ((10), (2), (11), (12); (13),) que Ca^{2+} est hepta-coordiné.

Enfin, notons que dans la Figure II-8 les atomes d'oxygène de Asp 90 (OD1) et Asp 92 (OD2) s'échangent pendant les étapes à $\lambda=0.95$, si bien que pendant le retour $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$ (jusqu'à $\lambda=0.35$) une octa-coordination du cation central est assurée par l'intermédiaire de Asp 90, qui est alors bidentate. Ce phénomène ne peut être justifié par

comparaison à l'expérience, dans le cas de la parvalbumine dont les sites EF-3 et EF-4 sont occupés par Mn^{2+} (structure cristallographique 2PAL (2)). Mn^{2+} ayant un rayon ionique, 0.80\AA , intermédiaire entre celui du Ca^{2+} et du Mg^{2+} , on peut s'attendre, lors de la variation de λ de l'état Ca^{2+} à l'état Mg^{2+} , à ce que le système passe par une géométrie avec les mêmes caractéristiques que celles observées dans le site EF-4 de 2PAL.

Dans cette structure les ligands Asp90 et Asp92 sont monodentates (2). Dans leur ensemble, les résultats présentés dans la Figure II-8 sont relativement en bon accord avec les changements conformationnels observés expérimentalement lors de l'échange Ca^{2+}/Mg^{2+} du site EF-4 de la parvalbumine. Le réarrangement de la chaîne latérale du Glu101 lors de cet échange est remarquablement bien prédit.

Nous avons cependant vu que les différences d'énergie libre calculées pour le site EF-4 donnent un rapport des constantes d'affinité $k_{Ca^{2+}}/k_{Mg^{2+}}$ en contradiction avec les données expérimentales. Comme le modèle que nous avons utilisé comprend seulement un sous-domaine de la protéine entouré par des atomes contraints, on pourrait penser que ce sont ces contraintes qui sont à l'origine de la faible valeur de différence d'énergie libre obtenue. Des résultats similaires sont observés lors de nos calculs à partir de la structure 3PAL (Tableau II-5). Cependant, si les calculs sont effectués à partir de la structure 1PAL, nous obtenons le résultat inverse, à savoir une valeur positive de 6.5 kcal/mol pour $\Delta\Delta A(Ca^{2+} \rightarrow Mg^{2+})$, proche de celle obtenue lors des calculs relatifs au site EF-3 ($\Delta\Delta A(Ca^{2+} \rightarrow Mg^{2+}) = 3$ à 5 kcal/mole). Comme nous l'avons vu le troisième site de 1PAL est occupé par un cation monovalent alors que dans celui des structures 3PAL et 4PAL il s'agit d'un cation divalent - Mg^{2+} . On peut donc s'interroger sur l'influence à longue distance de la nature chimique du cation dans le troisième site sur les affinités respectives du site primaire EF-4 pour Ca^{2+} et Mg^{2+} . Le troisième site est observé

seulement dans le cas des parvalbumines β . Sa présence est liée à celle d'un résidu Asp en position 61. Il est localisé dans l'environnement immédiat ($\approx 5\text{\AA}$) du site primaire EF-3. La cristallisation spécifique de la structure 4PAL de brochet 4.10, dans laquelle le site EF-4 est sélectivement remplacé par Mg^{2+} tandis que le site EF-3 est occupé par Ca^{2+} , et ce en présence d'un grand excès de Mg^{2+5} , suggère que dans ces conditions de cristallisation (occupation du troisième site par un Mg^{2+}) les deux sites primaires ont des affinités pour Ca^{2+} et Mg^{2+} différentes. Compte tenu de la distance qui les sépare, l'influence éventuelle de l'occupation du troisième site sur la capacité de fixation et de sélectivité du site EF-4 à l'égard du Ca^{2+} et du Mg^{2+} est difficile à justifier sur un plan structural ⁶. Nous pourrions cependant tester notre hypothèse en étendant notre étude au site EF-4 d'une parvalbumine sans « troisième site ». Ce calcul pourrait par exemple être réalisé à partir de la structure cristallographique de parvalbumine α de requin (11) dont la résolution est de 1.54 \AA .

2. Le site EF-3 occupé par Mg^{2+} .

Pour le moment, aucune structure tertiaire de parvalbumine n'est disponible dans laquelle les sites primaires (EF-3 et EF-4) sont tous deux occupés par Mg^{2+} . Comme on a pu le voir Figure II-6, notre approche a permis de générer la géométrie du site EF-3 occupé avec un Mg^{2+} . On doit cependant prendre en compte le fait que, dans notre étude, seulement un sous domaine de la protéine a été pris en compte et non la protéine entière. La structure totalement occupée en Mg^{2+} ainsi générée à partir de la structure 4PAL n'est donc pas une structure totalement réaliste. Cependant, nous pouvons considérer, compte tenu de ses caractéristiques géométriques, que la structure du site

⁵ des ions sulfate de magnésium ont été utilisés comme agent précipitant.

