



**HAL**  
open science

# Calcium Oxalate and Calcium Phosphate Biominerals : Formation and Stability studied by Molecular Dynamics

Théau Debroise

► **To cite this version:**

Théau Debroise. Calcium Oxalate and Calcium Phosphate Biominerals : Formation and Stability studied by Molecular Dynamics. Theoretical and/or physical chemistry. Sorbonne Université, 2019. English. NNT : 2019SORUS684 . tel-04059925

**HAL Id: tel-04059925**

**<https://theses.hal.science/tel-04059925v1>**

Submitted on 5 Apr 2023

**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.



# Sorbonne Université

École doctorale

*Laboratoire de Chimie de la Matière Condensée de Paris (LCMCP) / Équipe SMiLES*

## **Calcium Oxalate and Calcium Phosphate Biominerals: Formation and Stability studied by Molecular Dynamics**

Par Théau Debroise

Thèse de doctorat de Chimie Théorique

Dirigée par Frederik Tielens et Christian Bonhomme

Présentée et soutenue publiquement le 24/09/2019

Devant un jury composé de :

Florent BARBAULT	Maître de conférences	Rapporteur
Françoise REMACLE	Directrice de recherche	Rapporteur
Hazar GUESMI	Chercheur	Examineur
Franck DE PROFT	Professeur d'université	Examineur
Mahamadou SEYDOU	Maître de conférences	Examineur
Christian BONHOMME	Professeur d'université	Co-directeur de thèse
Frederik TIELENS	Professeur d'université	Directeur de thèse

# - Sommaire -

<b>Introduction Générale.....</b>	<b>3</b>
-----------------------------------	----------

<b>CHAPITRE 1 : Bibliographie et Méthodologie .....</b>	<b>10</b>
---	-----------

1.1. La construction de modèles moléculaires pour les matériaux .....	11
1.2. Description quantique des interactions interatomiques.....	12
1.3. Description des interactions interatomiques par un champ de force.....	24
1.4. L'optimisation de géométrie : à la recherche du minimum local .....	25
1.5. La dynamique moléculaire .....	26
1.6. Prédiction des morphologies cristallines des oxalates de calcium .....	36
1.7. Adsorption d'une catéchine sur les surfaces de l'oxalate de calcium .....	47
1.8. Étude de la nucléation des phosphates de calcium.....	53

<b>CHAPTER 2 : Morphologies of Calcium Oxalate Polyhydrates: A Quantum Chemical and Computational Study .....</b>	<b>72</b>
---	-----------

Abstract.....	73
Introduction .....	74
Methods.....	81
Results and discussion .....	84
Conclusion.....	94
Acknowledgements.....	95
References .....	95

<b>CHAPTER 3 : Effect of Catechin on the Calcium Oxalate Crystal Growth: A Molecular Dynamics Study .....</b>	<b>100</b>
---	------------

Abstract.....	101
Introduction .....	102
Acknowledgements.....	120
References .....	120

<b>CHAPTER 4 : Calcium Phosphate Nucleation: A Molecular Dynamics Study.....</b>	<b>125</b>
Abstract.....	126
Introduction .....	127
Methods.....	131
Results and discussion .....	139
Conclusion.....	146
Acknowledgements.....	147
References .....	147
<b>Conclusion Générale .....</b>	<b>151</b>
<b>Résumé.....</b>	<b>155</b>

# - Introduction Générale -

Le monde du vivant et le monde minéral sont intimement liés. Au cours d'un processus évolutif s'étalant sur plusieurs millions d'années, certains organismes ont pu développer des stratégies ingénieuses leur permettant d'utiliser, à leur avantage, les minéraux présents dans leur environnement. En orientant eux-mêmes la nucléation, ils sont capables de « confectionner » des matériaux sur mesure, aux propriétés et aux formes passionnantes. Ainsi, les diatomées utilisent les silicates pour se confectionner des carapaces poreuses de silice amorphe appelées frustules. La photosynthèse des diatomées consomme du CO<sub>2</sub> atmosphérique et joue un rôle crucial dans sa fixation<sup>1</sup>. Les frustules assurent d'une part une protection mécanique et d'autre part, permettent d'optimiser le processus de photosynthèse<sup>2</sup>.

La **biominéralisation** prend une part toute aussi importante dans la vie des vertébrés. Dans ce cas, ce sont les ions phosphates et calcium qui sont essentiels pour la génération d'un matériau également passionnant, l'os. Ces ions cristallisent pour former l'**hydroxyapatite** apportant à l'os une rigidité supplémentaire. Outre ses propriétés mécaniques, l'os constitue également un réservoir ionique permettant la régulation de la concentration de nombreux ions dans l'organisme<sup>3</sup>. La biominéralisation impacte également une grande partie du monde des plantes. Certaines d'entre elles utilisent les **oxalates** et les ions **calcium** dans la synthèse de raphides<sup>4</sup>, des cristaux leur permettant de se protéger des herbivores. Les exemples sont nombreux et démontrent la capacité de ces organismes à utiliser les ressources les entourant pour se forger une place dans le monde du vivant.

L'observation du phénomène de biominéralisation, soulève une interrogation d'ordre thermodynamique. Comment l'organisme parvient-il à déclencher la précipitation de la phase minérale alors que ce phénomène n'est pas observé dans le milieu dans lequel il évolue ? En effet, les silicates ne précipitent pas spontanément dans l'eau de mer pour former de la silice amorphe, et les phosphates de calcium ne précipitent pas non plus dans le milieu extracellulaire ou dans le sang pour former de l'hydroxyapatite. La biominéralisation est généralement un phénomène hors équilibre qui nécessite donc un apport énergétique en provenance de l'organisme vivant. Dans le cas de l'os, la cristallisation de l'hydroxyapatite est connue pour avoir lieu dans des zones spécifiques, à l'intérieur d'une matrice organique de protéines (dont

le collagène<sup>5</sup>) capable à elle seule de déclencher la croissance cristalline<sup>6</sup>. Un système de régulation extrêmement complexe est aussi mis en place par le biais de protéines inhibant et/ou promouvant la précipitation du phosphate de calcium<sup>7</sup>. Il a été démontré *in vitro*, que certaines protéines présentes dans l'os sont capables d'enclencher le processus de nucléation en augmentant la concentration locale en ions phosphate et calcium<sup>8</sup>. L'implication de vésicules dans le processus de minéralisation est observée dans le cas des diatomées<sup>9</sup>. D'une part, les vésicules peuvent jouer un rôle dans le transport de minéraux. D'autre part, il a été montré que ces vésicules sont capables de déclencher le processus de nucléation en augmentant la concentration intra-vésiculaire. La **supersaturation** semblent donc être le paramètre commun sur lequel les organismes vivants influent pour déclencher la biominéralisation de manière contrôlée et locale.

Il arrive cependant que la minéralisation échappe au contrôle de l'organisme vivant provoquant alors diverses **pathologies**. C'est le cas par exemple de la **lithiase urinaire**, une maladie provoquée par la formation et l'agrégation de cristaux à l'intérieur des reins. Des études sur la composition des **calculs rénaux** provenant de nombreux patients ont montré que les phases cristallines **whewellite** et **weddellite** de l'oxalate de calcium hydraté en sont les composants majoritaires (> 70% des échantillons analysés en Europe et en Amérique du Nord)<sup>10,11</sup>. Tout comme l'hypertension artérielle, les maladies cardiovasculaires ou le diabète, la lithiase urinaire fait partie des pathologies pour lesquelles les facteurs environnementaux jouent un rôle majeur<sup>12</sup>. L'alimentation en particulier peut à la fois provoquer ou prévenir la formation de calculs rénaux. Il a été montré par exemple qu'une consommation régulière de thé vert pouvait entraîner des changements dans la composition des calculs rénaux, favorisant la formation de weddellite au dépend de la whewellite<sup>13</sup>. De plus, il a été montré que le thé vert est capable de modifier la morphologie des cristaux *in vitro*<sup>14</sup>. Ces observations donnent aux composants du thé vert un intérêt thérapeutique majeur. En particulier, l'**epigallocatechin gallate** (EGCG)<sup>15</sup> est suspectée d'interagir directement avec les cristaux d'oxalates de calcium.

Dans le **chapitre 2**, nous prédisons par des **méthodes *ab initio*** les **morphologies cristallines** de la whewellite et de la weddellite, et proposons ainsi des modèles représentatifs des surfaces exposées. Puis, nous étudions l'interaction de ces surfaces avec l'eau par **dynamique moléculaire** (DM) classique.

Par la suite, nous étudions dans le **chapitre 3** l'interaction entre l'EGCG et ces surfaces, par DM et champ de force. Une méthode est ainsi proposée pour prédire l'influence de l'EGCG sur l'oxalate de calcium ouvrant la voie à l'étude du potentiel d'autres molécules organiques à inhiber la croissance des cristaux d'oxalate de calcium, présentant alors un intérêt thérapeutique dans la lithiase urinaire.

Au-delà des mécanismes biologiques extrêmement sophistiqués et complexes mis en place dans la biominéralisation, les mécanismes chimiques impliqués dans la nucléation des minéraux sont encore débattus dans la littérature. La **théorie classique de la nucléation** (CNT) offre un premier cadre pour tenter de décrire la transition de phase entre une solution ioniques et un cristal de taille nanométrique. Selon la CNT, une particule cristalline croit par des additions successives d'ions, lui permettant d'atteindre une taille critique à partir de laquelle le processus de croissance devient énergétiquement favorable. Cela est due à la compétition entre deux termes énergétiques associés à la croissance cristalline : **(i)** l'énergie libre de formation du « bulk » cristallin est un terme favorable, proportionnel au volume de la particule, et donc à  $r^3$  **(ii)** l'énergie libre de formation d'une interface solide-liquide est un terme défavorable, proportionnel à la surface de la particule, et donc à  $r^2$ . Ainsi, avant d'atteindre la taille critique  $r_c$ , c'est le terme surfacique qui prédomine. Dans le cas des phosphates de calcium, de récentes études affirment que le processus de nucléation ne peut pas se décrire de manière classique. Des observations par microscopie électronique (Cryo TEM)<sup>16,17</sup> ont montrées que les phosphates de calcium ont tendance à former des **complexes de prénucléation** et des **clusters de prénucléation**. L'agrégation de ces clusters permet ensuite la formation de **nanoparticules amorphes** (< 50nm) qui, à partir d'une certaine taille critique, sont alors soumises au processus de cristallisation. Ce passage par une phase amorphe est également observé *in vivo* dans l'émail des dents<sup>18</sup>.

Afin d'élucider le phénomène de **nucléation** des phosphates de calcium, nous présentons dans le **chapitre 4** une méthode originale et novatrice permettant d'étudier les processus impliqués. Nous proposons ainsi une coordonnée de réaction permettant de décrire la nucléation en DM et discutons des différents intermédiaires rencontrés le long de cette coordonnée.

La DM est une technique qui en théorie permet d'observer des transformations telles que l'adsorption de l'EGCG sur une surface d'oxalate de calcium ou la nucléation des phosphates de calcium. En revanche, les échelles de temps accessibles ne permettent pas d'observer ces transitions de manière réversible. Une première stratégie consiste à optimiser le calcul des

interactions. Ainsi, le choix d'un potentiel champ de force qui ne traite pas les électrons de manière explicite, permet d'accéder à des échelles de temps plus importantes (environ 100 ns) que les techniques dites *ab initio* (environ 100 ps avec par ex. la théorie de la fonctionnelle de la densité, DFT). Ces ordres de grandeurs sont basés sur des calculs effectués durant cette thèse, sur un modèle de 412 atomes représentant l'interface hydroxyapatite/eau (à titre d'exemple représentatif). La **métadynamique**<sup>19</sup> constitue une stratégie supplémentaire. Dans cette méthode, un biais de forme gaussienne est déposé à intervalle de temps régulier, le long d'une coordonnée appelée **variable collective** (CV) et correspondant à la transformation voulant être étudiée. La métadynamique permet ainsi d'améliorer l'échantillonnage des configurations le long de la CV en « décourageant » le système à retourner vers les configurations déjà visitées. D'autre part, la métadynamique permet de reconstruire l'énergie libre associée à la CV. Cette stratégie a été adoptée dans le **chapitre 3** afin de quantifier l'interaction entre l'EGCG et les surfaces d'oxalates de calcium, ainsi que dans le **chapitre 4** pour explorer le mécanisme de nucléation des phosphates de calcium. Enfin, la DM par **échange de réplique (REMD)**<sup>20</sup> est également un outil permettant d'améliorer l'échantillonnage des configurations. Durant la REMD, N DM sont lancées à N températures différentes. Chacune des simulations est appelée réplique. Tous les x étapes de la DM, deux répliques juxtaposées ont une certaine probabilité d'échanger leurs positions selon un critère de Metropolis-Hastings, dépendant de leur différence d'énergie potentielle. Cela permet d'explorer plus de configurations en dépassant certaines barrières énergétiques grâce à l'énergie thermique. La technique a été appliquée lors de l'étude de l'adsorption de l'EGCG sur l'oxalate de calcium.

Le manuscrit de cette thèse s'articule ainsi :

- Le **chapitre 1** expose les méthodes et les problématiques liées aux chapitres 2, 3 et 4.
- Le **chapitre 2** est consacré aux prédictions des morphologies cristallines de l'oxalate de calcium hydraté et à la construction de modèles représentatifs de ses surfaces.
- Le **chapitre 3** s'articule autour de la mise au point d'une méthodologie permettant de prédire, par la DM, l'affinité de l'EGCG pour les différentes surfaces de l'oxalate de calcium hydraté. Cette méthode tire avantage de la métadynamique et de la REMD pour améliorer l'échantillonnage des configurations.
- Le **chapitre 4** explore le mécanisme de nucléation du phosphate de calcium par DM en s'appuyant sur la métadynamique et l'utilisation de coordonnées variationnelles.



Enfin, nous concluons ce manuscrit en soulignant les points les plus originaux de notre travail de Doctorat, et en proposant des perspectives d'intérêt dans le domaine de la biominéralisation.

1. Yool, A. & Tyrrell, T. Role of diatoms in regulating the ocean's silicon cycle. *Global Biogeochem. Cycles* **17**, 1–21 (2003).
2. De Stefano, L., Rea, I., Rendina, I., De Stefano, M. & Moretti, L. Lensless light focusing with the centric marine diatom *Coscinodiscus walesii*. *Opt. Express* **15**, 18082 (2007).
3. Glimcher, M. J. Bone: Nature of the Calcium Phosphate Crystals and Cellular, Structural, and Physical Chemical Mechanisms in Their Formation. *Rev. Mineral. Geochemistry* **64**, 223–282 (2006).
4. Franceschi, V. R. & Horner, H. T. Calcium oxalate crystals in plants. *Bot. Rev.* **46**, 361–427 (1980).
5. Freeman Rutgers, J. W., Landis, W. J., Silver, F. H. & Freeman, J. W. Collagen as a scaffold for biomimetic mineralization of vertebrate tissues. *J. Mater. Chem.* **16**, 1495–1503 | (2006).
6. Wang, Y. *et al.* The predominant role of collagen in the nucleation, growth, structure and orientation of bone apatite. *Nat. Mater.* **11**, 724–733 (2012).
7. George, A. & Veis, A. Phosphorylated proteins and control over apatite nucleation, crystal growth, and inhibition. *Chem. Rev.* **108**, 4670–4693 (2008).
8. Glimcher, M. J. Mechanism of calcification: Role of collagen fibrils and collagen-phosphoprotein complexes in vitro and in vivo. *Anat. Rec.* **224**, 139–153 (1989).
9. Mahamid, J. *et al.* Bone mineralization proceeds through intracellular calcium phosphate loaded vesicles: A cryo-electron microscopy study. *J. Struct. Biol.* **174**, 527–535 (2011).
10. Moe, O. W. Kidney stones: pathophysiology and medical management. *Lancet* **367**, 333–344 (2006).
11. Liu, Y. *et al.* Differentiating calcium oxalate and hydroxyapatite stones in vivo using dual-energy CT and urine supersaturation and pH values. *Acad. Radiol.* **20**, 1521–5 (2013).
12. Daudon, M., Traxer, O., Lechevallier, E. & Saussine, C. La lithogénèse. *Progrès en Urol.* **18**, 815–827 (2008).
13. Rode, J. *et al.* Daily Green Tea Infusions in Hypercalciuric Renal Stone Patients: No Evidence for Increased Stone Risk Factors or Oxalate-Dependent Stones. *Nutrients* **11**, 256 (2019).

14. Chen, Z., Wang, C., Zhou, H., Sang, L. & Li, X. Modulation of calcium oxalate crystallization by commonly consumed green tea. *Cryst. Eng. Comm* **12**, 845–852 (2010).
15. Graham, H. N. Green tea composition, consumption, and polyphenol chemistry. *Prev. Med. (Baltim)*. **21**, 334–350 (1992).
16. Dey, A. *et al.* The role of prenucleation clusters in surface-induced calcium phosphate crystallization. *Nat. Mater.* **9**, 1010–1014 (2010).
17. Habraken, W. J. E. M. *et al.* Ion-association complexes unite classical and non-classical theories for the biomimetic nucleation of calcium phosphate. *Nat. Commun.* **4**, 1507–1512 (2013).
18. Beniash, E., Metzler, R. A., Lam, R. S. K. & Gilbert, P. U. P. A. Transient amorphous calcium phosphate in forming enamel. *J. Struct. Biol.* **166**, 133–143 (2009).
19. Barducci, A., Bonomi, M. & Parrinello, M. Metadynamics. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **1**, 826–843 (2011).
20. Rhee, Y. M. & Pande, V. S. Multiplexed-Replica Exchange Molecular Dynamics Method for Protein Folding Simulation. *Biophys. J.* **84**, 775–786 (2003).

# - Chapitre 1 -

## Bibliographie et Méthodologie

<b><i>Bibliographie et Méthodologie</i></b> .....	<b>10</b>
<b>1.1. La construction de modèles moléculaires pour les matériaux</b> .....	<b>12</b>
1.1.1. Le point de départ .....	12
1.1.2. Les conditions périodiques aux limites.....	12
1.1.3. Description physique des modèles moléculaires .....	13
<b>1.2. Description quantique des interactions interatomiques</b> .....	<b>13</b>
1.2.1. Les postulats de la mécanique quantique <sup>2,3</sup> .....	13
1.2.2. Résolution de l'équation de Schrödinger et le principe variationnel.....	14
1.2.3. Résolution de l'équation de Schrödinger d'un système moléculaire complexe .....	15
1.2.4. La DFT périodique.....	18
<b>1.3. Description des interactions interatomiques par un champ de force</b> .....	<b>24</b>
<b>1.4. L'optimisation de géométrie : à la recherche du minimum local</b> .....	<b>26</b>
<b>1.5. La dynamique moléculaire</b> .....	<b>27</b>
1.5.1. Principe général .....	27
1.5.2. Les ensembles thermodynamiques .....	28
1.5.3. La dynamique moléculaire avec un champ de force.....	29
1.5.4. La dynamique moléculaire <i>ab initio</i> .....	30
1.5.5. Principe d'ergodicité.....	31
1.5.6. Le problème des minima locaux en dynamique moléculaire .....	31
1.5.7. Replica Exchange Molecular Dynamics.....	32
1.5.8. La métadynamique.....	33
<b>1.6. Prédiction des morphologies cristallines des oxalates de calcium</b> .....	<b>36</b>
1.6.1. Problématique .....	36
1.6.2. État de l'art.....	39
1.6.3. Objectifs.....	44
1.6.4. Méthode .....	44
<b>1.7. Adsorption d'une catéchine sur les surfaces de l'oxalate de calcium</b> .....	<b>47</b>
1.7.1. Problématique .....	47
1.7.2. État de l'art.....	49
1.7.3. Objectifs.....	49
1.7.4. Méthode .....	49

<b>1.8. Étude de la nucléation des phosphates de calcium .....</b>	<b>53</b>
1.8.1. Problématique .....	53
1.8.2. État de l'art.....	57
1.8.3. Objectifs.....	58
1.8.4. Méthode .....	58

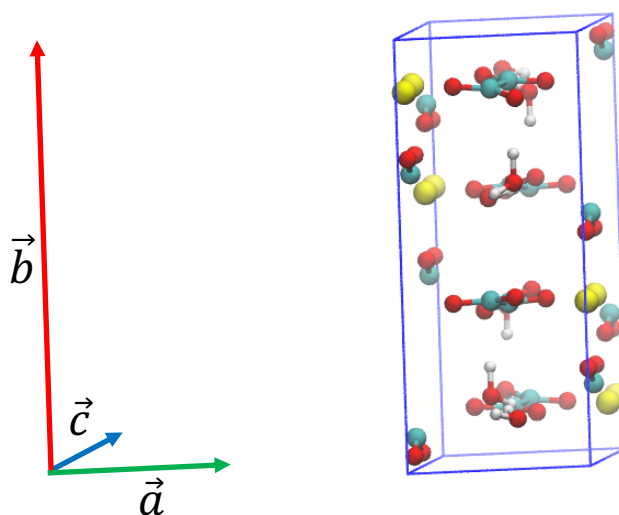
## 1.1. La construction de modèles moléculaires pour les matériaux

### 1.1.1. Le point de départ

Afin de pouvoir modéliser un matériau il faut être en capacité de construire un modèle calculable et proche de ce que l'on souhaite observer. Dans le domaine de la chimie des matériaux nous souhaitons la plupart du temps décrire les positions atomiques d'un système cristallin. Pour définir un tel modèle, les données issues de la cristallographie par diffraction des rayons X ou par diffusion des neutrons sont essentielles et constituent souvent le point de départ des simulations. En effet, la cristallographie nous donne les positions atomiques dans la maille élémentaire). Ce motif se répète par translation dans les directions données par les trois vecteurs de base définissant la maille. La maille cristalline de l'oxalate de calcium monohydrate est présenté en Figure 1.

### 1.1.2. Les conditions périodiques aux limites

Afin de simuler un système périodique infini, les simulations mises en œuvre doivent respecter les conditions périodiques aux limites. Ces conditions permettent de prendre en compte les interactions existantes entre les particules présentes à l'intérieure de la boîte de simulation et les particules présentes dans les images périodiques de la boîte.



**Figure 1 : exemple de la maille élémentaire de l'oxalate de calcium monohydrate (COM) issue d'une étude de cristallographie par diffraction des neutrons<sup>1</sup>. Légende : oxygène : rouge, carbone : bleu, calcium : jaune**

Un atome proche de l'une des faces de la maille cristalline sera également proche de la face opposée. Sur la Figure 1, on remarque que les atomes des ions oxalates proches des faces de

droite et de gauche sont représentés de part et d'autre de la maille cristalline. Il faut alors imaginer les images de ces atomes pour reconstituer la molécule d'oxalate. Cela souligne l'importance des conditions périodiques aux limites pour décrire les interactions entre les images des particules d'un réseau cristallin. En revanche, les conditions périodiques aux limites ne sont pas toujours souhaitables. C'est le cas lorsque l'on veut simuler un système non périodique. Il faut alors choisir un système suffisamment large afin de réduire au mieux les artefacts liés aux interactions entre images.

### 1.1.3. Description physique des modèles moléculaires

Une fois l'état initial du modèle moléculaire décrit, un cadre théorique est nécessaire afin de décrire les interactions existantes entre les atomes composant le système. Ce cadre est essentiel pour pouvoir calculer l'énergie du système et prédire son évolution dans le temps. Au cours de cette thèse, deux approches ont été utilisées : une approche quantique basée sur la théorie de la fonctionnelle de la densité (DFT) et une approche classique basée sur les champs de forces chimiques.

## 1.2. Description quantique des interactions interatomiques

### 1.2.1. Les postulats de la mécanique quantique<sup>2,3</sup>

La mécanique quantique est fondée sur plusieurs postulats.

- Tout d'abord, l'état quantique d'une particule est décrit complètement par une fonction  $\psi(\mathbf{r}, t)$  et  $\psi^*(\mathbf{r}, t)\psi(\mathbf{r}, t)dxdydz$  est la probabilité de la particule de se trouver dans le volume  $dxdydz$  à la position  $\mathbf{r}$  et au temps  $t$ . Du fait de ce postulat sur la probabilité, la fonction d'onde décrivant une particule doit être normée :

$$\iiint_{-\infty}^{+\infty} \psi^*(\mathbf{r}, t)\psi(\mathbf{r}, t)dxdydz = 1 \quad (1)$$

- A toute observable physique en mécanique classique correspond un opérateur hermitien linéaire en mécanique quantique. Par exemple, l'énergie cinétique  $K = \frac{p^2}{2m}$  en mécanique classique conduit à l'opérateur

$$\hat{K} = -\frac{\hbar^2}{2m} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) = -\frac{\hbar^2}{2m} \nabla^2 \quad (2)$$

en mécanique quantique. Pour l'énergie potentielle, l'opérateur reste inchangé par rapport à l'expression classique puisque celui-ci ne dépend que des positions du système. On peut alors écrire l'opérateur de l'énergie totale du système appelé hamiltonien :

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + U(x, y, z) \quad (3)$$

- Pour toute mesure faite sur une observable physique associée à un opérateur  $\hat{A}$ , les seules valeurs pouvant être observées sont les valeurs propres satisfaisant l'équation suivante :

$$\hat{A}\Psi = a\psi \quad (4)$$

Pour l'observable énergie cela conduit à l'équation de Schrödinger indépendante du temps :

$$\hat{H}\Psi = E_n\psi \quad (5)$$

Si l'on veut mesurer l'énergie du système, les seules valeurs que l'on peut obtenir sont les valeurs propres  $E_n$  de l'opérateur  $\hat{H}$ . D'un point de vue analytique, les fonctions propres décrivant les différents états quantiques du système peuvent être obtenue en résolvant l'équation de Schrödinger. On peut alors obtenir les différents niveaux d'énergies autorisés, associés aux valeurs propres de l'hamiltonien.

### 1.2.2. Résolution de l'équation de Schrödinger et le principe variationnel

L'équation de Schrödinger pour les hydrogénoïdes (espèces chimiques avec 1 seul électron) est soluble analytiquement. En revanche, à partir de l'hélium, le terme d'interaction inter-électronique rend l'équation de Schrödinger impossible à résoudre analytiquement. Une approximation permettant de résoudre ce problème est donnée par le principe variationnel<sup>2</sup>. La valeur propre correspondant à l'énergie fondamentale (l'énergie la plus basse) du système est donnée par l'équation suivante, obtenue à partir de l'équation de Schrödinger (équation 5):

$$E_0 = \frac{\int \psi_0^* \hat{H} \psi_0 dV}{\int \psi_0^* \psi_0 dV} \quad (6)$$

Le principe variationnel précise que si  $\psi_0$  est remplacée par une fonction d'onde quelconque  $\Phi$  alors l'énergie obtenue à partir de l'expression suivante :

$$E_\Phi = \frac{\int \Phi^* \hat{H} \Phi dV}{\int \Phi^* \Phi dV} \quad (7)$$

est supérieure à l'énergie de l'état fondamental :



$$E_{\Phi} > E_0 \quad (8)$$

On peut alors utiliser n'importe quelle fonction  $\Phi$  (appelée fonction d'essai) comportant des paramètres variationnels et minimiser  $E_{\Phi}$  par rapport à ces paramètres. Un choix judicieux des paramètres de  $\Phi$  permettra de faire converger  $\Phi$  vers la fonction d'onde fondamentale et donc  $E_{\Phi}$  vers  $E_0$ .

### 1.2.3. Résolution de l'équation de Schrödinger d'un système moléculaire complexe

Une approche commune est de définir une fonction d'essai comme une combinaison linéaire de plusieurs fonctions,  $f_i$  appelée fonction de base :

$$\Phi_{essai} = \sum_{i=1}^N c_i f_i \quad (9)$$

Le choix des fonctions de bases peut être arbitraire, elles peuvent contenir ou non des paramètres variationnels. En revanche, un choix judicieux est nécessaire pour obtenir la convergence de l'énergie  $E_{\Phi}$  vers  $E_0$ . Pour cela, la fonction d'essai doit « ressembler » au mieux à la fonction d'onde fondamentale du système. Ce choix sera discuté plus tard. Les coefficients  $c_i$  sont la plupart du temps des nombres réels et sont pris également comme paramètres variationnels. A chaque étape de la minimisation, l'énergie doit être calculée selon l'équation 7 ce qui revient à calculer les termes  $\int \sum_{i=1}^N c_i^* f_i^* \hat{H} \sum_{i=1}^N c_i f_i dV$  et  $\int \sum_{i=1}^N c_i^* f_i^* \sum_{i=1}^N c_i f_i dV$ . Cela revient à résoudre le déterminant séculaire suivant :

$$\begin{pmatrix} H_{11} - E_{\Phi} S_{11} & \cdots & H_{1N} - E_{\Phi} S_{1N} \\ \vdots & \ddots & \vdots \\ H_{N1} - E_{\Phi} S_{N1} & \cdots & H_{NN} - E_{\Phi} S_{NN} \end{pmatrix} = 0 \quad (10)$$

où  $H_{ij} = \int f_i^* \hat{H} f_j dV$  et  $S_{ij} = \int f_i^* f_j dV$ . Augmenter le nombre de fonctions de bases permet d'augmenter la précision du calcul de l'énergie mais demande une puissance de calcul accrue. On peut également remarquer ici, que choisir des fonctions élémentaires orthonormées simplifie le problème puisque  $S_{ij} = 0$  et  $S_{ii} = 1$ .

$$\begin{pmatrix} H_{11} - E_{\Phi} & \cdots & H_{1N} \\ \vdots & \ddots & \vdots \\ H_{N1} & \cdots & H_{NN} - E_{\Phi} \end{pmatrix} = 0 \quad (11)$$

Dans la grande majorité des calculs moléculaires quantiques, l'hamiltonien décrivant le système est simplifié par l'approximation de Born-Oppenheimer qui consiste à découpler le mouvement des électrons de celui des noyaux. Dans cette approximation, et à positions atomiques fixées,

l'énergie cinétique des noyaux est nulle et l'interaction noyau-noyau est une constante dépendant uniquement de la géométrie du système. Les seules sources de variation de l'énergie sont alors : l'énergie cinétique des électrons, l'interaction des électrons avec les noyaux et les interactions des électrons entre eux. Ces trois termes constituent l'hamiltonien électronique :

$$\hat{H}_{el} = \hat{T}_e + \hat{v}_{ee} + \hat{v}_{eI} \quad (12)$$

où  $\hat{T}_e$  représente l'énergie cinétique des électrons,  $\hat{v}_{ee}$  l'interaction inter-électron et  $\hat{v}_{eI}$  l'interaction des électrons avec le potentiels des noyaux. Les expressions analytiques de ces trois termes sont connues exactement :

$$\hat{T}_e = - \sum_i^N \frac{\hbar^2}{2m} \nabla_i^2 = \sum_i^N \frac{1}{2} \nabla_i^2 \quad (a.u) \quad (13)$$

$$\hat{v}_{ee} = \sum_{i,j} \frac{e^2}{4\pi\epsilon_0 |\vec{r}_i - \vec{r}_j|} = \sum_{i,j} \frac{1}{r_{ij}} \quad (a.u) \quad (14)$$

$$\hat{v}_{eI} = - \sum_{i,I} \frac{Z_I e^2}{4\pi\epsilon_0 |\vec{r}_i - \vec{R}_I|} = \sum_{i,I} \frac{Z_I}{r_{iI}} \quad (a.u) \quad (15)$$

où  $\vec{R}_I$  correspond à la position du noyau d'indice  $I$ ,  $Z_I$  son nombre atomique,  $\vec{r}_i$  la position de l'électron d'indice  $i$  et  $e$  la charge d'un électron.

Un problème survient du fait du terme d'interaction inter-électronique  $\hat{v}_{ee}$  lorsque l'on veut résoudre numériquement l'état fondamental électronique d'un système décrit par une fonction  $\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ . Pour calculer l'état électronique fondamental d'un système, il faut, d'après le principe variationnel, calculer l'élément de matrice  $H_{ij} = \int f_i^* \hat{H} f_j dV$ . Un ion phosphate  $\text{PO}_4^{3-}$  par exemple, comporte  $N=32$  électrons de valence. Supposons qu'il occupe un volume de  $5 \text{ \AA}$  divisée en  $5 \times 5 \times 5$  cubes, chaque électron à une probabilité non nulle d'occuper un de ces cubes. Il existe donc  $125^{32}$  configurations électronique possibles. La partie spatiale de la fonction d'onde multi-électronique du système dépend de  $3N = 96$  coordonnées cartésiennes. Il faut donc réaliser  $125^{96}$  opérations pour déterminer un élément de matrice  $H_{ij}$ . Actuellement les ordinateurs les plus puissants peuvent atteindre une vitesse de  $10^{18}$  opérations par secondes. Il faudrait donc  $10^{183}$  secondes pour calculer tous les éléments de matrice  $H_{ij}$ . A titre de comparaison, l'âge estimé de la terre en secondes est de  $10^{17}$ . Ainsi, il est clairement impossible



combinaison de fonctions monoélectroniques. En effet, la méthode HF néglige la corrélation existante entre les mouvements des électrons.

#### 1.2.4. La DFT périodique

##### 1.2.4.1. Principe général<sup>4</sup>

La DFT (ou théorie de la fonctionnelle de la densité) apporte une autre solution plus efficace en considérant les  $N$  électrons du système, non plus au travers d'une fonction d'onde multi-électronique mais *via* une fonction de densité dépendant des  $N$  électrons du système. La DFT implique l'existence d'un lien direct entre la fonctionnelle de la densité et l'énergie du système (premier théorème de Hohenberg-Kohn)<sup>4</sup>. On considère également que l'énergie du système est minimale lorsque  $n(\vec{r})$  est la densité de l'état fondamental (deuxième théorème de Hohenberg-Kohn). Néanmoins, la fonctionnelle exacte est inconnue et doit être approximée au mieux. Une première difficulté de la DFT réside dans la recherche d'une « bonne » fonctionnelle reliant densité et énergie. Les équations de la DFT se résolvent de manière itérative grâce aux équations mono-électroniques de Kohn-Sham :

$$h_j \varphi_j(\vec{r}) = \left[ \hat{T}_e + \hat{v}_{el} + e^2 \int d^3r' \frac{n(\vec{r}')}{|\vec{r} - \vec{r}'|} + V_{xc}(\vec{r}; [n]) \right] \varphi_j(\vec{r}) = \varepsilon_j \varphi_j(\vec{r}) \quad (18)$$

$$n(\vec{r}) = \sum_j c_j |\varphi_j(\vec{r})|^2 \quad (19)$$

Les deux termes  $\hat{T}_e$  et  $\hat{v}_{el}$  correspondent respectivement à l'énergie cinétique et à l'interaction électron-noyau. Ils ne diffèrent pas des termes de l'hamiltonien mono-électronique dans la méthode HF (voir équations **13** et **15**). En revanche les deux termes suivants représentant respectivement l'interaction inter-électronique et l'énergie d'échange dépendent désormais de  $n(\vec{r})$ . La densité  $n(\vec{r})$  dépend quant à elle des fonctions électroniques. Tout comme dans la méthode HF, le système d'équations n'est pas linéaire et est résolu de manière itérative grâce au cycle auto cohérent. Des fonctions d'essais sont prises comme point de départ et leurs paramètres variationnels sont recalculés à chaque cycle.

Depuis la découverte de la DFT, les efforts se sont concentrés : **(i)** d'une part, sur l'amélioration de la fonctionnelle de la densité, principalement en prenant en compte le terme d'échange  $V_{xc}$  de manière plus précise. Sans le terme d'échange, tout comme pour la méthode HF, la DFT néglige la corrélation existante entre le mouvement des électrons. Cette énergie doit contenir les effets quantiques d'échange et de corrélation électronique, et doit compenser le terme

d'auto-interaction présent dans le terme d'interaction inter-électronique dans les équations de Kohn-Sham (voir équation 18). (ii) D'autre part, sur le choix des fonctions d'essai qui est fondamental pour assurer la convergence de l'énergie à chaque itération et pour atteindre une bonne description chimique du système.

#### 1.2.4.2. Les fonctions de base et les pseudopotentiels

Chaque fonction monoélectronique  $\varphi_j(\vec{r})$  est exprimée comme une combinaison linéaire des fonctions de bases :

$$\varphi_j(\vec{r}) = \sum_{\alpha} c_{j,\alpha} f_{\alpha}(\vec{r}) \quad (20)$$

A chaque itération du cycle auto-cohérent, les paramètres variationnels  $c_{j,\alpha}$  sont optimisés pour approcher l'énergie fondamentale du système. Les fonctions  $f_{\alpha}(\vec{r})$  peuvent également contenir des paramètres variationnels. Plus le nombre de fonctions de base est important plus le calcul est précis, mais alors plus le temps de calcul est élevé. Un compromis doit être adopté.

Le choix de la nature de ces fonctions est essentiel pour assurer d'une part, la convergence de l'énergie lors de l'étape variationnelle, et d'autre part, une description correcte et adéquate du système. Celles-ci doivent reproduire le mieux possible le comportement de l'électron. Proche du noyau, l'électron ressent un potentiel très fort due au potentiel de coulomb et doit être localisé tandis que loin de celui-ci doit être délocalisé lorsqu'il s'éloigne du noyau.

#### **Les orbitales de Slater**

Les orbitales de Slater<sup>5</sup> sont des orbitales localisée sur l'atome. Elles ont dépendance radiale de la forme suivante :

$$R(r) = Nr^{n-1}e^{-\zeta r} \quad (21)$$

Où  $n$  est le nombre quantique principal ( $n = 1, 2, \dots$ ),  $N$  est une constante de normalisation,  $r$  est la distance de l'électron au noyau et  $\zeta$  est une constante déterminée par des règles empiriques et liée à la charge effective du noyau.

Ces fonctions décrivent bien le comportement de l'électron proche du noyau mais sont moins précises quand le nombre d'électrons augmente. De plus, leur intégration est compliquée mathématiquement et coûteuse numériquement. C'est pourquoi les fonctions d'onde de type gaussienne sont souvent préférées aux orbitales de Slater.

## Les orbitales gaussiennes

Les fonctions de base de type gaussienne (Gaussian Type Orbital ou GTO) sont également localisées sur l'atome et peuvent être exprimées sous la forme :

$$g_{\zeta,n,l,m}(r, \theta, \phi) = NY_{l,m}(\theta, \phi)r^{2n-n-l}\exp(-\zeta r^2) \quad (22)$$

$N$  est une constante de normalisation,  $Y_{l,m}$  représente les harmoniques sphériques et le terme  $\exp(-\zeta r^2)$  représente la dépendance en  $r$  (distance électron-noyau). La dépendance radiale de la fonction rend ce type de fonction plus facilement intégrable par comparaison aux orbitales de Slater.

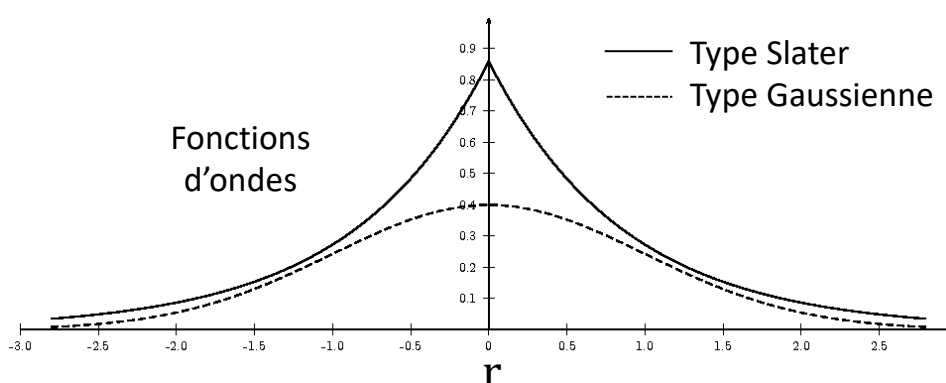


Figure 2 : Distribution électronique des orbitales de type Gaussienne et de type Slater. Adapté de Hoggan *et coll.*<sup>6</sup>

Les avantages d'une base constituée de fonctions gaussiennes sont :

- La petitesse de la base en général et sa croissance linéaire en fonction du nombre d'atomes composant le système,
- La facilité du calcul numérique des intégrales

Ses inconvénients sont :

- sa dépendance des positions atomiques,
- une mauvaise description de la densité électronique loin du noyau,

## Les ondes planes

Les ondes planes sont très utilisées pour décrire les systèmes périodiques. Contrairement aux fonctions de Slater et aux orbitales gaussiennes, les ondes planes ne sont pas centrées sur l'atome. En notation exponentielle, elles s'écrivent  $e^{i\vec{G}\cdot\vec{r}}$  avec  $\vec{G}$  un vecteur de l'espace réciproque.

Les fonctions décrivant le système sont alors une somme sur l'ensemble  $\{\vec{G}\}$  des points du réseau réciproque :

$$u_k(\vec{r}) = \sum_{\vec{G}} c_k(\vec{G}) e^{i(k+\vec{G})r} \quad (23)$$

Les ondes planes prises en compte dans la base ont une énergie cinétique inférieure à l'énergie de coupure  $E_{cut-off}$ . La taille de la base est alors définie par :

$$\frac{\hbar^2(k + \vec{G})^2}{2m} \leq E_{cut-off} \quad (24)$$

Les bases de fonctions d'onde planes présentent les avantages suivants :

- Le nombre de fonction de base est constant, quel que soit le nombre d'atomes
- La base ne dépend pas des positions atomiques
- Un seul critère est nécessaire pour la troncature de la base

Ses inconvénients sont :

- La base doit être très grande comparé aux bases de gaussiennes
- Les orbitales localisés sont mal décrites et l'usage de pseudopotentiel est nécessaire

### **L'approche mixte (GPW)**

Une base comportant à la fois des ondes gaussienne et des ondes planes est utilisée dans l'approche GPW<sup>7</sup> (Gaussian Plane Wave). Cela permet de profiter des avantages provenant de ces deux types de fonction. Deux représentations de la densité sont alors utilisées. **(i)** la base d'orbitales gaussiennes, centrée sur les atomes, permet de décrire correctement la fonction d'onde proche des noyaux ; **(ii)** la base d'ondes planes permet de décrire plus précisément la densité électronique éloignée des noyaux.

### **Les pseudopotentiels**

Le concept de pseudopotentiel consiste à substituer le potentiel d'interaction entre le noyau et les électrons de cœur par un potentiel effectif interagissant avec les électrons de valence. Cette approximation est acceptable car les électrons de cœurs, peu mobiles, participent peu aux phénomènes chimiques. Le pseudopotentiel permet alors de limiter le nombre de fonction d'ondes impliqué dans l'étape variationnelle du cycle auto cohérent, et donc, de gagner en efficacité de calcul.

Les pseudopotentiels de Goedecker, Teter et et Hutter (GTH)<sup>8,9</sup> sont souvent utilisés dans la méthode GPW. Ces pseudopotentiels conservent la norme et sont optimisés pour le calcul numérique.

#### 1.2.4.3. Les fonctionnelles

##### **L'approximation de la densité locale**

Une première approximation pour considérer le terme d'échange  $V_{xc}(\vec{r}; [n])$  (voir équation 18) est l'approximation LDA (Local Density Approximation) basée sur le modèle du gaz uniforme : la densité est traitée de manière locale. Les systèmes moléculaires étudiés, en particulier dans le domaine de la chimie des matériaux, sont très différents du modèle du gaz uniforme. En réalité la densité électronique présente une variation dans l'espace.

##### **Les méthodes de gradient corrigé**

La deuxième génération de fonctionnelles utilise l'approximation GGA (Generalized Gradient Approximation) qui prend en compte les effets de la variation de la densité électronique au niveau du terme d'échange. Les méthodes méta-GGA proposent une dépendance en densité d'énergie cinétique.

##### **Les méthodes hybrides**

Les fonctionnelles hybrides introduisent une fraction du terme d'échange basée sur la méthode de Hartree-Fock et une autre fraction sur les orbitales de Kohn-Sham. Pour la fonctionnelle de Becke<sup>10</sup> par exemple, le terme d'échange peut être écrit comme suit :

$$V_{xc}(\vec{r}; [n]) = (1 - a)E_x^{LDA} + a E_x^{HF} + bE_x^{GGA} + cE_c^{LYP} + (1 - c)E_c^{LDA} \quad (25)$$

Les paramètres  $a$ ,  $b$  et  $c$  sont optimisés empiriquement et dépendent de la fonctionnelle GGA choisie. L'énergie d'échange est la combinaison de trois termes :

- $E_x^{LDA}$  correspond à l'énergie d'échange obtenue par l'approche LDA,
- $E_x^{HF}$  correspond à l'énergie d'échange obtenue par un calcul de Hartree-Fock
- et  $E_x^{GGA}$  correspond à l'énergie d'échange obtenue par l'approche GGA.

Les méthodes qui choisissent d'introduire le terme  $E_x^{HF}$  calculé par la méthode d'Hartree-Fock deviennent plus coûteuse en temps de calcul mais sont plus précises dans le calcul de l'énergie.



## Choisir une fonctionnelle

Le choix d'une fonctionnelle dépend de la précision de calcul associée au système d'étude ainsi que du temps de calcul disponible. Les fonctionnelles implémentées dans les logiciels VASP et CP2K sont des fonctionnelles de type GGA (PBE, BLYP), méta GGA (BB95) et hybride méta GGA (B3LYP). J.P. Perdrew et K. Schmidt<sup>11</sup> ont mis au point une échelle de Jacob pour classer les fonctionnelles selon leur degré de sophistication. Cette échelle est présentée en Tableau 1.

SOLUTION EXACTE			
	Fonctionnelle	Dépendance	Exemple
4 <sup>ème</sup> échelon	Hybride	$\rho(\vec{r})\nabla\rho(\vec{r}), \Delta\rho(\vec{r}), \Phi_i(\vec{r})$	B3LYP
3 <sup>ème</sup> échelon	Méta-GGA	$\rho(\vec{r}), \nabla\rho(\vec{r}), \Delta\rho(\vec{r})$	BB95
2 <sup>ème</sup> échelon	GGA	$\rho(\vec{r}), \nabla\rho(\vec{r})$	PBE
1 <sup>er</sup> échelon	LDA	$\rho(\vec{r})$	SPWL

**THEORIE DE HARTREE-FOCK**

Tableau 1 : Échelle de Jacob des fonctionnelles adaptée de Perdew *et coll*<sup>11</sup>

Le niveau de sophistication le plus bas est donné par la méthode Hartree-Fock alors que le plus élevé correspond à la solution exacte. Plus les termes d'échange et de corrélation sont complets, plus la fonctionnelle est sophistiquée. Choisir une fonctionnelle avec un échelon élevé permet d'augmenter la probabilité d'améliorer l'accord théorie-expérience mais ce n'est pas toujours le cas. Par exemple, dans le domaine des matériaux, la fonctionnelle reproduisant au mieux les paramètres de mailles expérimentaux d'un cristal est souvent celle qui est choisie. Dans la thèse d'Ivan Petit<sup>12</sup>, il a par exemple été montré que la fonctionnelle GGA PBE reproduit mieux les paramètres de mailles expérimentaux des oxalates de calcium que la fonctionnelle hybride BLYP qui est pourtant plus haute dans l'échelle de Jacob. Le choix d'une fonctionnelle n'est donc a priori pas évident et doit faire l'objet de tests comparant l'expérience avec la théorie.