EF-3 est correcte. En effet, la transformation «alchimique» de Ca^{2+} en Mg^{2+} à partir de la structure cristallographique 4PAL satisfait pleinement les besoins de coordination de Mg^{2+} dans la main EF. Il est maintenant largement établi que le magnésium est hexacoordiné dans l'eau(14) comme au sein de petites molécules organiques (15).

Cette hexa-coordination correspond habituellement à un arrangement octaédrique des atomes d'oxygène autour du cation central. Le Glu62, qui est l'homologue de Glu101 dans le site EF-4, intervient alors comme un ligand monodentate au même titre que Asp51, Asp53 et Glu59.

Nous avons observé que la distance moyenne cation-oxygène est alors de 1.9 Å ce qui est légèrement plus court que la valeur moyenne expérimentale observée lorsque le site EF-4 est occupé par Mg^{2+} . Ce phénomène est sans doute dû à des insuffisances du potentiel utilisé pour traiter l'interaction cation-protéine.

3. Conformation de la chaîne latérale de Glu 12

Par ailleurs, un point à considérer de manière critique est le réarrangement conformationnel de la chaîne latérale du Glu 62 lors de la substitution de Ca^{2+} par Mg^{2+} dans le site EF-3. Les données cristallographiques ainsi que nos résultats sont résumés respectivement dans les tableaux I :2 et I :3. Contrairement, à la chaîne latérale de Glu101 dans le site EF-4 , tous les angles dièdre de la chaîne latérale de Glu62 varient pour s'adapter à la géométrie de la coordination de Mg^{2+} (cf. Figure II-4). Tandis que la chaîne latérale adopte la conformation gauche(+) avec $\chi_1 = -60^\circ$ dans la structure de départ 4PAL, à la fin de la transformation $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$, χ_1 a des valeurs comprises entre -30 et -40° qui correspondent à un rotamère gauche(+) de biais (aucune transition vers le

⁶ ≈ 15 Å séparent le cation du site EF-4 de celui du troisième site.

rotamère gauche (-) n'est observé). Cependant, si la trajectoire est prolongée de 100 ps, l'angle dièdre χ_1 effectue des changements rapides à $\chi_1 = 60^\circ$ puis reste stabilisé dans l'état gauche(-) pendant une période de temps suffisante pour qu'on puisse conclure que l'état gauche(-) est représentatif de la conformation de la chaîne latérale de Glu62 quand EF-3 est occupé par Mg^{2+} .

Comme on peut le voir dans la Figure II-5, le rotamère gauche(+) est systématiquement accompagné par une valeur de χ_3 de -170° tandis que la valeur de l'angle χ_2 reste en moyenne constant à près de -180° . Quant à l'angle dièdre χ_1 si le rotamère gauche(-) n'est pas l'unique état conformationnel de la chaîne latérale de Glu 62, nous devons alors considérer l'éventualité d'un mélange gauche(+) « légèrement de biais » / gauche (-). Notons que les trois angles dièdres sont fortement corrélés. Ainsi, nous aurions un premier conformère de Glu62 (quasi identique à celui observé expérimentalement pour Glu101 lorsque le site EF-4 est occupé par Mg^{2+}), avec $\chi_1 = 60^\circ$, $\chi_2 = -180^\circ$ et $\chi_3 = -180^\circ$ tandis qu'un second conformère pourrait correspondre à $\chi_1 = -30, -40^\circ$ (gauche (+) biaisé), $\chi_2 = -160^\circ$ et $\chi_3 = -100^\circ$. Aucun autre conformère ou groupe de conformères n'est observé pour Glu 62. Ceci suggère que le fait que ce résidu soit un ligand monodentate du Mg^{2+} limite le nombre de conformations possibles (s'il n'y a pas une conformation unique).

Une étude RMN de la parvalbumine de brochet 5.0 dans sa forme totalement occupée en magnésium a montré que les deux résidus Glu en position relative 12 (i.e. respectivement les Glu62 et Glu101) des sites EF-3 et EF-4 se comportent de manière symétrique concernant les états de l'angle dièdre χ_1 . Lorsque les sites sont occupés par Mg^{2+} , les deux Glu12 adoptent l'état gauche(-), alors qu'en présence de Ca^{2+} , ces derniers passent dans la forme gauche (+). (L'analyse RMN ne permet pas de distinguer

les conformères gauche(-) et trans ou le cas d'un équilibre de conformères). Dans leur ensemble, nos résultats sur la parvalbumine sont donc en accord avec les données RMN.

Si l'on considère le réarrangement conformationnel des deux chaînes latérales de Glu62 et Glu101 lors de l'échange $\text{Ca}^{2+}/\text{Mg}^{2+}$, nous pouvons donc conclure que, sur la base de nos résultats, les deux résidus qui occupent la position relative 12 doivent avoir un comportement identique, comme le suggère la présence d'un pseudo axe de symétrie d'ordre deux passant entre les motifs EF-3 et EF-4 de la parvalbumine.