### 1.2.4.4. Le terme de dispersion de Grimme

Les forces de London sont des forces faibles, intermoléculaires, issues de l'interaction entre deux dipôles induits. Ces forces résultent du caractère probabiliste de la densité électronique d'une molécule. Il y a une certaine probabilité que les électrons d'une molécule soient répartis à un instant t de telle manière qu'ils forment un moment dipolaire au sein de la molécule. Les

interactions entre deux dipôles ne sont pas prises en compte par les fonctionnelles usuelles, nécessitant l'ajout d'un terme de correction. Dans la méthode D2 de Grimme<sup>13</sup>, un terme empirique est ajouté :

$$E_{DFT-D} = E_{K-S} + E_{disp} \quad (26)$$

Ce terme prend la forme suivante :

$$E_{disp} = -\frac{1}{2} \sum_i^N \sum_j^N \frac{C_{6ij}}{r_{ij}^6} f_{d,6}(r_{ij}) \quad (27)$$

Les images périodiques sont bien sûr prises en compte et un rayon de coupure est choisi au-delà duquel les interactions ne sont plus prises en compte.  $C_{6ij}$  et  $r_{ij}$  sont des paramètres dépendant de la nature des deux atomes d'index  $i$  et  $j$  :

$$C_{6ij} = \sqrt{C_{6i}C_{6j}} \quad r_{ij} = r_{0i} + r_{0j} \quad (28)$$

Ces valeurs sont obtenues par de calculs quantiques de type Hartree-Fock : les paramètres  $C_{6i}$  sont obtenu par des calculs UDFT-PBE0/QZVP et les paramètres  $r_i$  par des calculs ROHF/TZV<sup>13</sup>. Ils sont tabulés pour chaque élément. La fonction  $f_{d,6}(r_{ij})$  est une fonction d'amortissement de type Fermi qui permet de moduler la force du terme de dispersion par rapport à la distance séparant les deux atomes impliqués<sup>14</sup>. Cette fonction est optimisée pour les différentes fonctionnelles de DFT. Dans la méthode D2, les paramètres dépendent seulement de la nature chimique des éléments, et non de leur environnement. De plus, un seul terme ne permet pas de rendre compte précisément des forces de dispersion. La méthode D3<sup>15</sup> tente de pallier à ce problème en ajoutant un terme à l'expression de l'énergie de dispersion et en rendant les paramètres  $C_{6ij}$  et  $C_{8ij}$  dépendant de l'environnement local des atomes :

$$E_{disp} = -\frac{1}{2} \sum_i^N \sum_j^N \frac{C_{6ij}}{r_{ij}^6} f_{d,6}(r_{ij}) - \frac{1}{2} \sum_i^N \sum_j^N \frac{C_{8ij}}{r_{ij}^8} f_{d,8}(r_{ij}) \quad (29)$$

### 1.3. Description des interactions interatomiques par un champ de force

Une approche beaucoup moins coûteuse en temps de calcul que la DFT est de ne plus traiter le système à l'échelle électronique, mais atomique. Dans l'approche dites des champs de force, l'effet des électrons est moyenné pour chaque atome et les interactions interatomiques sont définies par un potentiel, prenant en compte différents termes dépendant de la nature physique de l'interaction. Chacun des termes dépend de paramètres empiriques, optimisés afin de

reproduire les données expérimentales et/ou théoriques (DFT). C'est pour cette raison que cette méthode est qualifiée d'empirique. L'approche des champs de force permet de traiter des systèmes moléculaires plus importants en taille et sur une échelle de temps plus importante. D'autre part, le caractère empirique du champ de force permet d'obtenir de meilleurs résultats quantitatifs. En revanche, contrairement aux méthodes *ab-initio* (DFT, HF, ...) les paramètres du champ de force doivent être adaptés au système que l'on souhaite étudier.

Voici les différents termes présents dans la plupart des champs de forces :

- Un potentiel harmonique prend en compte les liaisons covalentes entre les différents types d'atomes. Ce terme dépend de deux paramètres par paire de type d'atomes  $i$  et  $j$  : la distance d'équilibre  $r_{ij}$  entre les deux atomes, et la constante de raideur  $k_{ij}$  du ressort fictif entre les deux atomes.  $r_{ij}$  est obtenu à partir de mesures de distances interatomiques expérimentales ou à partir de calculs *ab initio*.  $k_{ij}$  peut être obtenu grâce aux données de spectroscopie vibrationnelle (spectres IR) et peut également être calculé de manière théorique.
- Un potentiel harmonique permet également de prendre en compte les variations angulaires entre deux liaisons à l'intérieure d'une molécule. De la même manière que pour le terme impliquant les liaisons covalentes, les différents paramètres sont déterminés à partir de données expérimentales.
- Un potentiel de torsion est utilisé pour prendre en compte les angles dièdres. Ces angles sont définis par quatre atomes consécutifs, liés dans la même molécule.
- Un potentiel électrostatique permet de prendre en compte les interactions entre deux charges. Pour ce terme, à chaque type d'atome est associée une charge partielle.
- Un potentiel de Lennard-Jones est utilisé pour les interactions de van der Waals.

Il en résulte donc une énergie potentielle totale pouvant être décrite par la somme suivante :

$$E_{totale} = E_{liaison} + E_{angle} + E_{dihèdre} + E_{electrostatique} + E_{vdw} \quad (30)$$

Par exemple, pour le champ de force OPLS<sup>16</sup> (Optimized Potentials for Liquid Simulations) qui a été utilisé durant cette thèse pour décrire les phosphates de calcium, les termes énergétiques de liaisons et d'angles s'écrivent :

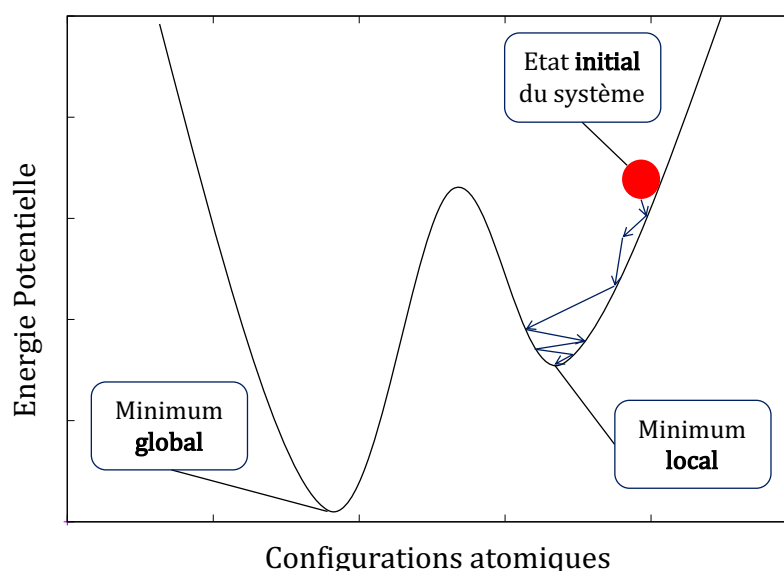
$$E_{liaison} = k_{ij}(r_0 - r_{ij})^2 \quad (31)$$

$$E_{angles} = k_{\theta}(\theta_0 - \theta_0)^2 \quad (32)$$

Pour ces deux termes, les paramètres  $k_{ij}$ ,  $r_0$ ,  $k_{\theta}$  et  $\theta_0$  dépendent de la nature de la molécule étudiée et des données expérimentales correspondantes. Ainsi,  $r_{ij}$  correspond à la longueur moyenne de la liaison formée par les atomes  $i$  et  $j$  et  $\theta_0$  correspond à l'angle d'équilibre formé par les trois atomes d'intérêt. Tous les autres termes sont également composés de paramètres empiriques optimisés pour reproduire les données expérimentales et *ab initio*.

#### 1.4. L'optimisation de géométrie : à la recherche du minimum local

Pour trouver la géométrie optimum d'un ensemble d'atomes, il est possible de minimiser son énergie potentielle en prenant comme paramètres les 3 coordonnées cartésiennes de chaque atome. Si le système est composé d'une centaine d'atomes il faut donc trouver le minimum de l'énergie dans un espace de 300 variables.



**Figure 3 : Schéma illustrant le problème des minimums locaux lors d'une optimisation géométrique**

Un algorithme couramment utilisé pour l'optimisation géométrique est celui de Broyden-Fletcher-Goldfarb-Shanno (BFGS)<sup>17</sup>. Il s'agit d'une méthode permettant de résoudre un problème d'optimisation non linéaire, sans contraintes. L'algorithme BFGS utilise la matrice hessienne (matrice des dérivés secondes de l'énergie par rapports aux positions) pour orienter l'optimisation géométrique vers le minimum dans l'espace des coordonnées du systèmes.

L'utilisation d'un tel algorithme en chimie théorique présente quelques désavantages. D'une part, en minimisant l'énergie potentielle seulement, aucune considération entropique n'est prise

en compte. L'algorithme de BFGS permet au mieux d'atteindre la configuration la plus stable à température 0K. D'autre part, la configuration finale obtenue dépend fortement de l'état initial du système. Ainsi, comme schématisé sur la Figure 3, l'algorithme ne permet pas de dépasser certaines barrières énergétiques mais seulement d'atteindre un minimum localement.

## 1.5. La dynamique moléculaire

### 1.5.1. Principe général

La dynamique moléculaire permet de prédire la position des atomes au cours du temps, et ainsi de calculer des propriétés dépendantes du temps.

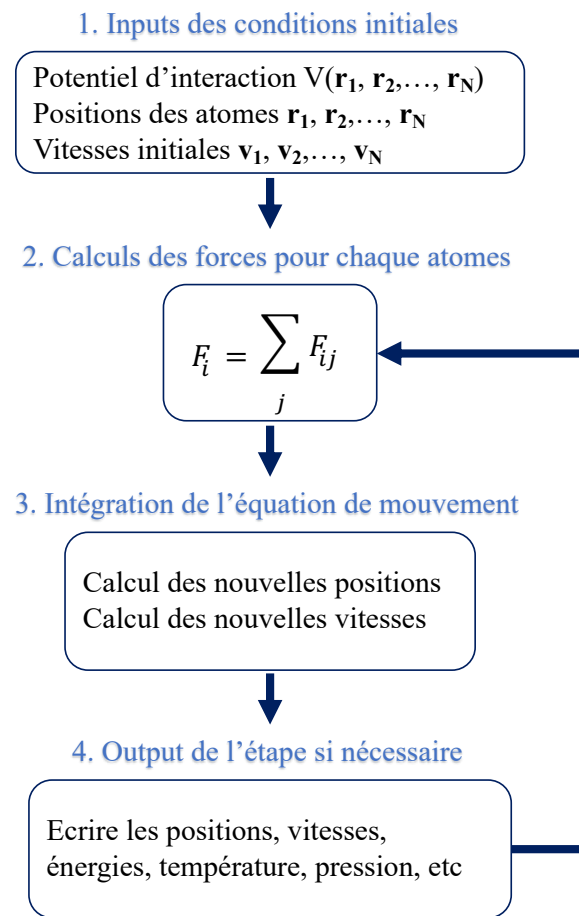


Figure 4 : Algorithme général de la dynamique moléculaire

Comme décrit en Figure 4, la dynamique moléculaire est un algorithme qui se découpe en étapes bien précises. Dans un premier temps, il faut définir les positions initiales de l'ensemble des atomes du système ainsi que les vitesses initiales. Les vitesses initiales sont en général déterminées de manière aléatoire selon une distribution de Maxwell-Boltzmann. Ensuite, la force agissant sur chacun des atomes est calculée à partir des positions atomiques et du potentiel interatomique choisi. Puis, l'équation de mouvement est intégrée pour obtenir les nouvelles

positions et vitesses du système. Lors de cette étape, plusieurs algorithmes peuvent être utilisés. Les plus couramment utilisés sont le Velocity Verlet (VV)<sup>18</sup> et le leap-frog<sup>19</sup> Verlet (LFV). Dans l'algorithme VV, en appliquant à  $\mathbf{r}(t + \Delta t)$  un développement de Taylor à l'ordre deux et à  $\mathbf{v}(t + \Delta t)$  un développement de Taylor à l'ordre un, les expressions suivantes sont obtenues pour les positions  $\mathbf{r}(t)$  et les vitesses  $\mathbf{v}(t)$  :

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \Delta t \mathbf{v}(t) + \frac{\Delta t^2}{2m} \mathbf{F}(t) \quad (33)$$

$$\mathbf{v}(t + \Delta t) = \mathbf{v}(t) + \frac{\Delta t}{2m} [\mathbf{F}(t) + \mathbf{F}(t + \Delta t)] \quad (34)$$

L'algorithme leap-frog<sup>19</sup> utilise les positions  $\mathbf{r}(t)$  à l'instant  $t$  et les vitesses  $\mathbf{v}(t)$  à l'instant  $t - \frac{1}{2}\Delta t$ . Il met à jour les positions et les vitesses en utilisant les forces  $\mathbf{F}(t)$  déterminées par les positions à l'instant  $t$  en utilisant les relations suivantes :

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \Delta t \mathbf{v}\left(t + \frac{\Delta t}{2}\right) \quad (35)$$

$$\mathbf{v}\left(t + \frac{\Delta t}{2}\right) = \mathbf{v}\left(t - \frac{\Delta t}{2}\right) + \frac{\Delta t}{m} \mathbf{F}(t) \quad (36)$$

Contrairement à la méthode VV, la méthode LFV est une méthode réversible dans le temps. Cette propriété reflète la réalité physique d'une simulation. La méthode d'intégration LFV a été choisie pour les dynamiques moléculaires réalisées dans cette thèse.

Au cœur de la dynamique moléculaire, se trouve la définition du potentiel interatomique et donc des forces régissant les interactions entre les atomes composant le système. Dans le cadre de la théorie des champs de force, ces forces sont calculées grâce au potentiel décrit dans l'équation 30. Dans le cadre de la DFT, les forces agissant sur les atomes sont calculées à partir du calcul de la structure électronique comme décrit dans l'équation 40.

### 1.5.2. Les ensembles thermodynamiques

Dans l'ensemble canonique (ou ensemble NVT), le nombre de particules  $N$ , le volume  $V$  et la température  $T$  du système sont fixés. L'ensemble canonique est approprié pour décrire un système fermé en contact thermique avec un bain de chaleur. Pour réaliser une simulation dans cet ensemble, la température du système doit être maintenue constante. Une approche utilisée

est de redéfinir les vitesses du système périodiquement de manière aléatoire selon une distribution de Maxwell-Boltzmann, centrée autour de la vitesse moyenne correspondant à la température considérée. Une autre méthode utilisée au cours de cette thèse est l'utilisation d'un thermostat de Nosé-Hoover qui consiste à introduire un coefficient de friction  $\xi$  dans les équations de la dynamique<sup>20,21</sup> :

$$m\ddot{r}_i = f_i - \xi p_i \quad (37)$$

$\xi$  varie avec le temps selon l'équation :

$$\dot{\xi} = v_T^2 \left[ \frac{\tau}{T} - 1 \right] \quad (38)$$

$\tau$  représente la température instantanée de la simulation, tandis que  $T$  représente la température fixe dans l'ensemble canonique. Quand  $\tau > T$ , le système a une température trop élevée.  $\dot{\xi} > 0$  et donc  $\xi$  augmente pour diminuer les forces et refroidir le système. L'inverse arrive lorsque le système est trop « froid ».

Dans l'ensemble isotherme-isobare (ou ensemble NPT), le nombre de particule  $N$ , la température  $T$  et la pression  $P$  sont fixés. Un barostat est alors utilisée en, plus du thermostat. Le barostat de Nosé-Hoover<sup>22</sup> peut être appliqué de manière similaire au thermostat. Le barostat de Parrinello-Rahman<sup>23</sup> apporte une extension supplémentaire en laissant chacun des trois vecteurs  $\mathbf{a}, \mathbf{b}$  et  $\mathbf{c}$  de la boîte de simulation, indépendants.

### 1.5.3. La dynamique moléculaire avec un champ de force

Avec un champ de force, les atomes sont considérés ponctuellement et seules les interactions de paires sont considérées comme décrit dans l'équation 30. Un terme d'interaction à trois corps peut être utilisé dans certains champs de force. Ce n'est pas le cas dans cette thèse.

Le logiciel GROMACS (GRoningen MACHine for Chemical Simulations)<sup>24</sup> a été utilisé pour toutes les dynamiques moléculaires avec un champ de force. Ce logiciel est largement utilisé par la communauté des chimistes théoriciens et a été guidé par mon entourage. De plus, le logiciel GROMACS s'est avéré adapté pour le calcul des modèles moléculaires étudié au cours de cette thèse.

#### 1.5.3.1. Les Condition Périodiques aux limites

Pour prendre en compte les conditions périodiques aux limites, GROMACS utilise la convention minimale de l'image : une seule image (la plus proche) de chaque atome est prise en compte pour les termes d'interaction à courte portée.

### 1.5.3.2. Les interactions à longue distance

Une distance de troncature (ou cut-off) est définie pour les interactions non covalentes. Au-delà de cette distance, le potentiel non covalent entre les deux atomes est nul. Cette distance ne doit pas dépasser la moitié de la taille du plus petit des vecteurs de la boîte de simulation, afin de respecter la convention minimale de l'image.

Le nombre de paires d'interactions à longue portée est beaucoup plus important que le nombre d'interactions à courte portée. GROMACS intègre la méthode de somme en réseau appelée somme d'Ewald<sup>25,26</sup> pour optimiser le calculs de ces interactions.

### 1.5.4. La dynamique moléculaire *ab initio*

Les méthodes de dynamique moléculaire *ab initio* sont largement utilisées du fait de leur pouvoir prédictif fort. On peut citer les méthodes de Car-Parrinello<sup>27</sup> et de Born-Oppenheimer comme les plus couramment utilisées. Ces méthodes dérivent toutes les deux de l'équation de Schrödinger dépendante du temps :

$$\hat{H}\Psi(\mathbf{r}, t) = i\hbar \frac{\partial \Psi(\mathbf{r}, t)}{\partial t} \quad (39)$$

En décorrélant le mouvement des électrons de celui des noyaux et en considérant les noyaux comme des particules ponctuelles il est possible de ramener l'équation de Schrödinger dépendante du temps à une équation où le mouvement des noyaux est décrit de manière classique dans un potentiel effectif due aux électrons :

$$M_I \ddot{R}_I(t) = -\nabla_I \int d\mathbf{r} \Psi^* H_e \Psi \quad (40)$$

Dans l'approche BOMD, Born-Oppenheimer, la structure électronique est calculée à chaque étape de la dynamique moléculaire. La structure électronique dépend de l'ensemble des positions nucléaires déterminées à un instant donné. Ainsi, la partie de la structure électronique est réduite à la résolution d'un problème quantique indépendant du temps. Autrement dit, l'équation de Schrödinger indépendante du temps est résolue parallèlement à la propagation des noyaux via la dynamique moléculaire classique. La dépendance temporelle de la structure électronique est alors une conséquence du mouvement nucléaire. La méthode Car-Parrinello a été développée initialement dans le cadre de la DFT en utilisant une base d'ondes planes uniquement. Elle permet un gain d'efficacité du calcul numérique car elle ne nécessite pas un calcul de la densité électronique à chaque intervalle de temps. En revanche, elle est inefficace pour le traitement de bases gaussiennes.



Comparé à l'approche de type champs de force, la dynamique moléculaire *ab-initio*, en considérant les électrons comme un degré de liberté supplémentaire, devient beaucoup plus coûteuse numériquement. Ainsi, les échelles de temps accessibles par ces méthodes sont bien plus faibles. D'après les calculs réalisés durant cette thèse, la dynamique moléculaire champs de force est plus performante que la dynamique BOMD basée sur la DFT par un facteur d'ordre  $10^3$ .

#### 1.5.5. Principe d'ergodicité

L'hypothèse d'ergodicité est un principe fondamental de la physique statistique. Il affirme qu'à l'équilibre, la valeur moyenne d'une grandeur calculée de manière statistique est égale à la moyenne d'un grand nombre de mesures effectuées dans le temps. Autrement dit, pour mesurer correctement la valeur moyenne d'une observable, il faut la mesurer pendant un temps suffisamment long.

#### 1.5.6. Le problème des minima locaux en dynamique moléculaire

Ce principe d'ergodicité n'est pas respecté lorsque la dynamique moléculaire rencontre le problème des minima locaux. Ce phénomène est similaire à celui décrit plus haut en Figure 3, pour l'algorithme de BFGS. En revanche, le problème a changé de forme. Contrairement à l'optimisation de géométrie, le problème n'est pas lié intrinsèquement à l'algorithme utilisé, mais uniquement à la puissance de calcul des ordinateurs utilisés pour résoudre numériquement l'équation de Newton. Les simulations moléculaires standard ne permettent pas d'accéder à des échelles de temps suffisamment grandes ou à des tailles de systèmes suffisamment grands pour observer certains événements chimiques. Par exemple, si on considère un événement chimique simple,  $A \rightarrow B$ , ayant une barrière d'activation de  $50 \text{ kJ.mol}^{-1}$  (voir Figure 5), la proportion de molécules A ayant une énergie supérieure ou égale à cette énergie d'activation peut être calculée grâce à la relation de Boltzmann :

$$e^{\frac{50}{RT}} = \frac{1}{4 \cdot 10^9} \quad (\text{à } 303K) \quad (41)$$

Il faudrait donc un système composé de  $10^9$  molécules de A pour pouvoir observer une transition de A vers B, et ce, une seule fois. Cela implique de calculer toutes les paires d'interactions entre les  $10^9$  molécules. En ordre de grandeur, cela donne donc  $10^{18}$  interactions à calculer. La puissance de calcul mis à disposition par les meilleurs centres de calculs ne permet pas de calculer un tel nombre d'interactions. L'ordre de grandeur atteignable pour un système est plutôt de  $10^4$  atomes. Ceci ramène à  $\frac{10^4}{10^9} = 10^{-5}$  la probabilité d'observer la transition de A

vers B à un instant  $t$ . Il est donc peu probable d'observer une telle transition au cours d'une simulation moléculaire où le nombre d'atome est de l'ordre de  $10^4$ . On qualifie la transition de A vers B comme un évènement rare.

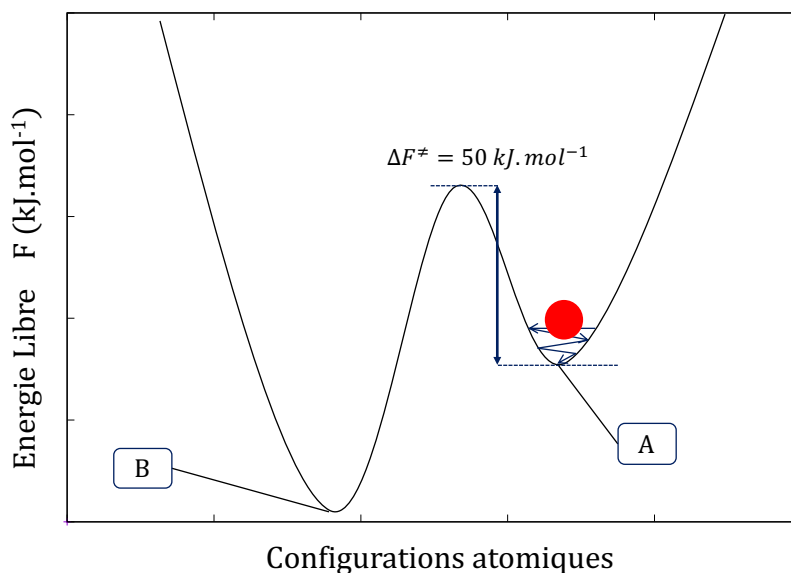


Figure 5 : Schéma illustrant le problème des minima locaux en dynamique moléculaire

Comme illustré en Figure 5, la dynamique moléculaire permet de faire converger le système vers un équilibre thermodynamique autour duquel il fluctue. Ces fluctuations ne permettent pas les transitions liées aux évènements *rare* même si elles sont favorables thermodynamiquement.

### 1.5.7. Replica Exchange Molecular Dynamics

#### 1.5.7.1. Le principe général

La méthode Replica Exchange Molecular Dynamics (REMD) apporte une première solution pour améliorer l'échantillonnage des configurations d'un système<sup>28-30</sup>. Cette méthode a donné satisfaction dans la simulation du processus de repliement des protéines<sup>29</sup>. En effet, le repliement d'une protéine est un évènement *rare* qui n'est pas observé lors d'une simulation moléculaire conventionnelle. Durant la REMD,  $N$  simulations sont lancées à  $N$  températures différentes. Chacune des simulations est appelée *réplique*. Tous les  $x$  étapes de la dynamique moléculaire, deux répliques juxtaposées ont une certaine probabilité d'échanger leurs positions dépendant de leur différence d'énergie potentielle. Cela permet d'explorer plus de configurations en dépassant certaines barrières énergétiques grâce à l'énergie thermique.

### 1.5.7.2. Paramétrisation

Une distribution exponentielle des températures des répliques est nécessaire pour obtenir un taux d'échange constant entre les différentes répliques :

$$T_i = T_0 e^{k.i} \quad (42)$$

Ici,  $i$  correspond à l'index de la réplique et  $k$  permet de paramétrer l'écart de température entre les différentes répliques.