4. Sites main EF d'autres calciprotéines à main EF.

Les deux sites de fixation Ca/Mg du lobe C-terminal de la troponine C (Tnc) correspondent à une paire de mains EF ayant une forte homologie structurale avec la paire de domaines main EF de la parvalbumine. Alors qu'on connaît une structure cristallographique à haute résolution de TnC dont les deux sites de ce lobe sont occupés par un Ca^{2+} , il n'y a pas l'équivalent pour la forme occupé en Mg^{2+} . Sur le plan physiologique les troponines C, dont les sites sont occupés par Mg^{2+} , sont associées au muscle au repos tandis que leurs formes chargées en Ca^{2+} sont à l'origine du déclenchement de la contraction musculaire suite à l'interaction avec d'autres protéines (cf. chapitre 1).

A partir d'une structure cristallographique résolue à 2.0 Å de la troponine C avec les deux sites EF- 3 et EF- 4 occupés par Ca^{2+} , nous avons généré une forme chargée en Mg^{2+} de la troponine C en utilisant l'approche « alchimique » qui a précédemment correctement prédit la géométrie des deux sites main EF de la parvalbumine lors de l'échange $\text{Ca}^{2+}/\text{Mg}^{2+}$. Ici encore nos simulations prédisent une hexa-coordination du Mg^{2+} au sein du site EF-3, Glu 117 en position relative 12 effectuant une transition de l'état bidentate à l'état monodentate. Le mécanisme d'adaptation de la chaîne latérale est



cependant différent de celui décrit lors de l'échange $\text{Ca}^{2+}/\text{Mg}^{2+}$ dans les sites primaires de la parvalbumine. On assiste en effet lors de ce passage monodentate $\leftarrow \rightarrow$ bidentate à des variations des trois dièdres que l'on peut comparer à celles observées expérimentalement dans la chaîne latérale du Glu 62 du site EF-3 de la parvalbumine occupé par Mn^{2+} (structure 2Pal). Un comportement similaire est aussi observé lors des simulations sur le site EF-1 de la calmoduline. En effet ici encore le rôle de l'acide glutamique en position relative 12 est essentiel lors de l'échange $\text{Ca}^{2+}/\text{Mg}^{2+}$.

Il s'agit de prédictions que nous ne pouvons confronter pour l'instant à aucune donnée expérimentale.

E. Conclusion

Nos résultats montrent clairement que l'acide glutamique invariant à la position relative 12 joue un rôle essentiel lors de l'échange $\text{Ca}^{2+}/\text{Mg}^{2+}$ dans les différents sites main EF que nous avons étudiés, puisque sa chaîne latérale adopte différents états conformationnels en fonction des besoins en coordination des deux cations Ca^{2+} et Mg^{2+} . Il est ainsi possible de coordonner ces cations à la protéine avec un minimum de modifications dans le site. Ainsi, le positionnement du cation au sein de la protéine est pratiquement invariant tandis que la chaîne latérale de Glu12 compense la différence de coordination du cation. Dans ce contexte, le répertoire des états conformationnels de la chaîne latérale de Glu 12 paraît être fortement limité. Dans le cas de la parvalbumine, les sites EF-3 et EF-4 mettent en jeu deux principaux conformères, qui diffèrent essentiellement par leur angle dièdre ϕ_1 gauche(+) et gauche(-) dans les formes respectivement chargées en Ca^{2+} et Mg^{2+} . Il est possible que l'inter conversion gauche(+)/gauche(-) ne soit pas systématiquement associée à l'échange $\text{Ca}^{2+}/\text{Mg}^{2+}$ comme le suggèrent les résultats obtenus en étudiant deux autres protéines à main EF ,

la troponine C et la calmoduline, bien que dans les deux cas Glu 12 passe de l'état monodentate à un état bidentate.

Enfin, l'étude d'une protéine à main EF contenant Asp au lieu de Glu en position 12 (site de la protéine sarcoplasmique) suggère que l'allongement de la chaîne latérale augmente les possibilités de variation conformationnelle de la position 12. En effet, Asp 12 reste à l'état bidentate malgré l'échange $\text{Ca}^{2+}/\text{Mg}^{2+}$ tandis que d'autres ligands du site main EF participent à la réorganisation imposée par les besoins en coordination différents des cations Ca^{2+} et Mg^{2+} . Là aussi la structure théorique n'a pas d'équivalent expérimental.

F. Tentative avec une nouvelle paramétrisation de l'interaction cation - oxygène

Le traitement de l'interaction cation - oxygène que nous avons utilisé (interaction effective, celle d'Åqvist) induit lors des les simulations i) une distance cation oxygène trop courte ii) une octa - coordination du calcium aussi bien dans l'eau (ce qui peut être exact) que dans les protéines (ce qui ne l'est pas). Afin de tenter d'améliorer ces points, nous avons entrepris de nouvelles simulations mettant en œuvre un autre type d'interaction effective. Cette interaction à la forme suivante :

$$V_{cation-O} = V_{elec} + \sum_i^{oxygène} \frac{A_{ion-o}}{r_{ion-o}^7} - \frac{B_{ion-o}}{r_{ion-o}^4}$$

Notons que le terme Lennard-Jones est maintenant en 7-4 (16) le terme en $1/r^4$ décrivant l'interaction charge - dipôle induit que l'on pense être importante dans ce type d'interaction. Les paramètres A_{ion-o} et B_{ion-o} sont dans le tableau suivant :

| | A_{ion-o} | B_{ion-o} |
|-------------|-------------|-------------|
| $Ca^{2+}-O$ | 6523.911 | 281.567 |
| $Mg^{2+}-O$ | 2423.231 | 212.246 |

Ces paramètres ont été obtenus de la manière suivante. Les paramètres (A_{ion-o} et B_{ion-o}) et la forme de la fonction ont été ajustés (16), afin de reproduire aux mieux l'énergie ab initio (HF) de l'interaction cation - site de la protéine pour différentes positions du cation autour de sa position d'équilibre. Le site utilisé⁷ pour cette extraction est le site EF-4 de la parvalbumine, car il s'agit du seul site dont la géométrie

⁷ Les coordonnées de ce site sont issues des structures cristallographiques 4pal (forme occupée par Mg^{2+}) et 1pal (forme occupée par Ca^{2+}) . Pour diminuer la taille du système lors du traitement ab initio, seul les atomes situés à moins de 4,5 Å du cation ont été considérés. Les liaisons rompues ont été saturées avec des hydrogènes.

expérimentale soit disponible à la fois pour les formes occupées en calcium (structure 1pal) et en Mg^{2+} (structure 4pal).

Après avoir implémenté dans le programme CHARMM22 cette nouvelle forme de potentiel, nous avons effectué : (i) des trajectoires sur les sites EF de la structure 4pal (PaCa.Mg.Mg) (ii) un échange alchimique Mg^{2+}/Ca^{2+} dans l'eau et dans le site EF-4 de la structure 4pal. Lors de ces trajectoires 1000 points ont été collectés durant 50 ps. Notons que ces 50 ps ont été précédées par 3 ps de thermalisation et 25 ps d'équilibration. Les distances moyennes cation oxygènes obtenues lors de ces simulations figurent dans le tableau ci-dessous.

| | Site EF-3 (Ca^{2+}) | Site EF-4 (Mg^{2+}) |
|-------------------|---------------------------------|---------------------------------|
| <i>Åqvist</i> | $\langle 2,30 \rangle \text{Å}$ | $\langle 1,90 \rangle \text{Å}$ |
| <i>Periole</i> | $\langle 2,43 \rangle \text{Å}$ | $\langle 2,00 \rangle \text{Å}$ |
| <i>expérience</i> | $\langle 2,43 \rangle \text{Å}$ | $\langle 2,11 \rangle \text{Å}$ |

Tableau II-8 : Distances moyennes cation-oxygène pour les différentes interaction cation-O (moyenne réalisée sur l'ensemble des oxygènes situés à moins de 3.0 Å du cation).

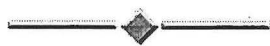
Comme on peut le voir ci-dessus, l'interaction « Periole » donne des distances moyennes Ca-O en très bon accord avec les données cristallographiques. Les distances moyennes Mg-O sont meilleures que celles obtenus avec le potentiel Åqvist, mais demeurent encore trop courtes par rapport à l'expérience ($\Delta d = 0.1 \text{ Å}$). En ce qui concerne le nombre et la nature des ligands du cation, on observe une hexacoordination autour du Mg^{2+} , et ce en accord avec les données expérimentales. Pour le calcium on constate encore une octa-coordination, du au passage à l'état bidentate de Glu 59, comme lors de l'utilisation de l'interaction développée par Åqvist.

Lors des calculs de différence d'énergie libre⁸, avec cette nouvelle forme d'interaction nous avons obtenu une différence d'énergie libre de solvation dans l'eau de -83.2 ± 0.5 kcal/mol et de -88.3 ± 0.4 kcal/mol dans la protéine, soit un $\Delta(\Delta A)$ de -5.1 ± 1 kcal/mol. Selon ce calcul le site EF-4 aurait donc une affinité pour calcium très inférieur à celle du magnésium!

Notons que la différence d'énergie libre de solvation $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$ est du même ordre de grandeur que la valeur expérimentale de -77.7 kcal/mol et ce bien que la paramétrisation ait été effectuée au sein du site de la protéine.