Pour avoir un taux d'échange suffisant, un recouvrement des énergies potentielles entre les différentes répliques est nécessaire. Sachant, que l'énergie potentielle augmente avec la température, il faudra adapter l'écart de température entre les répliques ( $k$ ) avec cette contrainte. Plus l'écart de température est réduit, plus il faudra choisir un nombre important de répliques pour atteindre une température maximale suffisante et permettant d'envisager les barrières souhaitant être franchies.

### 1.5.8. La métadynamique

#### 1.5.8.1. Principe général

La métadynamique<sup>31</sup> est une technique de choix pour étudier les événements *rare*s. Elle permet également de dépasser certaines barrières d'énergie qui bloquent le système dans un minimum local. La métadynamique est plus riche, en permettant de reconstruire le diagramme d'énergie libre associé à une ou plusieurs coordonnées de réactions. Dans cette approche, la dynamique moléculaire est biaisée selon une ou plusieurs coordonnée(s) bien précise(s) qui doit(vent) correspondre au phénomène étudié. Ces coordonnées sont appelées Variables Collectives (Collective Variables, CV).

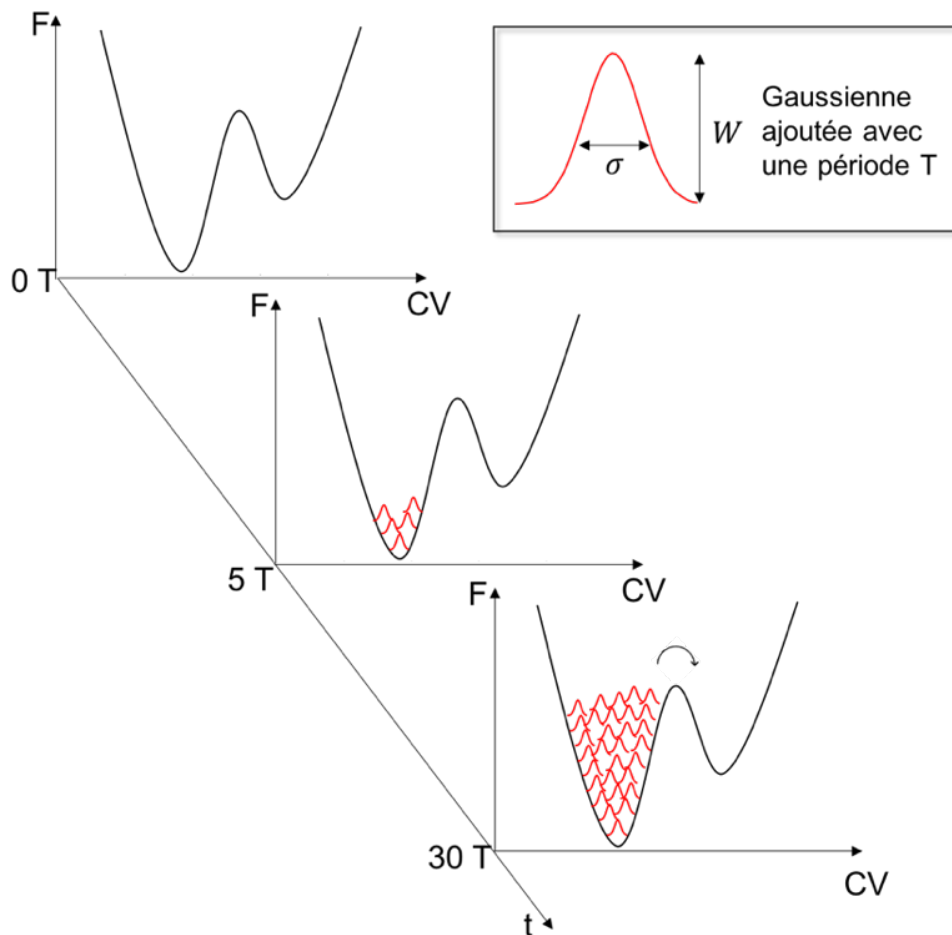
Un biais (ou une force) sous forme de gaussienne est ajouté à l'énergie potentielle du système toutes les périodes  $T$  de la dynamique, le long de la CV. Le biais « décourage » le système à revisiter les mêmes configurations, correspondant à une valeur particulière de la CV. Au fur et à mesure des ajouts de gaussiennes, le biais total ajouté empêche le système de rester dans le puits d'énergie libre minimum et l'encourage à aller visiter un autre puits comme schématisé en Figure 6. Le biais total imposé au système,  $V_G$  est la somme des gaussiennes ajoutées au cours du temps :

$$V_G = \int_0^t dt' \omega \exp\left(-\sum_{i=1}^d \frac{(CV_i(R) - CV_i(R(t')))^2}{2\sigma_i^2}\right) \quad (43)$$

$\omega$  est un taux énergétique, exprimé en fonction de la hauteur des gaussiennes  $W$  et la période d'ajout des gaussiennes  $T$  :

$$\omega = \frac{W}{T} \quad (44)$$

La somme de  $i$  à  $d$  correspond au nombre de CV utilisées dans la métadynamique.  $\sigma_i$  correspond à la largeur à mi-hauteur de la gaussienne.



**Figure 6 : Représentation schématique du principe de la métadynamique. Des gaussiennes de hauteur  $W$  et de largeur  $\sigma$  s'ajoutent le long d'une CV toute les périodes  $T$  de la dynamique moléculaire. Le système est contraint à sortir du puits de potentiel dans lequel il se trouvait.**

L'énergie potentielle totale du système est alors la somme de l'énergie potentielle usuelle et du biais énergétique  $V_G$  due à la métadynamique. En sommant l'ensemble des collines le long de la CV, il est ainsi possible de reconstruire la surface d'énergie libre le long de cette CV.

#### 1.5.8.2. Le choix des variables collectives (CV)

La définition des CV correspondant à la transition voulant être étudiée est souvent loin d'être triviale. Une mauvaise définition de la CV conduit à une mauvaise interprétation du diagramme d'énergie libre<sup>32</sup>. Une CV adéquate doit contenir tous les degrés de liberté associés à la transition considérée.

#### 1.5.8.3. Paramétrisation

Pour que la somme des gaussiennes puissent reproduire précisément le diagramme d'énergie libre, la hauteur  $W$  et la largeur  $\sigma_i$  des gaussiennes doivent être choisies judicieusement. La hauteur  $W$  des gaussiennes détermine l'erreur associée à la hauteur des barrières du diagramme d'énergie libre. Une bonne estimation de l'énergie libre d'activation associée à la transition voulant être étudiée permet de choisir une hauteur raisonnable. Par exemple, pour une transition associée à une barrière d'environ  $100 \text{ kJ.mol}^{-1}$ , une hauteur de gaussienne de  $1 \text{ kJ/mol}$  représente alors 1% de la barrière totale, ce qui est raisonnable. D'autre part, la largeur  $\sigma_i$  des gaussiennes doit être plus faible que la largeur du puits associé à la CV d'intérêt. Pour pouvoir reconstruire le puits précisément, il est aisé de comprendre intuitivement pourquoi les gaussiennes doivent être moins larges que le puits.

Après ce tour d'horizon des techniques de simulations utilisées au cours de la thèse, nous allons les appliquer aux trois problématiques suivantes :

- La Prédiction des morphologies cristallines des oxalates de calcium
- L'Adsorption d'une catéchine sur les surfaces de l'oxalate de calcium
- L'Étude de la nucléation des phosphates de calcium

## 1.6. Prédiction des morphologies cristallines des oxalates de calcium

### 1.6.1. Problématique

#### 1.6.1.1. Généralités

Les oxalates de calcium, de formule  $\text{CaC}_2\text{O}_4 \cdot n\text{H}_2\text{O}$  avec  $n = 0, 1, 2, 3$ , sont des minéraux présents dans le vivant. Les phases les plus observées sont les phases monohydratée (COM) et dihydratée (COD).

Ainsi, on retrouve les oxalates de calcium dans la plupart des familles de plantes sous la dénomination de raphides<sup>33</sup>. Les fonctions associées au minéral sont variées. Il peut être un moyen de protection contre les herbivores ou peut jouer un rôle mécanique au même titre que le squelette chez les vertébrés. Le processus de biominéralisation de ces cristaux est un processus physicochimique qui n'est pas aléatoire mais bien dirigé par l'organisme vivant. En effet, il a été montré que la quantité, la morphologie et la taille des cristaux d'oxalate de calcium dépend de la taxinomie du végétal et est donc soumis à un contrôle génétique. Les mécanismes par lesquels ces organismes vivants régulent la croissance cristalline sont complexes et peuvent survenir à plusieurs étapes.

Les oxalates de calcium ont également été très étudiés dans la littérature car ils jouent un rôle majeur dans une pathologie présente chez les humains : la lithiase urinaire. La pathologie est due à la formation de calculs rénaux dans le système urinaire. Les constituants principaux de ces calculs rénaux sont établis et correspondent à un nombre limité de phases cristallines bien définies : des oxalates de calcium (oxalate de calcium mono hydraté ou dihydraté), des phosphates d'ammonium, de calcium et de magnésium et de l'acide urique. Depuis plusieurs décennies, les calculs à base d'oxalate de calcium sont les plus fréquemment observés (> 70% de tous les échantillons analysés en Europe et en Amérique du Nord)<sup>34,35</sup>.

Il a été montré *in vitro* que divers composés, peuvent jouer le rôle d'inhibiteur de la croissance cristalline, comme des protéines (ostéopontine<sup>36</sup>, protéine de Tamm-Horsfall et albumine<sup>37</sup>), des molécules organiques (citrate<sup>38</sup> et l'acide phytique<sup>39</sup>) ou des sels inorganiques ( $\text{Mg}^{2+}$  en particulier<sup>40</sup>). Il est alors clair que l'environnement physicochimique joue un rôle crucial sur le contrôle de la croissance cristalline et sur la morphologie finale de l'oxalate de calcium.

#### 1.6.1.2. Structure cristalline

Il existe plusieurs hydrates de l'oxalate de calcium :

- La whewellite ou Calcium Oxalate Monohydrate (COM) ( $\text{CaC}_2\text{O}_4 \cdot \text{H}_2\text{O}$ ),

- La weddellite ou Calcium Oxalate Dihydrate (COD) ( $\text{CaC}_2\text{O}_4 \cdot 2\text{H}_2\text{O}$ )
- La caoxite ou Calcium Oxalate Trihydrate (COT) ( $\text{CaC}_2\text{O}_4 \cdot 3\text{H}_2\text{O}$ ).

Le COM est la phase thermodynamiquement la plus stable<sup>41</sup>. La phase COD peut être formée dans certaines conditions de synthèse bien précises et a tendance à se transformer en COM à température ambiante<sup>42</sup>. Au cours de cette thèse, nous nous sommes concentrés sur les phases COM et COD car ce sont les deux phases principales observées dans les plantes et dans les calculs rénaux<sup>34,35,43</sup>. La structure du COM a été résolue par Daudon *et coll.* par diffraction des neutrons<sup>1</sup> ainsi que dans d'autres études par diffraction des rayons X<sup>44-46</sup>. La structure du COD a été résolue par diffraction des rayons X<sup>45,47,48</sup>.

### 1.6.1.3. Morphologies cristallines dans les synthèses

Le COM a été observé par Microscopie Électronique à Balayage (MEB) dans plusieurs synthèses *in vitro* en milieu aqueux<sup>49-53</sup>. Les différentes morphologies observées dans ces études sont présentées en Figure 7.

<b>Morphologies et indices de Miller des surfaces (convention de Tazzoli <i>et coll.</i><sup>45</sup>)</b>	a) 	b) 	c) 
<b>Images MEB</b>			
<b>Conditions de synthèse</b>	Synthèse en milieu aqueux, équimolaire en ions oxalates et calcium, $T_{\text{amb}}$	Synthèse en milieu aqueux, $T=37^\circ\text{C}$	Synthèse dans un tube en U, rempli de gel de silice, équimolaire, $T_{\text{amb}}$
<b>Références</b>	Chen <i>et coll.</i> <sup>54</sup>	Millan <sup>50</sup>	Franchini-Angela <i>et coll.</i> <sup>51,53</sup>

Figure 7 : morphologie cristalline (observée en SEM) dans des synthèses en milieu aqueux.

La convention adoptée pour les indices de Miller des surfaces du COM est celle de Tazzoli *et coll.*<sup>45</sup>, elle est décrite dans l'article de Millan<sup>50</sup>. Cette convention a été utilisée tout au long de

cette thèse. Millan, ainsi que d'autres auteurs, décrivent également la formation de macles cristallines où plusieurs cristaux de COM sont imbriqués, conduisant à des morphologies particulières, comme celle de la pince à linge, appelée également «*twin crystal*»<sup>51</sup>. Par souci de simplicité, seuls les monocristaux sont présentés ici (les macles ne sont pas prises en compte). Quatre surfaces sont prédominantes dans les synthèses (Figure 7) : (100), (010), (021) et (12 $\bar{1}$ ). Au moins l'une des deux surfaces (021) ou (12 $\bar{1}$ ) est toujours présente (Figure 7a,b). Les deux surfaces peuvent également coexister (Figure 7c).

Le COD a souvent la forme d'une bipyramide dominée par la surface d'indice (101). Dans certaines conditions, la surface d'orientation (010), s'intercale entre les deux pyramides formées par la surface (101). Par exemple, les synthèses de Leroy<sup>55</sup> et de Giordani *et coll.*<sup>56</sup> montrent l'absence de la surface (010) alors que la surface est présente dans la synthèse de Chen *et coll.*<sup>52</sup> en présence de thé vert.

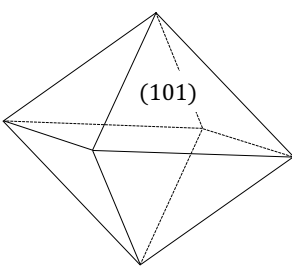
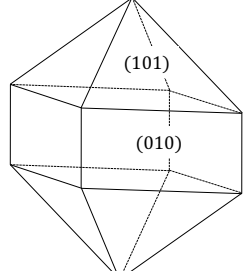
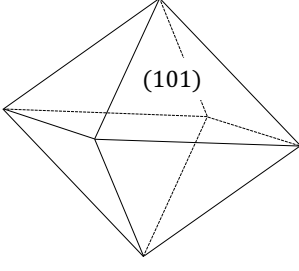
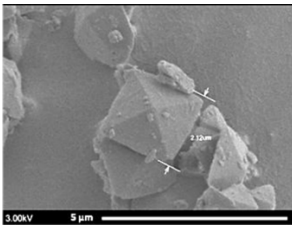
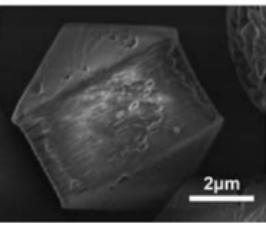
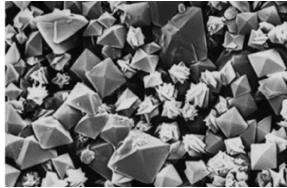
<b>Morphologies et indices de Miller des surfaces</b>	a) 	b) 	c) 
<b>Images MEB</b>			
<b>Conditions de synthèse</b>	Synthèse en milieu aqueux, en excès d'ions oxalates, T <sub>amb</sub>	Synthèse en milieu aqueux, en condition équimolaire, en présence de thé vert, T <sub>amb</sub>	Synthèse dans une solution synthétique d'urine en excès d'ions calcium
<b>Références</b>	Leroy <sup>55</sup>	Chen <i>et coll.</i> <sup>54</sup>	Giordani <i>et coll.</i> <sup>56</sup>

Figure 8 : morphologie cristalline du COD (observée en MEB) dans des synthèses en milieu aqueux



Il a été montré également que le polymère PEG-b-PMAA permet de stabiliser considérablement la surface (010)<sup>57</sup>. Il semble donc que la surface (010) nécessite une stabilisation par une molécule supplémentaire pour être exposée.

#### 1.6.1.4. Morphologies dans les minéraux

Comparé au COM synthétique, le COM minéral présente une plus grande diversité de morphologies<sup>58</sup>. Les surfaces (100), (010), (021) et (12 $\bar{1}$ ), ont encore une forte incidence, mais contrairement aux calculs synthétiques, la surface (001) fait aussi partie des surfaces les plus exposées. Beaucoup d'autres surfaces sont observées plus rarement dans quelques cristaux : (140), (13 $\bar{1}$ ), (12 $\bar{3}$ ), (14 $\bar{1}$ ), (161), (140), (011), (16 $\bar{1}$ ), (15 $\bar{1}$ ), (013), (121), (112) et (031). Le COD se retrouve dans certains sédiments marins et présente la même morphologie que le COD synthétique<sup>59</sup>.

#### 1.6.1.5. Morphologies dans les calculs rénaux

Un calcul rénal n'est jamais composé d'un unique cristal. C'est toujours un tissu complexe, composé d'une multitude de cristaux agrégés dans une matrice organique de lipides et de protéines. Pour le COM et le COD, des morphologies similaires à celles présentées en Figure 7a et Figure 7b sont décrites dans les calculs rénaux par Daudon *et coll.*<sup>60</sup>.

Il est important d'étudier les interactions que les cristaux d'oxalate de calcium peuvent avoir avec leur environnement chimique, afin de prédire son influence sur la croissance et la morphologie cristalline. La simulation moléculaire apporte un cadre théorique essentiel à la prédiction du comportement de tels cristaux. Dans cette partie, nous nous sommes intéressés à la prédiction de la forme des cristaux de l'oxalate de calcium par une approche *ab initio* pour proposer un modèle complet décrivant, au niveau moléculaire, les différentes surfaces du COM et du COD.

### 1.6.2. État de l'art

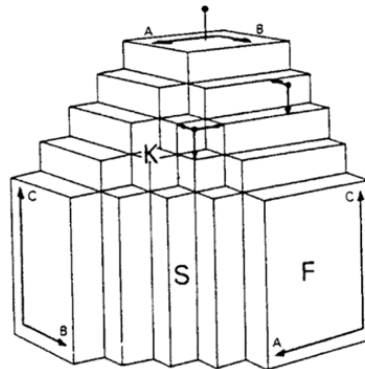
#### 1.6.2.1. Les modèles de prédiction de morphologie cristalline

Plusieurs méthodes existent afin de prédire la morphologie d'un cristal.

**(1) L'analyse de Bravais, Friedel, Donnay et Harker (BFDH)**<sup>61</sup> prend en considération les données géométriques de la maille cristalline pour déterminer les morphologies possibles. Dans cette méthode, seul le groupe d'espace et les dimensions de la maille cristalline sont utilisés comme critères pour prédire la morphologie<sup>61</sup>. La théorie soutient que la vitesse de croissance d'une surface est proportionnelle à la distance interréticulaire dans la direction considérée. Plus

cette distance est grande, plus la surface croît lentement et plus elle est exposée. Il faut comprendre ici que la croissance d'une surface permet d'agrandir les surfaces qui lui sont orthogonales. Ainsi, plus la croissance d'une surface est importante, plus les surfaces qui lui sont orthogonales sont représentées dans la morphologie cristalline. Au contraire, plus sa croissance est lente moins les surfaces orthogonales seront représentées.

**(2) La théorie de Hartman-Perdok (H-P)** ou Periodic Bond Chain (PBC) nous propose une autre approche pour prédire théoriquement la morphologie d'un cristal à partir de la connaissance de sa structure atomique<sup>62</sup>. Cette théorie part de l'observation que les cristaux ont des bords plats qui restent lisses lors de la croissance cristalline. La seule explication est qu'il existe un enchaînement ininterrompu de liaisons fortes (comme par exemple les interactions ioniques entre les ions oxalates et calcium) qui se forment lors du processus de croissance de certaines faces. Ces liaisons fortes forment une ou plusieurs chaînes, appelée PBC (Periodic Boundary Chains), qui sont parallèles à la face. H et P distinguent ainsi trois types différents de faces. Les faces K (kinked faces) n'ont pas de PBC. Ces faces peuvent croître rapidement, depuis n'importe quel site à la surface sans nécessité de nucléation. Les faces F (Flat faces) ont elles plus de deux PBC. Elles doivent donc croître par couches, ce qui est un processus plus lent. Enfin, les faces S n'ont qu'une PBC et sont donc un intermédiaire entre les faces K et F. Un schéma de ces faces est représenté ci-dessous.



**Figure 9 : Schéma des faces F, S et K dans la théorie PBC. Faces F : (100), (010) et (001) ; Faces S : (110), (101) et (011) ; Faces K : (111). Figure issue de Hartmann *et coll.* (1973a, p375)<sup>62</sup>**

Selon la théorie de H et P, les faces F croissent le plus lentement et sont donc les faces les plus représentées au niveau de la morphologie du cristal. Seules les liaisons de la première sphère de coordination sont prises en compte. L'effet d'un solvant n'est pas pris en compte.

(3) **Le modèle d'Ising** implique le calcul d'énergies sur les couches parallèles aux faces. Il est basé sur la température de transition ordre-désordre dans un réseau bidimensionnel de liaisons contenues dans la couche du cristal.

Dans les modèles H-P et d'Ising, seulement les interactions des premiers voisins sont prises en compte. Ces méthodes sont souvent plus pertinentes pour des cristaux covalents.

(4) **Le modèle de l'énergie d'attachement** calcule, pour chaque surface, l'énergie d'attachement d'une nouvelle couche sur cette dernière. Les faces présentant les énergies d'attachement les plus basses auront la croissance la plus lente et ont donc le plus d'importance morphologique. Ce modèle nécessite donc une description physique des interactions entre les atomes, en y incluant les interactions de longue distance : électrostatiques et Van der Waals.

Les quatre méthodes précédentes ont des limites clairement définies. Tout d'abord, elles ne prennent pas en compte un possible réarrangement des atomes présents à la surface du cristal. De plus, elles échouent à prendre en compte l'effet d'un solvant. Ces deux processus peuvent avoir un effet sur la stabilisation de certaines surfaces et donc avoir un impact important sur la morphologie cristalline finale.

(5) Une dernière approche est la **minimisation de l'énergie de surface du cristal**. Cette méthode nécessite le calcul de l'énergie de surface de toutes les surfaces possibles du matériau pour prédire *in fine* la morphologie. Celle-ci minimise alors l'énergie de surface totale du cristal.

#### 1.6.2.2. Prédictions de la morphologie du COM

Après une analyse PBC, les méthodes de BFDH, d'Ising et l'énergie d'attachement ont été appliquées au cas du COM dans l'étude de Millan<sup>50</sup>. Les morphologies prédites sont décrites en Figure 10.

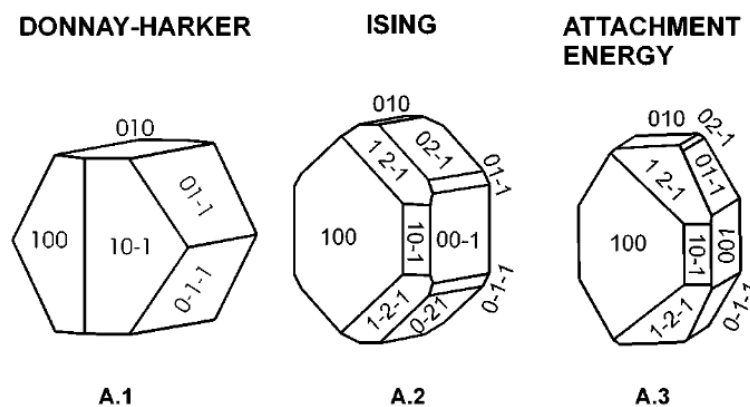


Figure 10 : Prédiction des morphologies du COM d'après Millan<sup>50</sup>

Ainsi les surfaces (100) et (010) qui sont décrites comme des surfaces de type F par la PBC, semblent être systématiquement présentes dans les prédictions de morphologies. Les surfaces (021) et (12 $\bar{1}$ ) sont également présentes dans les prédictions faites par les modèles d'Ising et de l'énergie d'attachement. Les résultats des deux dernières méthodes permettent d'expliquer la présence des surfaces (100), (010), (12 $\bar{1}$ ).