Bien que l'interaction « Periole » donne des distances calcium - oxygène en meilleur accord avec l'expérience, des insuffisances sont donc encore constatées. Si la détermination d'une interaction effective peut donner des résultats acceptables voire excellents dans l'eau, il semble bien que cette approche ne soit pas en mesure d'améliorer davantage nos résultats dans la protéine. Dans une prochaine étape, il sera sans doute nécessaire de tenir compte de manière explicite de la polarisabilité des chaînes latérales des acides aminés impliqués dans la coordination des cations, notamment celle des groupements carboxylate des aspartates et des glutamates. Jusqu'à présent les deux oxygènes de ces groupements étaient traités de manière équivalente (même charge partielle, notamment), que ceux-ci soient dans la configuration monodentate ou bidentate. Il est pourtant évident que la répartition des charges sur ces deux atomes doit changer d'une configuration à l'autre. On peut penser qu'une bonne description de ce changement devrait permettre de mieux décrire l'environnement du calcium dans le site des calciprotéines.

⁸ Nous avons utilisé le protocole décrit en début de chapitre, à savoir $\Delta\lambda=0.1$ (5ps d'équilibration et 10ps de trajectoire par cycle de transformation)



Références Bibliographiques

1. Aqvist, J. (1990) *J.Phys.Chem.* **94**(Iss 21), pp 8021-8024
2. Declercq, J. P., Tinant, B., Parello, J., and Rambaud, J. (1991) *J. Mol. Biol.* **220**(4), 1017-39
3. Blancuzzi, Y., Padilla, A., Parello, J., and Cave, A. (1993) *BIOCHEMISTRY* **32**(5), 1302-1309
4. Falke, J. J., Drake, S. K., Hazard, A. L., and Peersen, O. B. (1994) *Q Rev Biophys* **27**, 219-290
5. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) *J. Comput. Chem* **4**, 187
6. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) *J. Chem. Phys.* **79**, 926
7. Alary, F., and Durup, J. (1993) *J.Phys.Chem.* **97**, 13864-13876
8. Vijay-Kumar, S., and Cook, W. J. (1992) *J.Mol.Biol* **224**, 413
9. Moews, P. C., and Kretsinger, R. H. (1975) *J Mol Biol* **91**(2), 201-25
10. Declercq, J. P., Tinant, B., Parello, J., Etienne, G., and Huber, R. (1988) *J Mol Biol* **202**(2), 349-53
11. Roquet, F., Declercq, J. P., Tinant, B., Rambaud, J., and Parello, J. (1992) *J Mol Biol* **223**(3), 705-20
12. Swain, A. L., Kretsinger, R. H., and Amma, E. L. (1989) *J.Biol.Chem.* **264**, 16620-16628
13. McPhalen, C. A., Sielecki, A. R., Santarsiero, B. D., and James, M. N. (1994) *J Mol Biol* **235**(2), 718-32
14. Othaki, H., and Radnai, T. (1993) *Chem. Rev.* **93**, 1157-1204
15. Kaufman, K. A., Glusker, J. P., Beebe, S. A., and Bock, C. W. (1996) *J.A.C.S* **118**, 5752-573
16. Periole, X., Allouche, D., Daudey, J.-P., and Sanejouand, Y.-H. (1997) *résultats non publiés*

CONCLUSION





A. Protocole de simulation

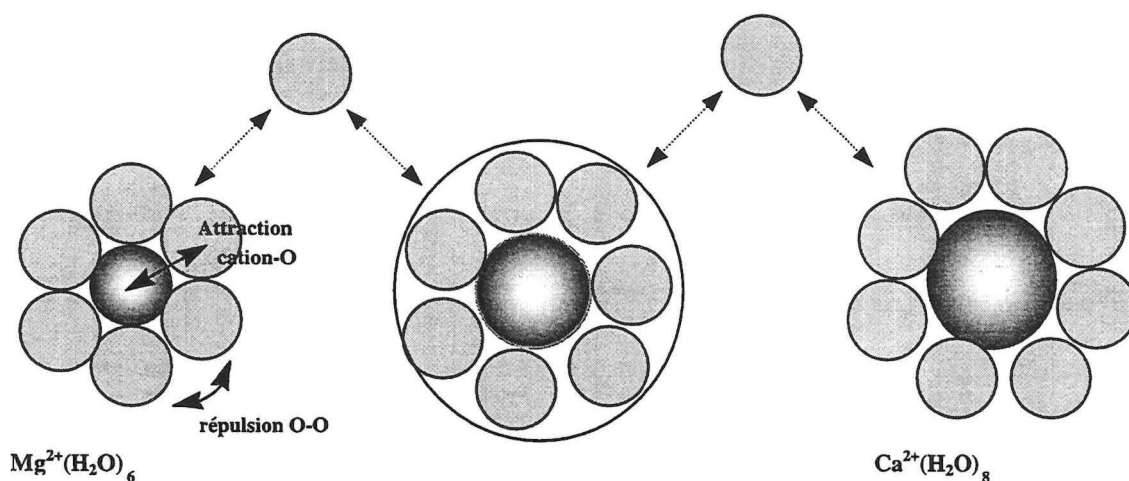
Nous avons vu lors de nos simulations de cations en solution que l'utilisation d'un petit cutoff conduit à des artefacts. Ce qui se traduit par une surévaluation de l'énergie libre calculée, ainsi que par une hystéresis plus importante, imputable probablement à une augmentation du temps de relaxation du système. La cause de tous ces phénomènes est sans doute la relative compartimentation que provoque l'utilisation d'un petit cutoff. Ainsi, les molécules d'eau constitutives des deux premières couches ressentent pleinement l'influence structurante du champ électrostatique du cation, alors qu'elles ne subissent celle des molécules d'eau plus éloignées que de manière indirecte. Compte tenu de ces éléments, il semble donc qu'il vaut mieux éviter l'emploi d'un petit cutoff lors des simulations mettant en jeu des cations en solution. Le cutoff doit être adapté de façon à ce que l'ensemble des molécules d'eau constitutives des couches de solvation du cation puissent interagir entre elles. C'est le cas lors des simulations que nous avons réalisées avec un cutoff de 14 Å.

La comparaison des ΔA calculés à un λ donné, avec des conditions initiales aussi différentes que possible, semble, être un critère beaucoup plus sensible pour estimer la qualité d'un calcul, que la simple donnée de l'hystéresis observée à la fin d'une transformation cation \rightarrow cation initial. En utilisant un critère aussi sensible, le choix d'un protocole de calcul pour un système donné ne pourra qu'être plus rigoureux.

B. Calcium et Magnésium dans l'eau

Bien qu'en milieu aqueux, nous ayons pu calculer avec une assez bonne précision la différence d'énergie libre de solvation $\text{Ca}^{2+} \rightarrow \text{Mg}^{2+}$, il est manifestement difficile de définir des règles générales pour établir un protocole de calcul de différence

d'énergie libre. Lors des calculs d'énergie libre nous avons vu que l'eau se comporte comme une matrice qui s'adapte en fonction de la taille de l'objet solvato. Cette constatation n'est pas nouvelle, néanmoins notre travail en fournit une belle illustration. Lors des trajectoires de cations dans l'eau nous avons montré que l'environnement du calcium est constitué de 8 molécules d'eau adoptant une géométrie type anti-prisme à base carrée alors que le magnésium est entouré de six molécules d'eau disposées en octaèdre. Nous avons suivi un chemin entre ces deux structures en transformant de manière alchimique le cation central. Ces simulations ont permis d'illustrer le lien, entre composition (nombre de coordination et géométrie) de la couche d'hydratation et énergie libre de solvatation. Notre étude, en accord avec un modèle comme celui de Marcus(1) montre que la différence d'énergie libre est le fruit de la réorganisation de l'environnement du cation suite à la modification de la distance d'équilibre cation-O.



Ainsi, le nombre de molécules en première couche est conditionné par la distance d'équilibre, qui est due à la taille du cation et à un équilibre subtil entre les forces attractives générées par la présence du cation et les forces répulsives entre molécules d'eau (cf. ci-dessus). Tout se passe comme si le cation jouait le rôle de clé de voûte centrale par l'intermédiaire du champ électrostatique qu'il génère et qu'en fonction de

son volume un nombre de molécules d'eau plus ou moins important peut s'en rapprocher. Lors de nos simulations, en faisant varier cette taille nous voyons augmenter ou diminuer le nombre d'oxygènes directement coordonnés au cation.

C. Calciprotéine à main EF

1. Lien structure - dynamique - fonction

Les calciprotéines à mains EF sont capables de solvater les cations calcium et magnésium aussi bien, voire mieux que l'eau. Cette propriété est la traduction au niveau moléculaire de leur capacité à fournir au cation le nombre d'oxygènes qui lui convient le mieux. Ce nombre semble de même que dans l'eau conditionné par la taille du cation ainsi que par l'équilibre entre forces attractives cation-ligand et les forces répulsives entre ligands. Nous avons vu lors de nos simulations que, de même que dans l'eau, le cation passe d'un nombre de coordination de 8 pour l'état calcium à 6 pour l'état Magnésium.

Nos résultats montrent clairement que l'acide glutamique invariant à la position relative 12 du motif main EF joue un rôle essentiel lors de l'échange $\text{Ca}^{2+}/\text{Mg}^{2+}$ dans les différents sites mains EF que nous avons étudiés, puisque sa chaîne latérale adopte différents états conformationnels en fonction des besoins en coordination des deux cations Ca^{2+} et Mg^{2+} . Cela rend la protéine capable de coordonner ces cations avec un minimum de modifications géométriques du site. Dans ce contexte, le répertoire des états conformationnels de la chaîne latérale du «Glu12» paraît être fortement limité. Dans le cas de la parvalbumine deux conformères sont observés dans les sites EF-3 et EF-4. Ceux-ci diffèrent essentiellement par l'angle dièdre χ_1 de la chaîne latérale du Glu 12 :

gauche(+) puis gauche(-) lors de la transition entre une forme occupée par Ca^{2+} et celle occupée par Mg^{2+} .