BFDH	Ising	Énergie d'attachement
(100)	(100)	(100)
(011)	(12 $\bar{1}$ )	(12 $\bar{1}$ )
(010)	(021)	(021)
<b>(10<math>\bar{1}</math>)</b>	(001)	(001)
	(010)	(010)

Tableau 2 : Surfaces du COM prédites par Millan<sup>50</sup>

En revanche un désaccord peut être soulevé pour la surface (001) qui est prédite mais jamais observée dans les synthèses. La surface (001) est cependant souvent observée dans les minéraux<sup>58</sup>. Millan suggère un effet cinétique. La surface (001) mettrait plus de temps à apparaître comparé aux autres surfaces<sup>50</sup>.

Surface	$\gamma_{vide}$ (J.m <sup>-2</sup> )	$\gamma_{hydraté}$ (J.m <sup>-2</sup> )
(001)	0.51	0.50
<b>(010)</b>	0.35	<b>0.11</b>
<b>(100)</b>	0.34	<b>0.14</b>
(110)	0.62	0.60
(011)	0.63	0.54
(111)	0.57	0.36
(101)	0.52	0.47
<b>(021)</b>	0.63	<b>0.37</b>
<b>(12<math>\bar{1}</math>)</b>	0.53	<b>0.20</b>

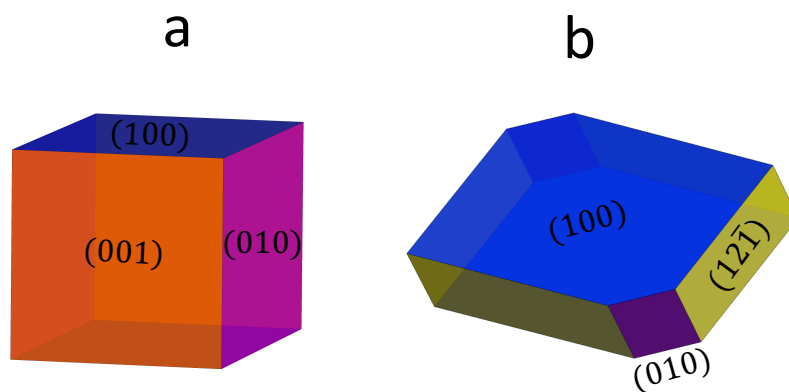
Tableau 3 : Prédiction des énergies de surface du COM dans l'article de Tommaso *et coll.*<sup>63</sup>

Enfin, la méthode de minimisation de l'énergie de surface a déjà été utilisée pour le COM dans l'étude de de Tommaso *et coll.* avec une description des interaction inter atomiques par champ de force.<sup>63</sup> Les énergies calculées sont présentée dans le Tableau 3.

Une fois les énergies de surface calculées, un diagramme de Wulff peut être construit pour prédire les morphologies cristallines. Le programme développé par Zucker *et coll.*<sup>64</sup> permet de construire un tel diagramme à partir des énergies de surfaces et en minimisant l'énergie totale du cristal :

$$E_{TOT} = \sum_i \gamma_i \quad (45)$$

Les prédictions obtenue par cette méthode, avec les énergies de Tommaso *et coll.*<sup>65</sup> sont présentées en Figure 11.



**Figure 11 : Prédiction de la morphologie du COM (a) selon les énergies de surface dans le vide (b) et selon les énergies de surface en présence d'eau, calculées par Tommaso *et coll.*<sup>65</sup> et obtenu avec le programme développé par Zucker *et coll.*<sup>64</sup>**

On remarque bien dans ces résultats l'importance cruciale de l'eau sur le calcul des énergies de surface et dans la prédiction de morphologie résultante. En effet, la surface  $(12\bar{1})$  n'apparaît que lorsque l'énergie de solvation (eau pure) est prise en compte (Figure 11b). Enfin, selon Tommaso *et coll.*<sup>65</sup>, l'effet du solvant permet de favoriser l'exposition de la surface (100) au dépend de la surface (001).

#### 1.6.2.3. Prédiction de morphologie du COD

La morphologie cristalline du COD a été prédite dans la littérature par les méthodes PBC, BFDH et de l'attachement de l'énergie<sup>59</sup>. L'analyse PBC prévoit l'existence de 5 surfaces de type F : (110), (101), (010), (121) et (211). Après une analyse de l'attachement de l'énergie,

seule trois surfaces ressortent : (101), (010) et (110). Aucune prise en compte de l'énergie d'hydrations des surfaces n'as été faite dans ces prédictions. La plupart des morphologies observées sont pourtant issues de synthèse en phase aqueuse.

### 1.6.3. Objectifs

Au cours de cette thèse, la méthode de minimisation de l'énergie de surface totale a été utilisée pour prédire la morphologie cristalline des phases monohydratée et dihydratée de l'oxalate de calcium. Pour cela, les énergies ont été calculées par des méthodes *ab-initio* (DFT) en prenant en compte les possible restructurations de surface et l'effet de l'eau. De tels calculs n'ont jamais été reportés pour les oxalates de calcium dans la littérature. Il sera intéressant de comparer les énergies de surfaces obtenue par DFT avec celles de Tommaso *et coll.*<sup>65</sup> obtenue par une méthode champ de force. Il sera également essentiel d'étudier l'impact de la solvation par l'eau sur le calcul des énergies de surfaces. L'influence de l'eau sur les énergies de surfaces du COD n'a jamais été étudiée.

### 1.6.4. Méthode

#### 1.6.4.1. Construction des modèles de surface

Dans un premier temps, il est nécessaire de construire un modèle moléculaire pour chacune des surfaces du matériau. Un modèle simple d'une telle surface peut être construit à partir des positions atomiques de la maille élémentaire en introduisant un espace vide entre les plans de Miller de la surface. Ainsi, deux surfaces de même indice sont créées comme cela est montré en Figure 12.

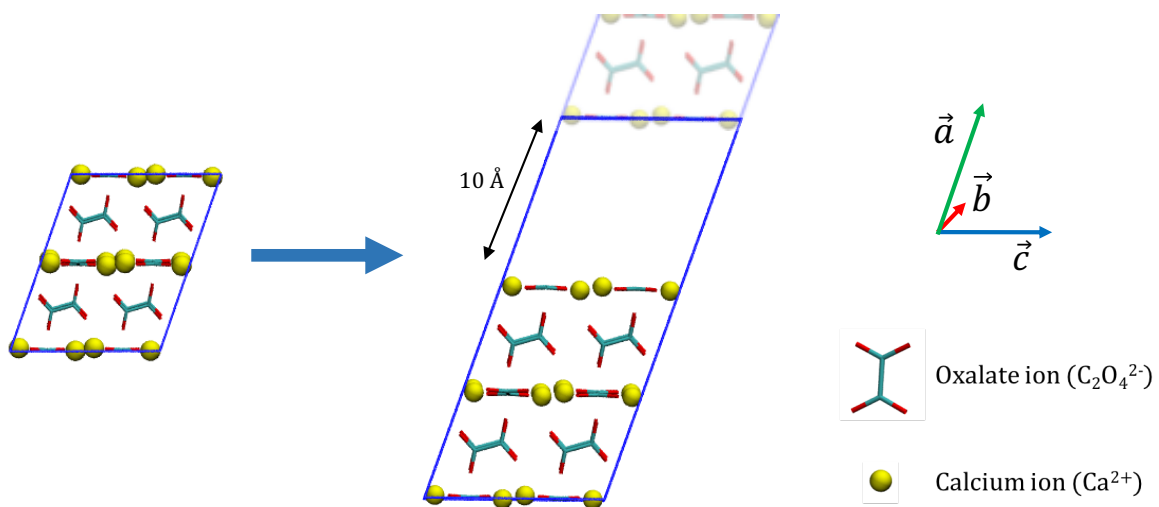


Figure 12 : exemple d'un modèle pour la surface (100) du COM.

L'énergie d'un tel modèle ainsi que l'énergie du bulk sont calculées par DFT après son optimisation géométrique. L'énergie de surface est alors calculée par l'équation suivante :

$$\gamma_{vide} = \frac{E_{surface} - E_{bulk}}{2A} \quad (46)$$

avec  $A$  l'aire de la surface,  $E_{surface}$  l'énergie du modèle de surface,  $E_{bulk}$  l'énergie relative au Bulk. Le facteur 2 est due à la création de deux surfaces de part et d'autre du « vide » ajouté. L'effet de  $N$  molécules l'eau peut être pris en compte en construisant un nouveau modèle d'une surface en interaction avec celles-ci. L'énergie de surface hydratée s'écrit alors :

$$\gamma_{hydratée} = \frac{E_{surface+H_2O} - E_{bulk} - NE_{H_2O}}{2A} \quad (47)$$

#### 1.6.4.2. Fonctionnelles utilisées

La fonctionnelles hybride PBE<sup>66</sup> a été utilisée avec une correction de Grimme de type D3<sup>15</sup>. Cette fonctionnelle a déjà été utilisée dans la littérature pour décrire les oxalates de calcium<sup>63,67,68</sup>. De plus, une étude comparative effectuée par Ivan Petit<sup>12</sup> montre que la fonctionnelle PBE permet de reproduire au mieux les paramètres de mailles expérimentaux comparé aux fonctionnelles BLYP<sup>69</sup> et OptPBE<sup>70</sup>. Le terme de dispersion de Grimme permet d'améliorer sensiblement la prédiction<sup>12</sup>. Nous avons tout de même testé ces deux dernières fonctionnelles pour comparer leur prédiction de l'énergie de surface.

#### 1.6.4.3. Optimisation géométrique

Tous les modèles (bulk et surfaces) ont été optimisés géométriquement avec le logiciel cp2k<sup>71</sup>. Les électrons sont donc décrits par une approche GPW. L'énergie obtenue après convergence est utilisée pour le calcul des énergies de surface dans le vide (équation 46).

#### 1.6.4.4. Prise en compte explicite de l'eau

Dans un premier temps, les molécules d'eau contenue dans la structure cristalline et se retrouvant à la surface du matériau du fait de la construction de la surface, sont supprimées. Puis, un nombre suffisant de molécule d'eau ( $1\text{kg.L}^{-1}$ ) a été ajouté pour combler le vide du modèle grâce à la commande *solvate* du logiciel GROMACS<sup>72</sup>. Une attention particulière a été apportée pour éviter que des molécules d'eau ne s'ajoutent à l'intérieur même de la structure. Le champ de force GROMOS 54a<sup>73</sup> a été utilisé afin de décrire les interactions du système. Le modèle *spce* a été choisi pour l'eau. Les positions des oxalates et des ions calciums sont fixées

pour ne pas dévier de la structure obtenue par l'optimisation géométrique en DFT. Le système est équilibré en dynamique moléculaire dans l'ensemble NVT. Ensuite, seules les deux premières couches de solvation ont été conservées pour l'optimisation géométrique en DFT en utilisant CP2K. L'optimisation de ces modèles a permis de calculer les énergies de surfaces hydratée (équation 47).

#### 1.6.4.5. Prédiction des morphologies cristallines

Le logiciel développé par Zucker *et coll.*<sup>64</sup> a été utilisé pour prédire la morphologie du COM et du COD en minimisant l'énergie de surface totale du cristal (équation 45).



## 1.7. Adsorption d'une catéchine sur les surfaces de l'oxalate de calcium

### 1.7.1. Problématique

#### 1.7.1.1. Bienfaits du thé vert

Le thé est l'une des boissons les plus populaires au monde grâce à ses saveurs mais également à ses vertus thérapeutiques. Un nombre considérable d'études épidémiologiques recensent les effets du thé vert sur la prévention du cancer<sup>74</sup>. La raison principale invoquée est liée à la présence d'une famille de molécules appelées *catéchines* dans le thé vert. En effet, les catéchines pouvant atteindre jusqu'à 30% de la masse de la feuille de thé<sup>75</sup> ont probablement des propriétés d'antioxydant. Leurs groupements phénols sont susceptibles d'agir sur des facteurs de transcriptions et sur l'activité enzymatique<sup>76</sup>, permettant de réduire le stress oxydatif des cellules. Ces résultats ont généré un intérêt certain pour cette famille de molécules dans la littérature. En particulier, l'EGCG ((-)-Epigallocatechin-3-gallate) est référencée comme étant la catéchine majoritaire dans le thé vert (entre 50 et 80% des catéchines)<sup>77</sup> (voir Figure 13).

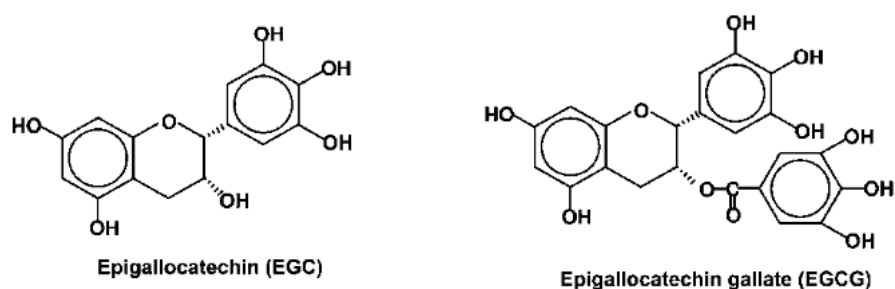
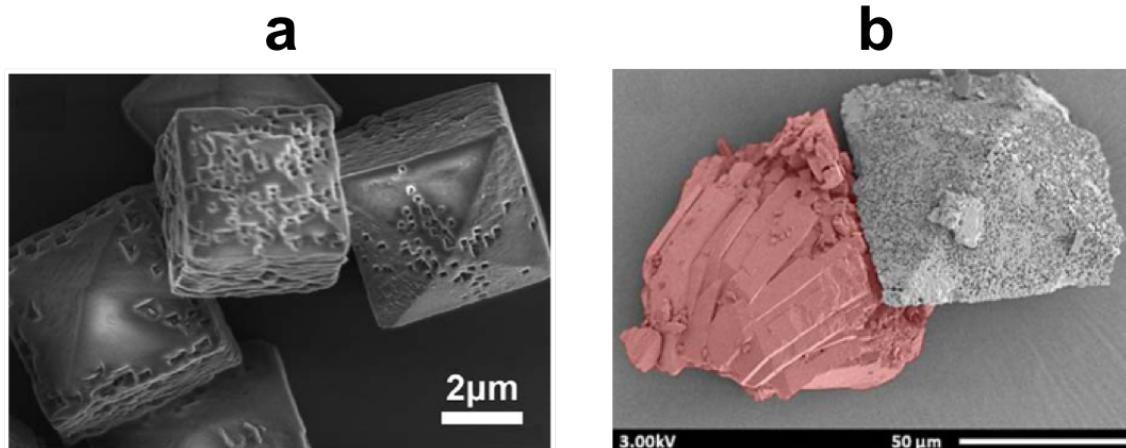


Figure 13 : Epigallocatechine (EGC) et (-)-Epigallocatechine-3-gallate (EGCG)

Par ailleurs, des études cliniques ont montré l'impact positif du thé vert sur l'inhibition de la précipitation de calculs rénaux<sup>78,79</sup>. En revanche, la raison reste floue. D'un côté, l'effet antioxydant des catéchines pourrait avoir une influence sur certains mécanismes biologiques ayant à leur tour un impact sur la régulation de la calcification dans le système urinaire<sup>78,79</sup>. D'un autre côté, les catéchines sont susceptibles d'avoir un effet direct sur la croissance cristalline des oxalates de calcium<sup>54</sup>.

Pour répondre à cette problématique, une première question posée est de savoir comment les catéchines sont métabolisées par l'organisme humain. Plusieurs études<sup>80,81</sup> mesurant par HPLC la concentration en catéchines dans l'urine après l'ingestion de thé vert, tendent à montrer que l'EGC se retrouve bien dans l'urine à une concentration détectable. L'EGCG en revanche, n'est pas détectée.

1.7.1.2. Effet du thé vert sur les oxalates de calcium



**Figure 14 :** (a) Image MEB de cristaux de COD en présence de thé vert. Résultats issus de Chen *et coll.*<sup>54</sup>. (b) COM (rouge) et COD synthétisés dans un gel d'agar en absence de thé vert puis plongé dans une infusion de thé vert. Figure issue de la thèse de Leroy<sup>55</sup>

Des synthèses d'oxalate de calcium en présence de thé vert ont permis d'observer l'influence de celui-ci sur leur processus de cristallisation.

(i) Le premier effet décrit concerne la sélectivité de la synthèse : en présence de thé vert, le produit cinétique COD est favorisé aux dépens du COM<sup>54,55,82</sup>. Le COD semble montrer une adhérence plus faible que le COM aux tissus biologiques des reins et serait alors éliminé plus efficacement que le COM, donnant ainsi aux catéchines un intérêt thérapeutique majeur<sup>83</sup>.

(ii) Selon Leroy<sup>55</sup>, le rendement des synthèses n'est pas modifié par la présence de thé vert. De plus, l'étude clinique de J. Rodes *et coll.*<sup>82</sup> ne montre pas de diminution de la quantité de calcul rénal chez les patients buvant du thé vert.

(iii) Néanmoins, des changements profonds sont observés sur l'apparence des cristaux au MEB<sup>54,55,82</sup> (voir Figure 14a). Des cavités de 200 à 500 nm apparaissent sur la surface (101) du COD. Les travaux de Leroy ont montré que le COM est beaucoup moins altéré que le COD lorsqu'il est plongé dans du thé vert (Figure 14b). Le thé vert dégraderait donc sélectivement le COD et non le COM.

(iv) Enfin, une élongation de la face (010) du COD est décrite dans l'étude de Chen *et coll.* (voir Figure 14a).

### 1.7.2. État de l'art

#### **Interactions de diverses molécules avec la surface des oxalates de calcium**

L'interaction entre plusieurs molécules et les surfaces d'oxalate de calcium a déjà été étudiée en dynamique moléculaire. Dans l'étude de Parvaneh *et coll.*<sup>68</sup>, l'interaction entre l'ion acétate et les surfaces (101) et (010) du COD est étudiée par DFT. Ils ont ainsi démontré une interaction préférentielle entre l'ion acétate et la surface (010) du COD.

De même, les interactions entre les surfaces du COM avec l'ion citrate ont été étudiées par une approche champ de force et *ab initio* dans l'étude de Qiu *et coll.*<sup>36</sup>.

Enfin, l'interaction entre une chondroïtine et le COM a été étudiée par champ de force dans l'étude de Hernandez *et coll.*<sup>84</sup>.

#### **Les catéchines en dynamique moléculaire**

L'interaction entre les catéchines et un modèle de bicouche lipidique a été étudiée par dynamique moléculaire<sup>85</sup>. Il a été montré que les catéchines peuvent interagir fortement avec la bicouche lipidique en formant des liaisons hydrogènes avec celle-ci.

Aucune étude théorique n'a été recensée décrivant les interactions entre une catéchine et les surfaces de l'oxalate de calcium.

### 1.7.3. Objectifs

L'objectif est de comprendre à l'échelle moléculaire l'interaction entre les catéchines et les surfaces d'oxalates de calcium COM et COD. Nous avons choisi d'étudier la catéchine EGCG (voir Figure 13) car elle est la plus représentée dans le thé vert. Plusieurs questions seront abordées. Pourquoi l'EGCG oriente la cristallisation vers le COD aux dépens du COM ? Comment l'EGCG interagit avec les surfaces du COM et du COD ? Quels sont les sites cristallographiques impliqués ? A quoi sont dus les « trous » observés sur la surface (101) du COD ? Pourquoi les catéchines permettent l'apparition de la surface (010) du COD ?

### 1.7.4. Méthode

#### 1.7.4.1. Construction des modèles

Pour le COM, trois modèles ont été construits concernant les surfaces (100), (010) et (12 $\bar{1}$ ). Pour le COD, deux modèles ont été construits concernant les surfaces (100) et (101). Le choix des surfaces étudiées est basé sur les morphologies observées dans les calculs rénaux et sur les

conclusions établies au chapitre 2. Comparé au chapitre 2 (Prédiction des morphologies cristallines des oxalates de calcium), la taille des modèles est largement augmentée afin d'éviter une possible interaction entre les images périodiques de l'EGCG. Les surfaces construites sont ainsi composées de plus de 6000 atomes, contre environ 200 pour les surfaces du chapitre 2.

#### 1.7.4.2. Description des interaction interatomiques

L'augmentation de taille des modèles n'est possible que si l'approche pour déterminer les interactions interatomiques est décrite par un champ de force. Le champs de force GROMOS 54a7<sup>86</sup> a été choisi afin de décrire les ions oxalates, les ions calciums et l'EGCG. Les paramètres de l'EGCG sont déterminés grâce à l'outil ATB (Automated force field Topology Builder)<sup>87</sup>. Cette approche est basée sur des calculs *ab initio* et sur des données expérimentales pour assurer la compatibilité avec les autres molécules décrites par le champ de force GROMOS. Nous avons conscience que les interactions entre la catéchine et l'oxalate de calcium pourrait faire l'objet d'une optimisation. Nous pensons cependant que cette description simple des interactions permettra de tirer des conclusions sur la sélectivité de l'EGCG vis-à-vis des différentes surfaces de l'oxalate de calcium. L'eau est décrite par le modèle *spce*.

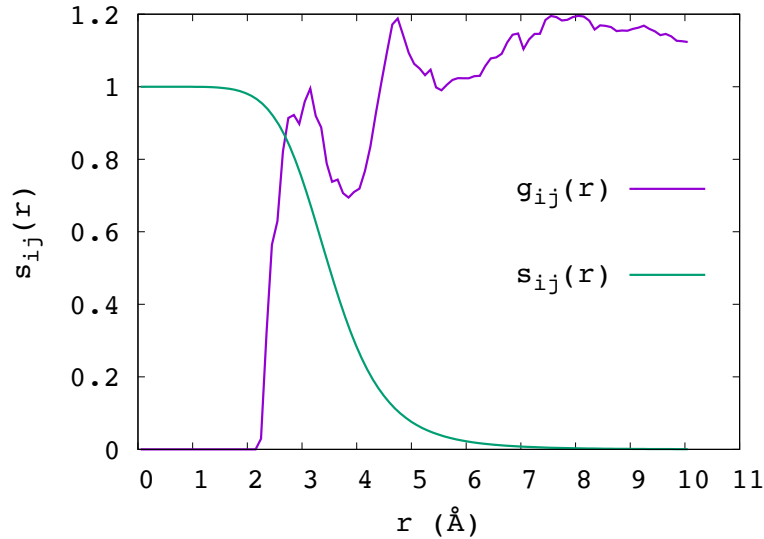
#### 1.7.4.3. Coordonnées utilisées pour décrire le processus d'adsorption

Deux coordonnées ont été choisies pour décrire l'adsorption de l'EGCG sur une surface de CaOx:

- **D** : la distance entre le centre de masse de l'EGCG et la surface de l'oxalate de calcium
- **CN** : le nombre de coordination entre les oxygènes de l'EGCG et les calcium et oxygènes à la surface de l'oxalate de calcium

$$CN = \sum_i \sum_j s_{ij} \quad s_{ij} = \frac{1 - \left(\frac{r_{ij}}{r_0}\right)^n}{1 - \left(\frac{r_{ij}}{r_0}\right)^m} \quad (48)$$

où  $i$  correspond aux atomes oxygène de l'EGCG et  $j$  aux atomes à la surface de l'oxalate de calcium.  $s_{ij}$  est une fonction comprise entre 0 and 1 qui décrit le contact entre l'atome  $i$  et l'atome  $j$ .  $s_{ij} = 0$  correspond à une absence de contact tandis que  $s_{ij} = 1$  correspond à un contact. Le paramètre  $r_0$  est la distance à laquelle  $s_{ij}$  est égal à 0.5. Les exposants  $n$  and  $m$  sont deux paramètres supplémentaires. Ces trois paramètres sont choisis pour prendre en compte totalement le premier pic de la fonction de distribution radiale des atomes  $i$  et  $j$  comme présenté en Figure 15.



**Figure 15:** Fonction  $s_{ij}(r)$  décrivant le contact entre les atomes d'oxygènes de l'EGCG (i) et les atomes de surface de CaOx (j) et fonction de distribution radiale  $g_{ij}(r)$  correspondant à ces deux types d'atome

Nous pensons que ces deux CV décrivent correctement le processus de désorption de la molécule. Il est associé à augmentation de la première CV et à une diminution conjointe de la deuxième CV.

#### 1.7.4.4. Amélioration de l'échantillonnage des configurations

La dynamique moléculaire standard ne permet pas de simuler l'adsorption de l'EGCG de manière réversible. C'est pourquoi nous avons fait l'usage de méthodes basées sur la métadynamique et la REMD permettant d'atteindre l'ergodicité. Quatre méthodes ont ainsi été comparées en prenant comme cas particulier l'adsorption de l'EGCG sur la surface (010) du COM.