Néanmoins, il est possible que l'inter conversion gauche(+)/gauche(-) ne soit pas systématiquement associée à l'échange $\text{Ca}^{2+}/\text{Mg}^{2+}$, comme le suggèrent les résultats obtenus dans deux autres sites de protéines à main EF -, la troponine C et la calmoduline. Cependant, dans ces deux cas «Glu12» passe aussi d'un état de ligand monodentate (en présence de Mg^{2+}) à l'état de ligand bidentate (en présence de Ca^{2+}). Nous avons vu lors de l'étude d'une protéine à main EF contenant Asp au lieu de Glu à position 12 (site de la protéine sarcoplasmique) que l'allongement de la chaîne latérale via la substitution d'un méthyl dans Asp par un diméthyl dans (Glu) semble optimiser fortement la variabilité conformationnelle de la position 12, «Asp12» restant à l'état bidentate malgré l'échange $\text{Ca}^{2+}/\text{Mg}^{2+}$ tandis que d'autres ligands du site main EF - participent à la réorganisation imposée par les différents besoins en coordination des cations Ca^{2+} et Mg^{2+} .

Quoiqu'il en soit nos simulations sur les sites EF4 et EF3 de la parvalbumine sont en très bon accord avec les données expérimentales en ce qui concerne au moins la réorganisation de ces sites dans la forme chargée en Mg^{2+} . Pour les autres sites étudiés, nos résultats sont des «prédictions», puisque nous ne pouvons pour l'instant les confronter à l'expérience.

2. Compréhension de la différence d'affinité des calciprotéines pour Ca^{2+} et Mg^{2+} .

L'affinité calculée du site EF -3 de la parvalbumine est en parfait accord avec les données expérimentales. Par ailleurs, nous avons vu que dans les calculs d'énergie libre mettant en oeuvre les sites de la parvalbumine, nous avons un comportement très

différent du site EF-4, selon la nature de l'occupation du site secondaire (troisième site). En effet, lorsque ce site est occupé par Mg^{2+} , le site EF-4 est plus affine pour Mg^{2+} , tandis que lorsqu'il est occupé par NH_4^+ , ce site est plus affine pour Ca^{2+} . Nous avons émis l'hypothèse que ce phénomène était dû à une différence de conformation de l'orientation relative des sites dans les structures de départ. Ceci reste à confirmer. Des éléments de réponse pourraient être obtenus de la façon suivante :

- i) en effectuant une étude comparée de l'orientation des sites dans les structures de départ afin de mettre en évidence des différences éventuelles. (Nous n'avons rien vu qui semble significatif jusqu'à présent).
- ii) en étudiant l'influence du cation du troisième site, c'est-à-dire en effectuant, par exemple, une double transformation $PaCaCaNH_4^+ \rightarrow PaCaCaMg^{2+} \rightarrow PaCaMgMg^{2+}$
- iii) en étudiant des parvalbumines α (qui n'ont pas de troisième site)
- iv) en calculant cette différence d'affinité en prenant cette fois en compte la totalité de la protéine. Un tel calcul serait aussi une tentative pour estimer le rôle des interactions entre sites lors de l'échange Ca^{2+}/Mg^{2+} ainsi que celui de l'adaptation d'ensemble de la protéine à la taille du cation, son éventuelle rigidification, etc.

En effet, il est clair que la taille réduite du modèle de solvatation que nous avons utilisé ne permet de prendre en compte que des ajustements très localisés de la protéine. Dans le cas de la parvalbumine, une protéine globulaire, ceci est sans doute justifié et explique peut être la qualité des résultats que nous avons obtenus. Dans le cas de protéines en « forme d'haltères », telles que la calmoduline et la troponine C, des changements conformationnels de grande amplitude ont été mis en évidence, qu'il est peut être nécessaire de prendre en compte pour bien décrire l'échange Ca^{2+}/Mg^{2+} .

3. Interaction cation - oxygène

Nous avons fait le choix au départ d'utiliser pour le traitement des interactions cation - oxygène des paramètres incluant la polarisabilité de manière implicite. Nous avons utilisé deux types d'interaction (i) celle développée par Åqvist, par ajustement progressif par rapport à des données expérimentales (ii) celle développée spécifiquement pour les sites EF au laboratoire, à partir de calculs ab initio. Si les deux ensembles de paramètres sont capables de donner des résultats acceptables dans l'eau, dans le site de la protéine, on constate que les distances cation - oxygène sont trop courtes et, systématiquement qu'il y a un oxygène en trop autour du Ca^{++} , qui est alors octa-coordiné.