Dans la **méthode 1**, l'EGCG est tout d'abord positionnée dans le vide au-dessus de la surface selon cinq orientations différentes. Une dynamique NVT est alors effectuée pour équilibrer l'EGCG et générer ainsi cinq configurations différentes où l'EGCG est adsorbée sur la surface. Puis, de l'eau est ajoutée pour compléter le vide de la boîte de simulation. Le système est alors rééquilibré dans l'ensemble NVT. Enfin, une simulation de métadynamique est réalisée en prenant comme variable collective **D** et **CN**. La position de l'EGCG est contrainte dans les directions x et y. Ainsi, l'EGCG ne peut que s'éloigner ou se rapprocher de la surface. La méthode 1 est une première représentation (naïve) de l'adsorption qui permettrait le calcul de l'énergie libre d'adsorption associée à une unique configuration adsorbée.

Dans la réalité, l'EGCG est probablement libre de visiter une multitude de configurations adsorbées. C'est pourquoi, dans les trois méthodes suivantes, nous avons décidé de ne plus contraindre les positions de l'EGCG. Différentes stratégies ont alors été comparées dans le but d'améliorer l'échantillonnage des configurations :

- Dans la **méthode 2**, la REMD seule a été utilisée.
- Dans la **méthode 3**, la métadynamique seule a été utilisée pour biaiser la DM le long des coordonnées (ou variables collectives) **D** et **CN**.
- Dans la **méthode 4**, nous avons tenté de tirer avantage des deux techniques en combinant la REMD avec la métadynamique.

## 1.8. Étude de la nucléation des phosphates de calcium

Dans le dernier chapitre, nous avons voulu étudier les premiers instants de la nucléation. Nous aurions pu continuer à travailler sur l'oxalate de calcium mais nous avons préféré nous concentrer dans un premier temps sur les phosphates de calcium, pour plusieurs raisons. Tout d'abord, les études expérimentales sur la nucléation sont plus importantes pour les phosphates de calcium que pour les oxalates de calcium. De plus, nous avons pu tirer avantage d'un champ de force conçu spécifiquement pour décrire les interactions du biominéral avec l'eau. Enfin, le laboratoire LCMCP et en particulier l'équipe SMILES s'intéresse actuellement à la prénucléation des phosphates de calcium. La méthodologie développée dans ce chapitre pourra cependant s'appliquer aux oxalates de calcium dans le futur.

### 1.8.1. Problématique

#### 1.8.1.1. Généralité sur les phosphates de calcium

L'hydroxyapatite (HAP), de formule brute  $\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$ , est la phase thermodynamiquement la plus stable du phosphate de calcium, dans les conditions physiologiques<sup>88,89</sup>. Elle constitue l'un des composants principaux de l'os (~60% en masse)<sup>90,91</sup> en s'intégrant dans une matrice organique composé de collagène (~22% en masse), de protéines non-collagéniques (~3% en masse), et de cellules (~2% en masse)<sup>90</sup>. Le rôle biologique de l'HAP est d'augmenter la rigidité du tissu, du fait de son module de Young élevé (80-120 GPa). L'os est également un matériau dynamique en interaction avec son milieu biologique. Il se forme et se dissout constamment et joue le rôle de réservoir ionique pour l'organisme vivant. L'hydroxyapatite se retrouve aussi dans l'émail des dents ou dans des calcifications pathologiques comme les calculs rénaux et la plaque de Randall<sup>92</sup>. Le processus de biominéralisation du tissu osseux a fait l'objet de nombreuses études ces dernières années. Différentes théories ont été émises pour tenter de décrire le processus de croissance de l'os. Comment l'organisme vivant parvient-il à orienter la formation d'HAP dans des tissus spécifiques ? Comprendre les mécanismes sous-jacents est également cruciale pour comprendre les facteurs pouvant déclencher des calcifications pathologiques.

Les ostéoblastes, des cellules présentes dans les tissus osseux sont connues pour être fortement impliquées dans le processus de biominéralisation. Il a été proposé dans les années 1920-30 que la nucléation des phosphates de calcium prendrait place à l'intérieur de vésicules en provenance des ostéoblastes. La nucléation serait alors déclenchée par une augmentation de la concentration intra vésiculaire en ions phosphates et calcium<sup>93,94</sup>. Cependant, des études dans les années 50

rendue possible par l'avènement de la microscopie électronique, ont permis l'observation de cristaux d'HAP dans le tissu osseux<sup>95-97</sup>. L'HAP a été retrouvé dans des zones spécifiques appelée zone de « gap » où la nucléation serait déclenchée. De plus, la croissance des cristaux d'HAP est orientée le long des fibrilles de collagène suggérant un mécanisme de nucléation hétérogène. Selon cette théorie, le collagène à lui seul serait capable de déclencher le processus de nucléation<sup>98,99</sup>. Des protéines non-collagéniques sont également suspectées d'avoir un rôle le déclenchement de la nucléation. Les protéines SIBLINGs par exemple, en interagissant avec les ions calcium, sont capables de d'augmenter la vitesse de nucléation<sup>100</sup>.

La stratégie commune à ces théories consiste à augmenter localement la concentration en phosphate et en calcium pour déclencher la nucléation. Au-delà de la complexité des mécanismes biologiques mise en place par l'organisme vivant pour réguler la biominéralisation, les mécanismes chimiques impliqués dans la nucléation des phosphates de calcium ne sont pas encore totalement compris.

#### 1.8.1.2. La théorie classique de la nucléation

La théorie de la nucléation classique (ou Classical Nucleation Theory, CNT) offre un premier cadre pour tenter de décrire cette transformation. Selon la CNT, la croissance du cristal devient favorable lorsque sa taille est suffisamment grande pour dépasser une barrière d'énergie libre. Cette barrière dépend d'une part de l'énergie libre négative (favorable) liée à la formation du bulk ( $V\Delta g$ ), et d'autre part, de l'énergie libre positive (défavorable)  $S\sigma$ , liée à la formation d'une nouvelle interface liquide/solide<sup>101</sup>. L'énergie libre de bulk est proportionnelle au volume  $V$  du cluster tandis que l'énergie libre de surface est proportionnelle à la surface  $S$ . L'énergie libre totale du système s'écrit alors :

$$\Delta G = V\Delta g + S\sigma \quad (49)$$

Si l'on approxime la forme du cluster à une sphère on obtient l'expression suivante :

$$\Delta G = \frac{4}{3}\Pi r^3 \Delta g + 4\Pi r^2 \sigma \quad (50)$$

Ainsi, pour une taille faible ( $r \ll$ ), le terme défavorable de l'énergie de surface est prépondérant tandis que le terme favorable d'énergie de bulk prend de l'importance avec l'augmentation de la taille du cluster. Le cluster doit alors dépasser une certaine taille critique pour que sa croissance devienne favorable énergétiquement. Cette taille peut être obtenue en dérivant l'expression de  $\Delta G$  dans l'équation 50 :



$$\frac{\partial(\Delta G)}{\partial r} = 0 \quad \rightarrow \quad r_c = \frac{-2\sigma}{\Delta g} \quad (51)$$

Pour dépasser la barrière, le système doit donc présenter une fluctuation de densité suffisante pour lui permettant d'atteindre  $r_c$ .

#### 1.8.1.3. La théorie non classique de nucléation

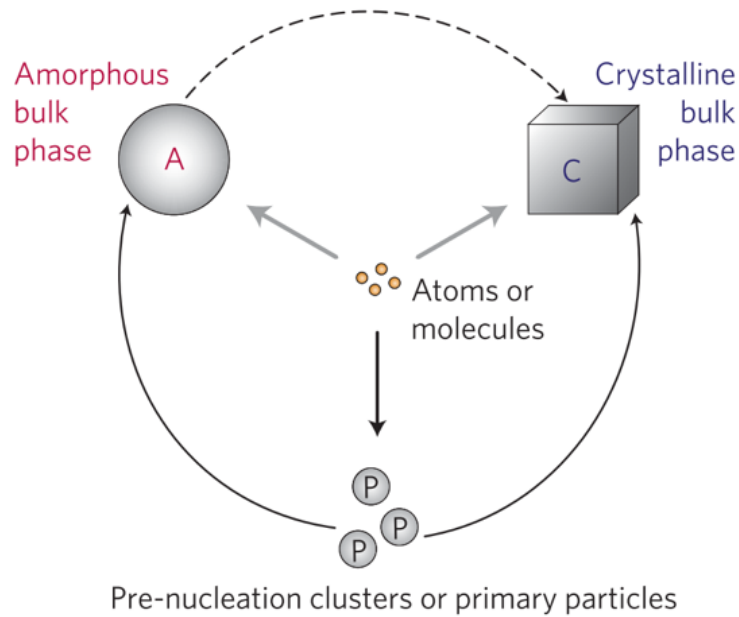


Figure 16 : les différents chemins envisageables dans un processus de nucléation. Figure tirée de la revue de De Yoreo<sup>102</sup>

Depuis une dizaine d'année, plusieurs études expérimentales ont reporté la présence de clusters de prénucléation dans des solutions ioniques supersaturés<sup>103-105</sup>. Ces études proposent qu'il existe un autre chemin de nucléation plus complexe que dans la théorie classique, où la solution supersaturée passerait par un état métastable constitué de nanoparticules ou de clusters qui s'agrègeraient ensuite entre eux pour former une phase intermédiaire amorphe. Cette phase amorphe se transformerait alors en une phase cristalline. La Figure 16 issue de la revue de De Yoreo<sup>102</sup> illustre les différents chemins envisageables pour un processus de nucléation.

#### 1.8.1.4. Etudes de la nucléation des phosphates de calcium

Dans le cas des phosphates de calcium, la théorie classique de la nucléation a été remise en question dans plusieurs études expérimentales, pour plusieurs raisons :

(i) la présence de **complexes de prénucléation** métastables.

Des complexes de prénucléation sont décrits dans l'article d'Habraken *et coll.*<sup>106</sup> comme étant de formule brute  $[\text{Ca}(\text{HPO}_4)_3]^{4-}$  (voir Figure 17) après une mesure par titration du rapport Ca/P durant les premiers instant du processus de nucléation. Le rapport mesuré est Ca/P=0,3–0,4. L'article propose également la présence d'un équilibre entre ces complexes et la paire d'ions  $[\text{Ca}(\text{HPO}_4)]$ .



Figure 17 : complexe  $[\text{Ca}(\text{HPO}_4)_3]^{4-}$ . Image adaptée de Mancardi *et coll.*<sup>107</sup>

Ces mesures ont été faites à 19°C, loin de la température physiologique.

L'étude par BOMD (PBE) de Mancardi *et coll.*<sup>107</sup> confirme la stabilité de l'entité  $[\text{Ca}(\text{HPO}_4)_3]^{4-}$  comparée aux autres complexes  $[\text{Ca}(\text{HPO}_4)_2]^{2-}$  et  $[\text{Ca}(\text{HPO}_4)]$  et aux ions libres solvatés.

(ii) la présence de **clusters de prénucléation** métastables.

Dey *et coll.* observent la formation de clusters de taille  $0.87 \pm 0.2$  nm dans une solution de SBF, à 37°C. Ces clusters s'agglomèrent ensuite entre eux pour former un objet décrit comme une chaîne polymérique par Habraken *et coll.*<sup>106</sup>. Ces motifs présentés en Figure 18 sont observés dans les deux études de Habraken *et coll.* et Dey *et coll.*

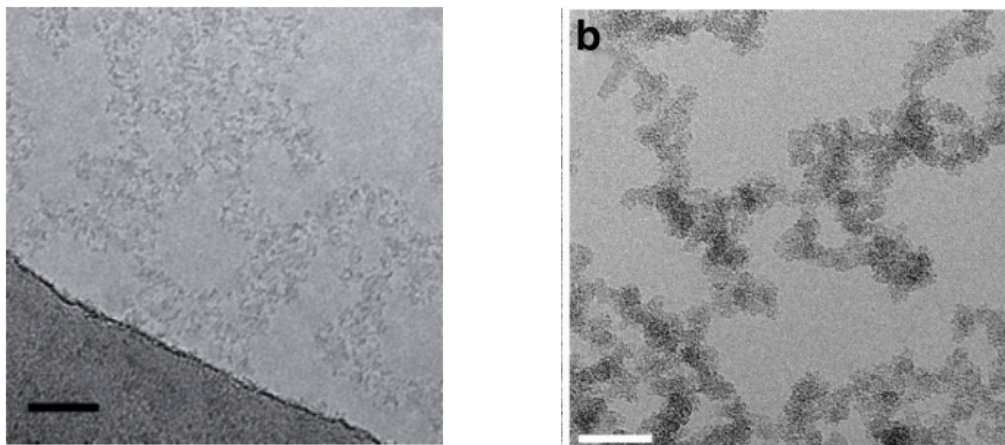


Figure 18 : image CryoTEM de l'étape 2 de nucléation des phosphates de calcium. La première image correspond à l'étude de Dey *et coll.*<sup>104</sup>, la deuxième image à celle de Habraken *et coll.*<sup>106</sup>. Les barre d'échelle sont toutes les deux de 50nm

Durant cette étape, le rapport Ca/P mesuré par Habraken *et coll.* reste autour de 0,3 suggérant une association de complexes  $[\text{Ca}(\text{HPO}_4)_3]^{4-}$  (Figure 17) constituant alors les sous-unité du cluster de prénucléation. Cette dernière hypothèse semble étonnante étant donnée la charge du complexe et les répulsions électrostatiques attendues par l'agglomération de plusieurs complexes de ce type.

(iii) le passage par une phase **amorphe (Amorphous Calcium Phosphate, ACP)** avant la cristallisation

Dans l'étude de Dey *et coll.*<sup>108</sup>, en présence d'une monocouche d'acide arachidique, une densification est observée uniquement près de la surface. Cette densification se transforme en particules de tailles allant de 40 à 80 nm. Une analyse par aux rayons X montre que ces particules sont composées d'ACP. Ces mêmes particules d'ACP sont décrites par Habraken *et coll.*<sup>106</sup> Il est soutenu par ces derniers que l'ACP se forme à partir des chaînes polymériques par l'incorporation de calcium pour augmenter le ration Ca/P jusqu'à atteindre celui de l'ACP. Des particules de taille moyenne 120 nm composé cette fois ci d'une phase cristalline sont observés plus tard par Dey *et coll.* confirmant l'hypothèse selon laquelle la cristallisation de l'HAP passe d'abord par la phase d'ACP.

De plus, une forme transitoire amorphe de phosphate de calcium a également été observée *in vivo* dans plusieurs matériaux biologiques en formation (dent, os, coquillages)<sup>109</sup>.

Il a été démontré par Posner que l'ACP est composé d'une sous unité de formule brute  $\text{Ca}_9(\text{PO}_4)_6$  appelée cluster de Posner ayant une taille autour de 1 nm<sup>110</sup>. Cette taille est proche de celle des clusters observé expérimentalement par Dey *et coll.*<sup>108</sup>. Les clusters de prénucléation seraient donc peut-être des clusters de Posner. Cette hypothèse rentre cependant en contradiction avec les complexes de prénucléation décrits par Habraken *et coll.* (voir Figure 17). En effet, la présence de cluster de Posner de formule brut  $\text{Ca}_9(\text{PO}_4)_6$  durant cette étape impliquerait un rapport Ca/P de 1,5 contre 0,3 selon l'hypothèse d'Habraken *et coll.*

### 1.8.2. État de l'art

Étant donné les limitations dues à la puissance de calcul, il est impossible de simuler en dynamique moléculaire des systèmes caractérisés par les concentrations décrites dans les expériences<sup>106,108</sup>. Une première approche pour étudier les processus de prénucléation des phosphates de calcium est de simuler par MD une solution aqueuse supersaturée en ions phosphates et calcium. Une telle approche a été étudiée par Mancardi *et coll.*<sup>65</sup>. Dans cet article, une boîte de  $50 \times 50 \times 50$  Å remplie d'eau et avec 44 ions  $\text{Ca}^{2+}$  et un mélange de  $\text{HPO}_4^{2-}/\text{H}_2\text{PO}_4^-$

permet de simuler une solution ionique de concentration  $[Ca^{2+}] = 0,65$  M. La supersaturation suffit alors à observer la formation de clusters de prénucléation décrit comme ressemblant aux clusters de Posner.

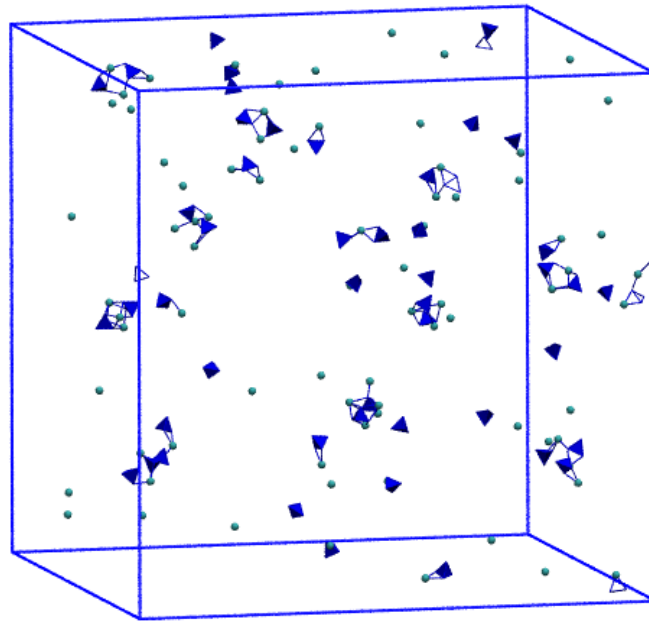
### 1.8.3. Objectifs

Actuellement, le rôle des clusters de prénucléation de phosphate de calcium n'est pas clairement établi. L'objectif est d'apporter des clarifications sur le mécanisme et sur la stabilité des structures intermédiaires dans le processus de nucléation. Pour ce faire, nous nous sommes appuyés sur une technique originale (décrite plus loin) pour simuler en DM la transformation d'une solution supersaturée de phosphates de calcium vers une nanoparticule de taille nanométrique ( $\sim 2$  nm), *réversiblement*. Notre méthode ouvre la voie à l'étude de l'influence des paramètres clés (température, concentration, additifs ...) de la nucléation.

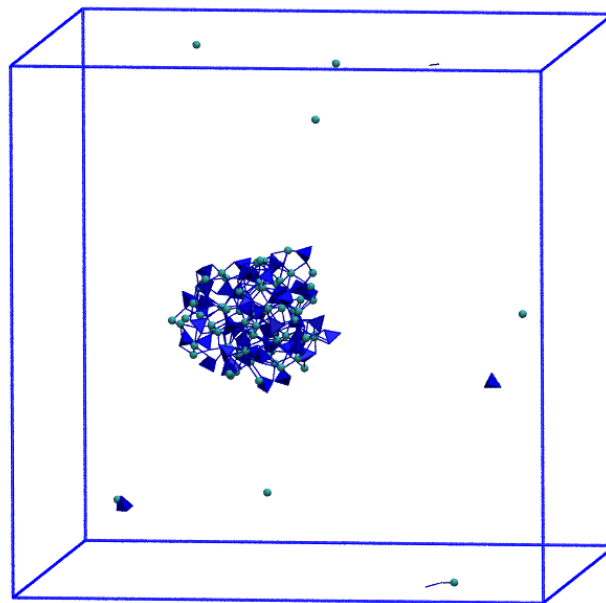
### 1.8.4. Méthode

#### 1.8.4.1. Construction des modèles

Deux boîtes de simulation de taille  $80 \times 80 \times 80$  Å ont été construites. La première boîte contient une solution ( $1 \text{ kg} \cdot \text{L}^{-1}$  d'eau) supersaturée avec 78 ions  $Ca^{2+}$ , 52 ions  $PO_4^{3-}$  correspondant au rapport Ca/P du Tri Calcium Phosphate. La deuxième boîte comporte une nanoparticule sphérique de diamètre 2 nm construite à partir de la structure cristallographique du TCP<sup>111</sup> grâce au logiciel *CrystalMaker*. Le choix de cette phase a été réalisé afin de respecter le rapport Ca/P de 1,5 et assurer la neutralité de la boîte de simulation. Les modèles de la solution supersaturée et de la nanoparticule sont présentés en Figure 19.



Supersaturated solution



Amorphous nanoparticle



Figure 19 : haut : solution supersaturée de phosphate de calcium, bas : nanocristal de phosphate de calcium  $[Ca_{72}(PO_4)_{50}]^{6-}$  dans une boîte de  $80 \times 80 \times 80 \text{ \AA}$ . Le diamètre du cluster est de  $20 \text{ \AA}$

Pour les deux modèles la procédure suivante est utilisée :

Dans un premier temps, la géométrie du système est optimisée avec la minimisation d'énergie GROMACS. Puis, une DM NPT de 1 ns est réalisée afin d'équilibrer la densité du système. Enfin, une DM NVT de 1 ns pour la solution supersaturée et de 20 ns pour la nanoparticule sont réalisées pour équilibrer les deux systèmes. La DM de la solution supersaturée a été arrêtée à 1 ns pour éviter la formation de clusters qui peuvent se former spontanément.

#### 1.8.4.2. Champs de force de Lin *et coll.*

Les paramètres du champ de force développé par Lin *et coll.*<sup>112</sup> ont été optimisés afin de reproduire les données expérimentales relatives à l'HAP, ainsi que les énergies d'hydratation des surfaces. Ce champ de force a été utilisé dans cette étude. Les paramètres du champ de force ont été dérivés avec le modèle tip3p de l'eau. Ce même modèle a été utilisé dans toutes les simulations. Nous avons dans un premier temps fait une étude comparant le champ de force avec une méthode DFT (BLYP). Dans cette étude comparative, les énergies libres associées à la dissolution de la surface (100) de l'HAP ont été calculées et comparées entre le champ de force et la DFT. La procédure sera plus détaillée au chapitre 4.

#### 1.8.4.3. Path Permutation Invariant Vector (*path-PIV*)

Comme dit précédemment, nous souhaitons simuler réversiblement la transition entre le modèle de la solution supersaturée et le modèle de la nanoparticule (voir Figure 19). La dynamique moléculaire basée sur les champs de force est en théorie capable de décrire une telle transformation. Cependant les barrières cinétiques impliquées sont probablement trop grandes pour décrire la transformation de manière réversible. C'est pourquoi nous avons utilisé la métadynamique<sup>32</sup> pour améliorer l'échantillonnage des configurations et reconstruire le diagramme d'énergie libre. Le choix de bonnes variables collectives est alors essentiel pour décrire correctement la transformation. Les variables collectives utilisées doivent être capables de distinguer correctement l'état final et l'état initial de la transformation ainsi que tous les états intermédiaires impliqués. Pour cela, le vecteur de permutation invariant (ou Permutation Invariant Vector, *PIV*) est une métrique pertinente introduite par Gallet *et coll.*<sup>113</sup> et par Pipolo *et coll.*<sup>114</sup> Les composants du vecteur sont obtenus selon l'équation 52 :

$$v_{kk'}^{\beta\beta'} = c_{kk'} f \left( \sqrt[3]{\frac{\Omega_0}{\Omega}} \| \mathbf{R}_{\beta k} - \mathbf{R}_{\beta' k'} \| \right) \quad (52)$$

$v_{kk'}^{\beta\beta'}$  correspond à une paire entre le  $\beta$ -ème atome de type  $k$  et le  $\beta'$ -ème atome de type  $k'$ .  $\| \mathbf{R}_{\beta k} - \mathbf{R}_{\beta' k'} \|$  est la distance entre ces deux atomes et  $f$  est une fonction permettant de décrire le contact entre les deux atomes. Cette fonction est la même que celle décrite dans l'équation 48. La constante  $c_{kk'}$  peut être égale à 0 ou 1 et est utile pour retirer certaines paires d'atomes dans le calcul du PIV.  $\Omega$  et  $\Omega_0$  sont respectivement le volume de la boîte de simulation et un volume de référence.  $\sqrt[3]{\frac{\Omega_0}{\Omega}}$  est un paramètre permettant d'éviter de trop grandes variations du volume de la boîte de simulation. Ce paramètre n'est pas pertinent dans le cas d'une simulation dans l'ensemble NVT où il est égal à 1. Pour obtenir le PIV, le vecteur est divisé en compartiments  $v_{kk'}$  correspondant à un couple de types d'atomes. Puis, chaque compartiment est trié en ordre croissant pour assurer l'invariance du vecteur lors de l'échange des positions de deux atomes de même type. La différence entre deux configurations peut être obtenue en prenant la distance euclidienne entre leur deux PIV. Dans notre cas, les atomes de phosphore et les atomes de calcium ont été impliqués dans le calcul du PIV.

Le PIV peut être combiné avec les variable collective de chemin (path collective variable, *path-CV*). Dans cette approche, les *path-CV* sont définie comme ci-dessous :

$$spath(t) = \frac{e^{-\lambda D(\mathbf{R}(t), \mathbf{R}_i)} + 2e^{-\lambda D(\mathbf{R}(t), \mathbf{R}_f)}}{e^{-\lambda D(\mathbf{R}(t), \mathbf{R}_i)} + e^{-\lambda D(\mathbf{R}(t), \mathbf{R}_f)}} \quad (53)$$

$$zpath(t) = -\frac{1}{\lambda} \log \left( e^{-\lambda D(\mathbf{R}(t), \mathbf{R}_i)} + e^{-\lambda D(\mathbf{R}(t), \mathbf{R}_f)} \right) \quad (54)$$

$\mathbf{R}_i$  représente l'ensemble des positions de l'état initial, c'est-à-dire la solution supersaturée.  $\mathbf{R}_f$  représente l'ensemble des positions de l'état final, c'est-à-dire du modèle de la nanoparticule.  $\lambda$  est un paramètre d'échelle.  $D(\mathbf{R}(t), \mathbf{R}_{i/f})$  est la distance euclidienne entre le PIV à un instant  $t$  et le PIV de la configuration initiale/finale. *spath* est le chemin direct reliant l'état initial à l'état final. *spath* = 1.1 quand le système correspond à la solution supersaturée ( $D(\mathbf{R}(t), \mathbf{R}_i) = 0$  et  $\lambda D(\mathbf{R}(t), \mathbf{R}_f) \approx 2.3$ ) et *spath* = 1.9 quand le système correspond à la nanoparticule ( $\lambda D(\mathbf{R}(t), \mathbf{R}_i) \approx 2.3$  et  $\lambda D(\mathbf{R}(t), \mathbf{R}_f) \approx 0$ ). Ajouter un biais le long de cette CV déplace le système entre l'état initial et final. Au contraire, ajouter un biais le long de *zpath*

éloigne le système à la fois de l'état initial et de l'état final. Cela permet d'explorer d'autres chemins impliquant des états de transitions différents à la fois de l'état initial et de l'état final. Dans la méthode *path-PIV*, seuls les états initial et final de la transformation sont nécessaires pour définir les CV, sans faire aucune hypothèse sur le chemin reliant les deux états. L'utilisation de ces CV a déjà fait ses preuves dans l'étude du diagramme de phase de l'eau<sup>114</sup>. Nous l'avons appliqué pour la première fois à la problématique de la nucléation des phosphates de calcium.



1. Daudon, M. *et al.* Examination of whewellite kidney stones by scanning electron microscopy and powder neutron diffraction techniques. *J. Appl. Crystallogr.* **42**, 109–115 (2009).
2. McQuarrie, D. A. (Donald A. *Quantum chemistry*. (University Science Books, 2008).
3. Atkins, P. & Friedman, R. *Molecular quantum mechanics*. (2011).
4. Finocchi, F. *Density Functional Theory for Beginners Basic Principles and Practical Approaches*. (2011).
5. Slater, J. C. Atomic Shielding Constants. *Phys. Rev.* **36**, 57–64 (1930).
6. Hoggan, P. E. & Ruiz, M. B. Molecular Integrals over Slater-type Orbitals . From pioneers to recent progress. (2009).
7. Lippert, G., Hutter, J. & Parrinello, M. The Gaussian and augmented-plane-wave density functional method for ab initio molecular dynamics simulations. *Theor. Chem. Accounts Theory, Comput. Model. (Theoretica Chim. Acta)* **103**, 124–140 (1999).
8. Goedecker, S., Teter, M. & Hutter, J. Separable dual-space Gaussian pseudopotentials. *Phys. Rev. B* **54**, 1703–1710 (1996).
9. Hartwigsen, C., Goedecker, S. & Hutter, J. Relativistic separable dual-space Gaussian pseudopotentials from H to Rn. *Phys. Rev. B* **58**, 3641–3662 (1998).
10. Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **98**, 5648–5652 (1993).
11. Perdew, J. P. & Schmidt, K. Jacob's ladder of density functional approximations for the exchange-correlation energy. in *AIP Conference Proceedings* **577**, 1–20 (AIP, 2001).
12. Petit, I. Vers une meilleure description des interfaces entre biominéraux et milieux biologiques par une approche combinée théorique et expérimentale. (Sorbonne Université, 2016).
13. Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **27**, 1787–1799 (2006).
14. Grimme, S., Ehrlich, S. & Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **32**, 1456–1465 (2011).

15. Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate *ab initio* parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **132**, 154104 (2010).
16. Jorgensen, W. L., Maxwell, D. S. & Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. (1996).
17. BROYDEN, C. G. The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations. *IMA J. Appl. Math.* **6**, 76–90 (1970).
18. Verlet, L. Computer “Experiments” on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.* **159**, 98–103 (1967).
19. Hockney, R. ., Goel, S. . & Eastwood, J. . Quiet high-resolution computer models of a plasma. *J. Comput. Phys.* **14**, 148–158 (1974).
20. Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **52**, 255–268 (1984).
21. Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **31**, 1695–1697 (1985).
22. Hoover, W. G. Constant-pressure equations of motion. *Phys. Rev. A* **34**, 2499–2500 (1986).
23. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
24. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
25. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
26. Essmann, U. *et al.* A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577 (1995).
27. Car, R. & Parrinello, M. Unified Approach for Molecular Dynamics and Density-Functional Theory. *Phys. Rev. Lett.* **55**, 2471–2474 (1985).
28. Hukushima, K. & Nemoto, K. Exchange Monte Carlo Method and Application to Spin Glass Simulations. *J. Phys. Soc. Japan* **65**, 1604–1608 (1996).
29. Fukunishi, H., Watanabe, O. & Takada, S. On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: Application to protein structure prediction. *J.*

- Chem. Phys.* **116**, 9058–9067 (2002).
30. Okamoto, Y. Generalized-ensemble algorithms: enhanced sampling techniques for Monte Carlo and molecular dynamics simulations. *J. Mol. Graph. Model.* **22**, 425–439 (2004).
  31. Laio, A. & Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci.* **99**, 12562–12566 (2002).
  32. Barducci, A., Bonomi, M. & Parrinello, M. Metadynamics. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **1**, 826–843 (2011).
  33. Franceschi, V. R. & Nakata, P. A. Calcium Oxalates in Plants: Formation and Function. *Annu. Rev. Plant Biol.* **56**, 41–71 (2005).
  34. Moe, O. W. Kidney stones: pathophysiology and medical management. *Lancet* **367**, 333–344 (2006).
  35. Daudon, M. Épidémiologie actuelle de la lithiase rénale en France. *Ann. Urol. (Paris)*. **39**, 209–231 (2005).
  36. Qiu, S. R. *et al.* Molecular modulation of calcium oxalate crystallization by osteopontin and citrate. *Proc. Natl. Acad. Sci.* **101**, 1811–1815 (2004).
  37. Chen, W.-C., Lin, H.-S., Chen, H.-Y., Shih, C.-H. & Li, C.-W. Effects of Tamm-Horsfall Protein and Albumin on Calcium Oxalate Crystallization and Importance of Sialic Acids. *Mol. Urol.* **5**, 1–5 (2001).
  38. Ruiz-Agudo, E. *et al.* A non-classical view on calcium oxalate precipitation and the role of citrate. *Nat. Commun.* **8**, (2017).
  39. Grases, F., Rodriguez, A. & Costa-Bauza, A. Efficacy of Mixtures of Magnesium, Citrate and Phytate as Calcium Oxalate Crystallization Inhibitors in Urine. *J. Urol.* **194**, 812–819 (2015).
  40. Riley, J. M., Kim, H., Averch, T. D. & Kim, H. J. Effect of magnesium on calcium and oxalate ion binding. *J. Endourol.* **27**, 1487–92 (2013).
  41. Nancollas, G. H. *Biological Mineralization and Demineralization: Report of the Dahlem Workshop on Biological Mineralization and Demineralization Berlin 1981, October 18-23.* (Springer Berlin Heidelberg, 1982).
  42. Conti, C. *et al.* Stability and transformation mechanism of weddellite nanocrystals studied by X-ray diffraction and infrared spectroscopy. *Phys. Chem. Chem. Phys.* **12**, 14560 (2010).

43. Sorokin, I. *et al.* Epidemiology of stone disease across the world. *World J. Urol.* **35**, 1301–1320 (2017).
44. Arnott, H. J., Pautard, F. G. E. & Steinfink, H. Structure of Calcium Oxalate Monohydrate. *Nature* **208**, 1197–1198 (1965).
45. Tazzoli, V. & Domeneghetti, C. The crystal structures of whewellite and weddellite: re-examination and comparison. *Am. Mineral.* **65**, 27–334 (1980).
46. Deganello, S. & IUCr. The structure of whewellite, CaC<sub>2</sub>O<sub>4</sub>. H<sub>2</sub>O at 328 K. *Acta Crystallogr. Sect. B Struct. Crystallogr. Cryst. Chem.* **37**, 826–829 (1981).
47. Sterling, C. & IUCr. Crystal structure analysis of weddellite, CaC<sub>2</sub>O<sub>4</sub>.(2+x)H<sub>2</sub>O. *Acta Crystallogr.* **18**, 917–921 (1965).
48. Izatulina, A., Gurzhiy, V. & Frank-Kamenetskaya, O. Weddellite from renal stones: Structure refinement and dependence of crystal chemical features on H<sub>2</sub>O content. *Am. Mineral.* **99**, 2–7 (2014).
49. Bouropoulos, N., Weiner, S. & Addadi, L. Calcium Oxalate Crystals in Tomato and Tobacco Plants: Morphology and in Vitro Interactions of Crystal-Associated Macromolecules. *Chem. Eur. J.* **7**, 1881–1888 (2001).
50. Millan, A. Crystal Growth Shape of Whewellite Polymorphs: Influence of Structure Distortions on Crystal Shape. *Cryst. Growth Des.* **1**, 245–254 (2001).
51. Aquilano, D. & Franchini-Angela, M. Twin laws of whewellite, CaC<sub>2</sub>O<sub>4</sub>.H<sub>2</sub>O. A structural and growth approach. *Phys. Chem. Miner.* **7**, 124–129 (1981).
52. Chen, Z., Wang, C., Zhou, H., Sang, L. & Li, X. Modulation of calcium oxalate crystallization by commonly consumed green tea. *CrystEngComm* **12**, 845 (2010).
53. Franchini-Angela, M. & Aquilano, D. Theoretical growth morphology of whewellite CaC<sub>2</sub>O<sub>4</sub>.H<sub>2</sub>O. *Phys. Chem. Miner.* **10**, 114–120 (1984).
54. Chen, Z., Wang, C., Zhou, H., Sang, L. & Li, X. Modulation of calcium oxalate crystallization by commonly consumed green tea. *Cryst. Eng. Comm* **12**, 845–852 (2010).
55. Leroy, C. Oxalates de calcium et hydroxyapatite : des matériaux synthétiques et naturels étudiés par techniques RMN et DNP. (Université P. et M. Curie, 2016).
56. Giordani, P., Modenesi, P. & Tretiach, M. Determinant factors for the formation of the

- calcium oxalate minerals, weddellite and whewellite, on the surface of foliose lichens. *Lichenologist* **35**, 255–270 (2003).
57. Zhang, D., Qi, L., Ma, J. & Cheng, H. Morphological Control of Calcium Oxalate Dihydrate by a Double-Hydrophilic Block Copolymer. (2002). doi:10.1021/cm010768y
  58. Goldschmidt, V. Atlas der Krystallformen (1913 edition) | Open Library. *Heidelberg: Winter*
  59. Heijnen, W. M. M. & Van Duijneveldt, F. B. The theoretical growth morphology of calcium oxalate dihydrate. *J. Cryst. Growth* **67**, 324–336 (1984).
  60. Daudon, M., Jungers, P. & Traxer, O. *Lithiase urinaire*. (Médecine sciences publications-[Lavoisier], 2012).
  61. Docherty, R., Clydesdale, G., Roberts, K. J. & Bennema, P. Application of Bravais-Friedel-Donnay-Harker, attachment energy and Ising models to predicting and understanding the morphology of molecular crystals. *J. Phys. D: Appl. Phys.* **24**, 89–99 (1991).
  62. Haussühl, S. P. Hartman (Ed.): Crystal Growth: An Introduction. North-Holland series in crystal growth Vol. 1. North-Holland Publishing Company, Amsterdam-London 1973. 531 S., Preis: \$ 21.10. *Berichte der Bunsengesellschaft für Phys. Chemie* **78**, 1276–1276 (1974).
  63. Tommaso, D. Di, Hernández, S. E. R., Du, Z. & Leeuw, N. H. de. Density functional theory and interatomic potential study of structural, mechanical and surface properties of calcium oxalate materials. *RSC Adv.* **2**, 4664 (2012).
  64. Zucker, R. V., Chatain, D., Dahmen, U., Hagège, S. & Carter, W. C. New software tools for the calculation and display of isolated and attached interfacial-energy minimizing particle shapes. *J. Mater. Sci.* **47**, 8290–8302 (2012).
  65. Mancardi, G., Hernandez Tamargo, C. E., Di Tommaso, D. & de Leeuw, N. H. Detection of Posner's clusters during calcium phosphate nucleation: a molecular dynamics study. *J. Mater. Chem. B* **5**, 7274–7284 (2017).
  66. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
  67. Colas, H. *et al.* Whewellite, CaC<sub>2</sub>O<sub>4</sub>.H<sub>2</sub>O: structural study by a combined NMR, crystallography and modelling approach. *Crystengcomm* **15**, 8840–8847 (2013).

68. Parvaneh, L. S., Donadio, D. & Sulpizi, M. Molecular Mechanism of Crystal Growth Inhibition at the Calcium Oxalate/Water Interfaces. *J. Phys. Chem. C* **120**, 4410–4417 (2016).
69. Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **38**, 3098–3100 (1988).
70. Klimeš, J., Bowler, D. R. & Michaelides, A. Chemical accuracy for the van der Waals density functional. *J. Phys. Condens. Matter* **22**, 022201 (2010).
71. VandeVondele, J. *et al.* Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Comput. Phys. Commun.* **167**, 103–128 (2005).
72. Van Der Spoel, D. *et al.* GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).
73. Reif, M. M., Hünenberger, P. H. & Oostenbrink, C. New Interaction Parameters for Charged Amino Acid Side Chains in the GROMOS Force Field. *J. Chem. Theory Comput.* **8**, 3705–3723 (2012).
74. Bettuzzi, S. *et al.* Chemoprevention of Human Prostate Cancer by Oral Administration of Green Tea Catechins in Volunteers with High-Grade Prostate Intraepithelial Neoplasia: A Preliminary Report from a One-Year Proof-of-Principle Study. *Cancer Res.* **66**, 1234–1240 (2006).
75. Graham, H. N. Green tea composition, consumption, and polyphenol chemistry. *Prev. Med. (Baltim)*. **21**, 334–350 (1992).
76. Higdon, J. V. & Frei, B. Tea Catechins and Polyphenols: Health Effects, Metabolism, and Antioxidant Functions. *Crit. Rev. Food Sci. Nutr.* **43**, 89–143 (2003).
77. Singh, B. N., Shankar, S. & Srivastava, R. K. Green tea catechin, epigallocatechin-3-gallate (EGCG): Mechanisms, perspectives and clinical applications. *Biochem. Pharmacol.* **82**, 1807–1821 (2011).
78. Hirose, M. *et al.* Renal tubular epithelial cell injury and oxidative stress induce calcium oxalate crystal formation in mouse kidney. *Int. J. Urol.* **17**, 83–92 (2010).
79. Itoh, Y. *et al.* Preventive Effects Of Green Tea On Renal Stone Formation and the Role of Oxidative Stress in Nephrolithiasis. *J. Urol.* **173**, 271–275 (2005).

80. Lee, M.-J. *et al.* Pharmacokinetics of tea catechins after ingestion of green tea and (-)-epigallocatechin-3-gallate by humans: formation of different metabolites and individual variability. *Cancer Epidemiol. Biomarkers Prev.* **11**, 1025–32 (2002).
81. Yang, C. S. *et al.* Blood and urine levels of tea catechins after ingestion of different amounts of green tea by human volunteers. *Cancer Epidemiol. Biomarkers Prev.* **7**, 351–4 (1998).
82. Julie, R. *et al.* Influence of daily green tea infusions on hypercalciuric renal stone patients: biological and morpho-constitutional analysis. **46**, 8.27 (2016).
83. Benzerara, Y., Bazin, D., Daudon, M., Haymann, J.-P. & Rode, J. Catechin for treating Renal Lithiasis. (2015).
84. Hernandez, S. E. R. & De Leeuw, N. H. Effect of Chondroitin 4-Sulfate on the Growth and Morphology of Calcium Oxalate Monohydrate: A Molecular Dynamics Study. *Cryst. Growth Des.* **15**, 4438–4447 (2015).
85. Sirk, T. W., Brown, E. F., Sum, A. K. & Friedman, M. Molecular Dynamics Study on the Biophysical Interactions of Seven Green Tea Catechins with Lipid Bilayers of Cell Membranes. *J. Agric. Food Chem.* **56**, 7750–7758 (2008).
86. Schmid, N. *et al.* Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur. Biophys. J.* **40**, 843–856 (2011).
87. Malde, A. K. *et al.* An Automated Force Field Topology Builder (ATB) and Repository: Version 1.0. *J. Chem. Theory Comput.* **7**, 4026–4037 (2011).
88. Wang, K. *et al.* Theoretical analysis of protein effects on calcium phosphate precipitation in simulated body fluid. *CrystEngComm* **14**, 5870–5878 (2012).
89. Zawacki, S. J., Nancollas, G. H., Lore, M., Richardson, C. & Perez, L. Mineral phases of calcium phosphate. *Anat. Rec.* **224**, 234–241 (2005).
90. Zapanta LeGeros, R. Apatites in biological systems. *Prog. Cryst. Growth Charact.* **4**, 1–45 (1981).
91. Weiner, S. & Wagner, H. D. THE MATERIAL BONE: Structure-Mechanical Function Relations. *Annu. Rev. Mater. Sci.* **28**, 271–298 (2002).
92. Bazin, D. & Daudon, M. Nouvelles Méthodes d'étude des calculs et plaques de Randall. *Actual. Néphrologiques Jhamburg. Maison Chim.* 75–98 (2010).

93. Robison, R. The Possible Significance of Hexosephosphoric Esters in Ossification. *Biochem. J.* **17**, 286–93 (1923).
94. Robison, R. & Soames, K. M. The Possible Significance of Hexosephosphoric Esters in Ossification. *Biochem. J.* **18**, 740–754 (1924).
95. Robinson, R. A. & Watson, M. L. Collagen-crystal relationships in bone as seen in the electron microscope. *Anat. Rec.* **114**, 383–409 (1952).
96. Robinson, R. A. An electron-microscopic study of the crystalline inorganic component of bone and its relationship to the organic matrix. *J. Bone Joint Surg. Am.* **34-A**, 389–435; passim (1952).
97. Glimcher, M. J. Molecular Biology of Mineralized Tissues with Particular Reference to Bone. *Rev. Mod. Phys.* **31**, 359–393 (1959).
98. Freeman Rutgers, J. W., Landis, W. J., Silver, F. H. & Freeman, J. W. Collagen as a scaffold for biomimetic mineralization of vertebrate tissues. *J. Mater. Chem.* **16**, 1495–1503 | (2006).
99. Wang, Y. *et al.* The predominant role of collagen in the nucleation, growth, structure and orientation of bone apatite. *Nat. Mater.* **11**, 724–733 (2012).
100. Glimcher, M. J. Mechanism of calcification: Role of collagen fibrils and collagen-phosphoprotein complexes in vitro and in vivo. *Anat. Rec.* **224**, 139–153 (1989).
101. Gibbs, J. W. A. du texte. The collected works / of J. W. Gibbs,... (1928).
102. De Yoreo, J. More than one pathway. *Nat. Mater.* **12**, 284–285 (2013).
103. Pouget, E. M. *et al.* CaCO<sub>3</sub> Formation Revealed by Cryo-TEM. *Science (80-. )*. **323**, 1455–1458 (2009).
104. Dey, A. *et al.* The role of prenucleation clusters in surface-induced calcium phosphate crystallization. *Nat. Mater.* **9**, 1010–1014 (2010).
105. Gebauer, D., Kellermeier, M., Gale, J. D., Bergström, L. & Cölfen, H. Pre-nucleation clusters as solute precursors in crystallisation. *Chem. Soc. Rev.* **43**, 2348–2371 (2014).
106. Habraken, W. J. E. M. *et al.* Ion-association complexes unite classical and non-classical theories for the biomimetic nucleation of calcium phosphate. *Nat. Commun.* **4**, 1507–1512 (2013).
107. Mancardi, G., Terranova, U. & de Leeuw, N. H. Calcium Phosphate Prenucleation Complexes



- in Water by Means of ab Initio Molecular Dynamics Simulations. *Cryst. Growth Des.* **16**, 3353–3358 (2016).
108. Dey, A. *et al.* The role of prenucleation clusters in surface-induced calcium phosphate crystallization. *Nat. Mater.* **9**, 1010–1014 (2010).
  109. Beniash, E., Metzler, R. A., Lam, R. S. K. & Gilbert, P. U. P. A. Transient amorphous calcium phosphate in forming enamel. *J. Struct. Biol.* **166**, 133–143 (2009).
  110. Betts, F. & Posner, A. S. An X-ray radial distribution study of amorphous calcium phosphate. *Mater. Res. Bull.* **9**, 353–360 (1974).
  111. Yashima, M., Sakai, A., Kamiyama, T. & Hoshikawa, A. Crystal structure analysis of  $\beta$ -tricalcium phosphate  $\text{Ca}_3(\text{PO}_4)_2$  by neutron powder diffraction. *J. Solid State Chem.* **175**, 272–277 (2003).
  112. Lin, T.-J. Force Field Parameters and Atomistic Surface Models for Hydroxyapatite and Analysis of Biomolecular Adsorption At Aqueous Interfaces. (2013).
  113. Gallet, G. A. & Pietrucci, F. Structural cluster analysis of chemical reactions in solution. *J. Chem. Phys.* **139**, 074101 (2013).
  114. Pipolo, S. *et al.* Navigating at Will on the Water Phase Diagram. (2017).

## - Chapter 2 -

# Morphologies of Calcium Oxalate Polyhydrates: A Quantum Chemical and Computational Study

*Théau Debroise<sup>1</sup>, Thomas Sedzik<sup>1</sup>, Christian Bonhomme<sup>1</sup> and Frederik Tielens<sup>2,\*</sup>*

<sup>1</sup> Sorbonne Université, CNRS, Laboratoire Chimie de la Matière Condensée de Paris, LCMCP, 4 Place Jussieu, F-75005 Paris, France

<sup>2</sup> General Chemistry (ALGC) Materials Modeling Group, Vrije Universiteit Brussel (Free University Brussels-VUB), Pleinlaan 2, 1050 Brussel, Belgium

\*Author to whom correspondence should be addressed: [frederik.tielens@vub.be](mailto:frederik.tielens@vub.be)

<b>Abstract</b> .....	<b>73</b>
<b>Introduction</b> .....	<b>74</b>
<b>Methods</b> .....	<b>81</b>
<b>Results and discussion</b> .....	<b>84</b>
<b>Conclusion</b> .....	<b>94</b>
<b>Acknowledgements</b> .....	<b>95</b>
<b>References</b> .....	<b>95</b>

## **Abstract**

Monohydrated (COM) and Dihydrated (COD) **Calcium Oxalate** have been extensively studied in the literature because of their role in **urolithiasis**, a mammal pathology responsible for the formation of stones in the kidney. It is now clear that the physicochemical environment plays a crucial role in the control of crystal growth and the morphologies of calcium oxalates. To study these effects, it is crucial to have reliable models for the calcium oxalate's faces, exposed to water and potential additives.