Pour améliorer le comportement des sites de calciprotéines lors de simulations, il pourrait donc s'avérer nécessaire de prendre en compte de manière explicite la polarisabilité, au moins en ce qui concerne les groupements interagissant directement avec le cation. Pour ce faire il serait intéressant, par exemple, de transposer le principe utilisé dans des modèles d'eau, tels MCHO (2) ou ceux de Berne (3).

i) approche MCHO

MCHO est un modèle d'eau rigide polarisable ajusté à partir de calculs ab initio. Dans MCHO, chaque atome est lié à une charge ponctuelle par un potentiel harmonique, la position des charges étant optimisée de sorte que la somme des forces que chacune subit soit nulle. Ce modèle a jusqu'à présent été utilisé lors de simulations Monte Carlo. Pour transposer ce modèle, deux difficultés importantes seront à surmonter. (i) Comment traiter les charges en dynamique ? (ii) Comment limiter le nombre et déterminer les paramètres nécessaires à la description de ce modèle ?

ii) approche Berne :

Cette fois ce n'est pas la position des charges qui est optimisée mais la valeur de la charge elle-même, en fonction du champ électrostatique qu'elle subit, et de l'électronégativité de l'atome qui la porte. Un des avantages d'un tel modèle : la valeur des charges peut fluctuer autour de la valeur qu'elles ont dans un champ de force déjà connu.

Evidemment, une combinaison de ces deux types de modèle - des charges mobiles dont la valeur change en fonction du champ électrostatique externe - serait encore plus réaliste... mais bien plus difficile à paramétrer et à mettre en œuvre. Tout d'abord, il s'agira donc, de modifier aussi peu que possible un champ de force connu, de façon à obtenir une bien meilleure description des interactions ion-biomolécules. Cette modification pouvant par la suite servir de base à l'élaboration de champs de forces de « seconde génération » (le principe des champs de force tel celui de CHARMM date du début des années 80).



Références bibliographiques

1. Marcus, Y. (1994) *Biophys Chem* **51**(2-3), 111-127
2. Saint-Martin, H., Medina-Llanos, C., and Ortega-Blake, I. (1990) *J.Chem.Phys.* **93**, 6448-6452
3. Rick, S. W., Stuart, S. J., and Berne, B. J. (1994) *J.chem.phys.* **101**, 6141-6156

RÉSUMÉ

La modulation de la concentration en calcium intervient dans bon nombre de processus, dits « Ca^{2+} dépendants », au cours desquels Ca^{2+} va entre autre interagir avec des calci-protéines à mains EF pour former des complexes qui vont activer d'autres maillons de ce qui formera une réponse biologique. Dans les conditions physiologiques, Ca^{2+} n'est pas le seul cation divalent présent dans le milieu, puisque Mg^{2+} est présent en relativement grande quantité. Une des questions centrales concernant le fonctionnement de ces calci-protéines est par conséquent la compréhension à l'échelle moléculaire de leurs différences d'affinité pour le calcium et le magnésium.

Lors de ce travail, nous nous sommes intéressés à la différence de sélectivité de diverses calci-protéines vis à vis de Ca^{2+} et Mg^{2+} . Pour ce faire, nous avons calculé des différences d'énergie libre à partir de simulations de dynamique moléculaire durant lesquelles on transforme pas à pas Ca^{2+} en Mg^{2+} et inversement. Notre premier travail a consisté à suivre l'évolution structurale de l'environnement du Ca^{2+} durant sa transformation en Mg^{2+} , et ceci, dans l'eau, et dans les sites des protéines étudiées. Les transformations progressives du Ca^{2+} en Mg^{2+} au sein des sites primaires de la parvalbumine ont permis de mettre en évidence une contraction générale de la sphère de coordination des cations, ce qui se traduit par la diminution de la distance moyenne cation - oxygène. Cette contraction s'accompagne d'une diminution du nombre d'oxygènes coordonnés au cation qui passe de 8 pour le Ca^{2+} à 6 pour le Mg^{2+} et ceci à la suite de la réorganisation ultime d'un ligand bidentate (Glu en position relative 12), qui devient monodentate. Cette réorganisation a été observée lors d'études expérimentales.

Suite aux résultats obtenus sur la parvalbumine nous sommes avons étudié la troponine C et la calmoduline, pour lesquelles aucune structure tridimensionnelle de ces protéines complexées à Mg^{2+} n'est disponible. L'analyse structurale de nos simulations a permis de montrer que ces protéines ont le même type de réorganisation structurale que celui observé précédemment.

Sur le plan thermodynamique, nos résultats se sont avérés moins satisfaisants, dans la mesure où le rapport $k_{\text{Ca}^{2+}}/k_{\text{Mg}^{2+}}$ calculé n'était pas toujours en accord avec l'expérience. C'est pourquoi parallèlement aux travaux précédents, de nouveaux ensembles de paramètres pour décrire ces cations ont été développés à partir de calculs *ab initio*.

Dynamique Moléculaire
calcul d'énergie libre

Modélisation
calcium

cations en solution
Magnésium

calci-protéine
main EF