Here we have used total surface energy minimization to predict **crystal morphology** of the calcium oxalate monohydrate and dihydrate phases. Surface energies were calculated at **Density Functional Theory (DFT)** level taking into account surface relaxation and **solvation effect**. An excellent agreement was found between predictions and experimental morphologies observed by SEM. Our results demonstrate the importance of water on the morphology predictions.

## **Introduction**

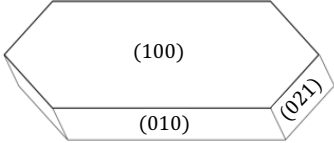
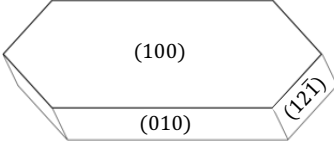
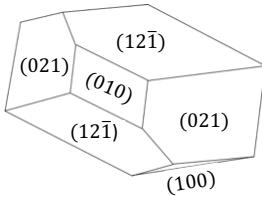
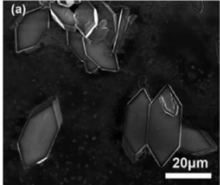
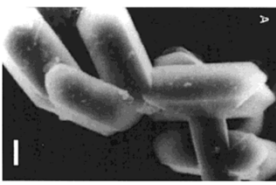
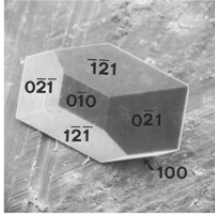
Calcium oxalate polyhydrates are ionic crystals present in minerals and living organisms. They are found in most plant families under the name raphides<sup>1</sup>. The functions associated with the calcium oxalate crystals in plants are diverse. They can be used for protection against herbivores or play a structural role. Calcium oxalates biomineralization is not a random physicochemical process but is well directed by living organisms. Indeed, it has been shown that the quantity, morphology and size of calcium oxalate crystals depends on the taxonomy of the plant and is therefore subject to genetic control<sup>2</sup>. The mechanisms by which these living organisms regulate crystal growth are complex, can occur at several stages and are mainly due to interactions between COM surfaces and additives (ions, organic molecules, proteins). Calcium oxalates have also been extensively studied because of their role in urolithiasis. This pathology originates from the formation of stones in the urinary system called kidney stones<sup>3,4</sup>. The main constituents of kidney stones are now established and correspond to a limited number of well-defined crystalline phases: calcium oxalates (with a different hydration rate), ammonium, calcium and magnesium phosphates and uric acid. In recent decades, calcium oxalate-based stones are the most frequently observed (> 70% of all samples analyzed, in Europe and North America)<sup>3,4</sup>.

### ***COM and COD crystallographic structures***

From the crystallographic point of view, the calcium oxalate family is classified according to their hydration rate: anhydrous calcium oxalate (ACO), whewellite ( $\text{CaC}_2\text{O}_4 \cdot \text{H}_2\text{O}$ ; COM), weddellite ( $\text{CaC}_2\text{O}_4 \cdot 2\text{H}_2\text{O}$ ; COD) and caoxite ( $\text{CaC}_2\text{O}_4 \cdot 3\text{H}_2\text{O}$ ; COT). An amorphous phase is also described in nanoparticles<sup>5</sup>. In this study we focused on COM and COD as they are the two main phases observed in kidney stones<sup>4,6</sup>. ACO and COT are not observed in kidney stones. COM structure has been solved by X-Ray and neutron diffraction<sup>7-12</sup>. Two COM polymorphs are described: the first structure is stable in the range 318-425K and the second in the range 293-318K. Since, we were interested in studying kidney stones obtained at  $\approx 310$  K, only the second structure was studied. It is a monoclinic structure belonging to the space group  $P2_1/c$ . The COD structure was solved by X-ray diffraction<sup>9,13,14</sup>. It is a tetragonal structure belonging to the space group  $I 4/m$ .

### COM morphologies *in vitro*

COM has been observed by Scanning Electron Microscopy (SEM) in several *in vitro* syntheses in aqueous medium<sup>15–18</sup>. Four faces came out as predominant from these studies: (010), (100), (021) and (12 $\bar{1}$ ). Tazzoli *et al.*<sup>9</sup> convention is used for Miller indices of COM. (as described by Millan<sup>16</sup>). The (010) and (100) faces are always observed. One of the (021) or (12 $\bar{1}$ ) face is always present as well, sometimes both of them. The corresponding morphologies are presented in Table 1.

Morphology	a) 	b) 	c) 
SEM			
Experimental conditions	Synthesis in water, equimolar, T <sub>amb</sub>	Synthesis in water, T=37°C	Synthesis in U tubes filled with silica gel, equimolar, T <sub>amb</sub>
Reference	Chen <i>et al.</i> <sup>19</sup>	Millan <sup>16</sup>	Franchini-Angela <i>et al.</i> <sup>17,18</sup>

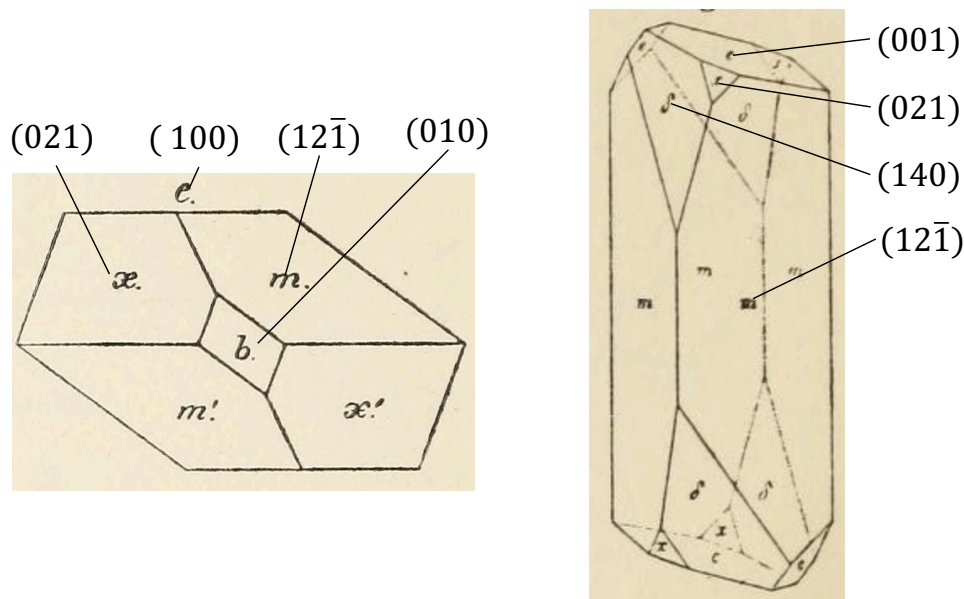
**Table 1: COM morphologies observed by SEM for different synthetic procedures<sup>16–19</sup>.**

### COM morphologies *in vivo*

Kidney stones are never composed of a single crystal. They always consist of crystals of different chemical compositions in an organic matrix of proteins and lipids<sup>20</sup>. COM morphologies have been observed in kidney stones<sup>12,20</sup>. In this environment, COM crystals present morphologies similar to those presented in Table 1a, b. (010), (021) and (12 $\bar{1}$ ) faces are less expressed giving crystals a flat aspect. COM crystals are stacked on (100) face with proteins between them.

### ***COM morphologies in minerals***

Compared to synthetic and kidney stone whewellite, mineral whewellite crystals exhibit a wider diversity in morphology<sup>7</sup>. (100), (010), (021) and  $(12\bar{1})$  faces have still a strong occurrence but contrary to synthetic and kidney stones COM, (001) face appears to be one of the most exposed face in minerals. An example of mineral COM morphology can be seen in Figure 1. A large family of other faces is observed more rarely in few crystals: (140),  $(13\bar{1})$ ,  $(12\bar{3})$ ,  $(14\bar{1})$ , (161),  $(140)$ , (011),  $(16\bar{1})$ ,  $(15\bar{1})$ , (013), (121), (112) and (031).



***Figure 1: Example of two mineral COM crystal morphologies from Goldschmidt<sup>7</sup>.***

### ***COM morphologies predictions***

Hartman-Perdok (HP), Bravais-Friedel-Donnay-Harker (BFDH), Ising Model and attachment energy method have been used by Millan<sup>16</sup> to predict COM morphology. HP theory<sup>21</sup> is based on the concept of periodic bound chain (PBC) which is a sequence of strong bonds formed by the crystalline particles (molecules, ions, atoms) parallel to the face. HP theory distinguishes three kinds of faces (*ie* K,S,F): K faces have no PBC, S faces have a PBC in one dimension and F faces have more than two PBC. According to HP theory, only F faces are exposed. The basic idea is that F faces grow more slowly because they need to form strong bonds in two dimensions. BFDH analysis<sup>22</sup> gives a simple procedure to establish a hierarchy between F face. It assumes that face importance is related to its interplanar distance. The greater the interplanar distance, the slower the growth and the more the face is expressed. According to this simple

model, (100), (001), (021), (12 $\bar{1}$ ) and (10 $\bar{2}$ ) faces are predicted. In Ising model<sup>23</sup>, energy calculations are done on crystal slices parallel to the face. It is based on the order-disorder transition temperature in a two-dimensional network of bonds contained in the crystal layer. The faces mentioned above were predicted from this model, as well. In the attachment energy method<sup>24</sup>, the potential energy associated to the addition of a layer to each surface is calculated. Except for (10 $\bar{2}$ ) face (which is absent), the same faces were predicted again with this last method.

Finally, Tommasso *et al.*<sup>25</sup> have predicted COM morphology in a force field study. They found a stabilization of (100) face due to a solvation effect (pure water). Such a stabilization is not observed on (001) face. This would explain why (001) face is not observed in aqueous syntheses.

Except for this last study, the main difference between theory and experiment originates from the (001) face which is always predicted but never observed in syntheses nor in kidney stones. This may be due to the growth kinetics. Indeed, Milan<sup>16</sup> observed that faces appeared in a defined order: first (100), then (010) and later (021) and (12 $\bar{1}$ ). Furthermore, the (001) face is prevalent in minerals<sup>7,16</sup> suggesting that this face may take more time to appear.

### ***COD morphologies in vitro***

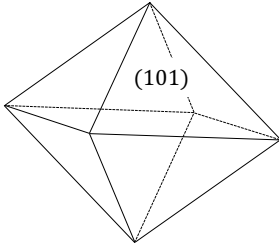
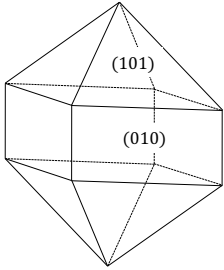
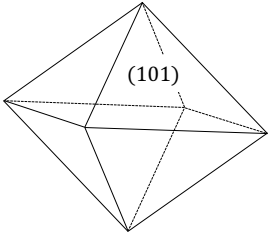
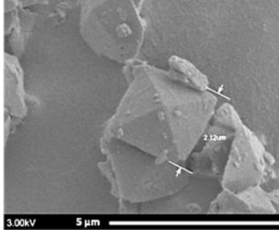
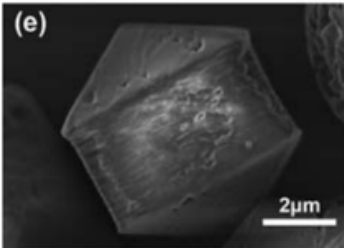
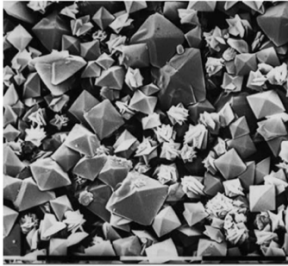
COD morphologies have also been reported in many syntheses. As shown in Table 2a, c, COD crystals have often the form of a bipyramid dominated by (101) faces. Under certain conditions, the (010) face is intercalated between the two pyramids formed by the (101) faces (see Table 2b). Leroy<sup>26</sup> and Giordani *et al.*<sup>27</sup> syntheses reported the absence of the (010) face while this face is observed only in the presence of green tea, as in the Chen *et al.*<sup>19</sup> synthesis (see Table 2b). Green tea is assumed to have therapeutic properties, that is why its effect was investigated by Chen *et al.* Catechins, components of green tea are known to interact with calcium oxalate crystals. In another study, PEG-b-PMAA significantly stabilizes the (010) face<sup>28</sup>. Therefore, it appears that the (010) face exposition requires an additional stabilization by interacting molecules.

### ***COD morphologies in vivo***

COD is less observed than COM in kidney stones but it is also able to aggregate in order to form COD stones<sup>29</sup>. COD crystals in kidney stones, adopt the bipyramid morphology with the predominant (101) face<sup>29,30</sup> (see Table 2a,c).

### ***COD morphologies in minerals***

COD crystals occurring in marine sediments have the same morphology than COD from synthetic and kidney stones ones, with predominant (101) face<sup>31</sup>.

Morphology	a) 	b) 	c) 
SEM		(e) 	
Experimental conditions	Synthesis in water, excess of oxalate, $T_{amb}$	Synthesis in water, equimolar conditions, presence of green tea, $T_{amb}$	Synthesis in a synthetic solution of urine in excess of calcium ions
Reference	Leroy thesis <sup>26</sup>	Chen <i>et al.</i> <sup>19</sup>	Giordani <i>et al.</i> <sup>27</sup>

***Table 2: COD morphologies observed by SEM for different synthetic procedures<sup>19,26,27</sup>.***

### ***COD morphology predictions***

COD morphologies have been predicted in the literature by HP, BFDH and attachment energy methods<sup>31</sup>. HP analysis predicts the existence of five F-type faces: (110), (101), (010), (121) and (211). From attachment energy analysis, only three faces are expressed: (101), (010) and (110). Differences between experiment and theory are evident for (110) and (010) faces. In the case of (010), we have showed above that the presence of additional molecules was necessary.



### ***Effect of additives on calcium oxalates crystal growth***

Many studies have focused on the effect of additives on the COM/COD crystal growth. It has been shown that various species can act as COM/COD growth inhibitors *in vitro*, such as proteins (osteopontin<sup>32</sup>, Tamm-Horsfall protein and albumin<sup>33</sup>), organic molecules (citrates<sup>34</sup>, catechins<sup>35</sup>, phytate<sup>36</sup>) or inorganic salts ( $Mg^{2+}$ )<sup>37</sup>. Furthermore, it has been demonstrated that phytate and citrate are able to inhibit COM crystal growth in urine<sup>36,38</sup> and that osteopontin plays an important role in kidney stone regulation<sup>39</sup>, *in vivo*. It is clear that the physicochemical environment is essential to understand the control of COM/COD crystal growth and morphologies, both *in vitro* and *in vivo*.

### ***Crystal Growth mechanism***

Calcium oxalate have a very low solubility ( $\approx 6.10^{-3} \text{ g.L}^{-1}$ )<sup>40</sup>. Like other low solubility derivatives (calcium carbonates and phosphates), the crystal growth is suspected to adopt a non-classical pathway<sup>5</sup> mainly due to kinetics reasons. Indeed, it has been demonstrated that calcium oxalate is able to precipitate first into an amorphous precursor phase in aqueous solution<sup>41</sup> before performing a transition from the amorphous phase to a crystalline phase. However, those amorphous particles are always observed in competitions with COM or COD crystalline phases. It is still not clear which pathway is preferred for the precipitation. Nevertheless, this phenomenon suggests that calcium oxalate precipitation should be sensitive to kinetics effects.

### ***Goals of the current work and novelty***

The scope of this contribution is to develop a theoretical approach being able to predict growth inhibition potential of targeted molecules' ability to inhibit calcium oxalate growth, in a drug design perspective. Theoretical methods such as molecular dynamics should be relevant for that purpose. Nevertheless, it is crucial to build relevant models for calcium oxalates exposed faces before studying the interaction between COM/COD and additives at an atomic level. Morphology prediction methods described above have strong limitations. First, they do not take into account a possible rearrangement of ions presents on the surface of the crystal. Second, they do not take into account explicitly the effect of the solvent. These two contributions may have an effect on the stabilization of particular faces and thus have an impact on the crystal morphology.

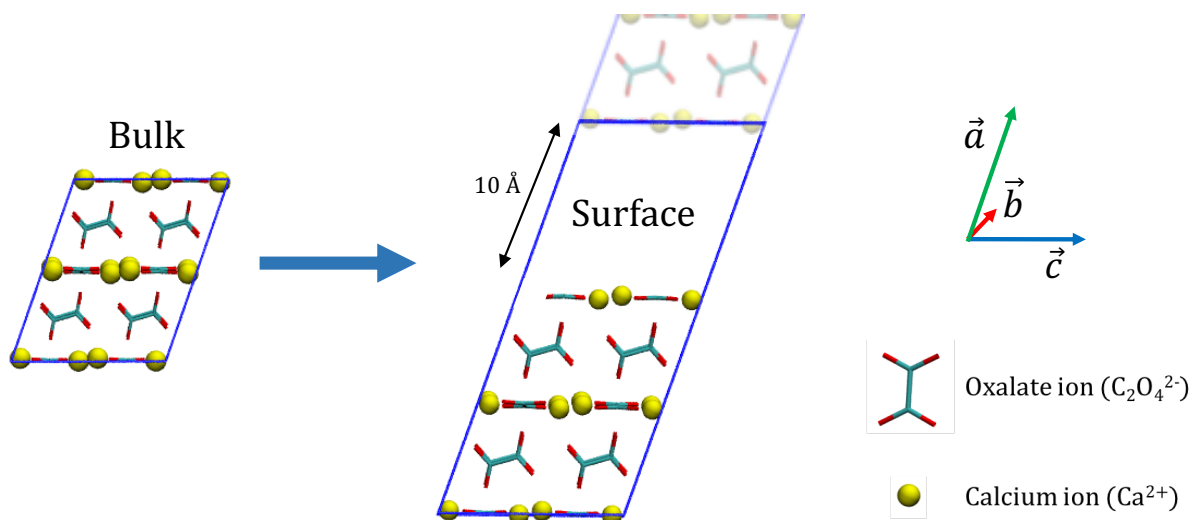
To take into account these external factors explicitly, we have used total surface energy minimization in order to predict COM and COD morphologies. This method predicts the

morphology which corresponds to the lowest total surface energy. Surface energies were calculated at a DFT level in the presence, or absence, of a solvation layer above the surfaces and with geometric relaxation.

## Methods

### Surfaces models

Bulk structures of COM and COD were obtained respectively from Daudon *et al.* (neutron diffraction)<sup>12</sup> and Tazzoli *et al.* (X-ray diffraction)<sup>9</sup> studies, respectively. Calcium oxalate surfaces were built by adding a vacuum of 10 Å in the bulk unit cell, along the low Miller indices (see Figure 2).



**Figure 2: COM (100) surface construction from COM bulk. Initial data from Daudon et al.**  
12

Finally, 11 surface models, with different miller indices, were built for COM and 11 for COD. The COM and COD surfaces are composed of 160 and 226 atoms, respectively.

### Geometry optimization

All geometry optimizations were carried out with the freely available DFT package CP2K/quickstep<sup>42</sup>. This software is based on the hybrid Gaussian and plane wave method<sup>43</sup>. Generalized Gradient Approximation Perdew-Burke-Ernzerhof (PBE) + D3 functional<sup>44</sup> was used. BLYP-D3<sup>45</sup> and OptPBE-vdW<sup>46</sup> functionals were tested in addition in order to compare surface energies calculations. Since our models involve oxalates ions, we introduced dispersion forces, by using the Grimme D3 method<sup>47</sup>. Analytic Goedecker–Teter–Hutter pseudopotentials<sup>48</sup>, a DZVP level basis set for the orbitals, and a density cutoff of 300 Ry were used.

### ***Hydration of surfaces***

Water molecules were added to fill the vacuum present in each surface model super cell with *solvate* command of the GROMACS simulation package (v 5.1.2)<sup>49</sup>. The average density of water molecules was checked to be at 1 kg.L<sup>-1</sup>. Gromos force field 54a7<sup>50</sup> was used to describe interactions between oxalate and calcium ions, and water molecules. SPC/E model<sup>51</sup> was chosen for water. Calcium and oxalate positions were frozen using P-LINCS algorithm<sup>52</sup> and a 500 ps simulation was performed to equilibrate water molecules. All simulations were run in the canonical ensemble (NVT) with a Nosé-Hoover thermostat to ensure thermal stability<sup>53,54</sup>. Particle Mesh Ewald (PME) method was used to optimize calculation of long-range interactions<sup>55,56</sup> with a grid spacing of 0.12 nm and an interpolation of 4. A real space cutoff of 0.5 nm was used, and the Lennard-Jones interactions were truncated at the same distance. For reasons of calculation costs, only the first two layers of solvation were kept before performing the geometry optimization of the hydrated surfaces at the DFT level. Indeed, one geometry optimization could take up to 48 hours of calculation on 128 cores, on GENCI supercomputer. A geometry optimization had to be performed for all “dry” and hydrated surface models, as well as for bulks. This led to 46 geometry optimizations in total.

### ***Surface energy calculations***

Surface energy  $\gamma_{vacuum}$  of dry surfaces was calculated as following:

$$\gamma_{vacuum} = \frac{E_{surface} - E_{bulk}}{2A} \quad (1)$$

with  $A$  the surface area,  $E_{surface}$  and  $E_{bulk}$  the potential energies of the surface and the bulk model.

Hydrated surfaces energies were calculated as following:

$$\gamma_{H_2O} = \frac{E_{surface,H_2O} - E_{bulk} - N \cdot E_{H_2O}}{2A} \quad (2)$$

$N$  correspond to the number of water molecules in solvation layers.  $E_{surface,H_2O}$  is the potential energy of the solvated surface.  $E_{H_2O}$  is the energy of one water molecule. It was obtained from

a NVT Born Oppenheimer Molecular Dynamic (BOMD) of 10 ps followed by a geometry optimization of 93 water molecules in a  $10 \times 10 \times 10 \text{ \AA}^3$  box with PBE-D3 functional.

### ***Crystal morphology prediction***

Crystal morphology predictions were made by the mean of Wulff diagrams and using either surface or hydrated surface energies. The predictions are based on the Gibbs<sup>57</sup> hypothesis that the equilibrium shape of a crystal is the one minimizing the total surface energy of the crystal.

$$\Delta G = \sum_j \gamma_j A_j \quad (3)$$

with  $\gamma_j$  the surface energy of a specific face, and  $A_j$  the area of the face. And so,  $\Delta G$  represent the difference in free energy between ions/molecules in a real crystal (with surfaces) and ions in an infinite crystal. Additionally, according to Gibbs<sup>58</sup>, the length of the vector  $h_j$  normal to the face and drawn between the center of the crystal and the face, is proportional to the surface energy:

$$h_j = k \cdot \gamma_j \quad (4)$$

Equations 3 and 4 lead to the Wulff-Gibbs theorem which is applied here. Software developed by Zucker *et al.*<sup>59</sup> was implemented for that purpose.

### ***Interface between calcium oxalate and bulk water***

Larger models were built in order to investigate the interface between predominant COM/COD surfaces and bulk water by using GROMOS force field 54a7<sup>50</sup>. These models were built from DFT relaxed surface models. The surface area was greatly increased by extension of the simulation box in two dimensions, parallel to the surface. These corresponding models have a surface 40 times larger than the DFT models and are composed of about 6000 atoms. Surface splitting was also increased from 10 to 50  $\text{\AA}$ . After water molecules additions (to completely fill the vacuum), the simulation box contained around 50000 atoms. Water density profiles were calculated from 1 ns NVT simulations. COM and COD atoms were constrained in the DFT equilibrium state.

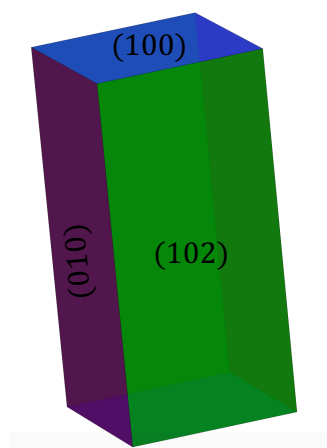
## **Results and discussion**

### ***The COM case***

In vacuum, the (010) face is the most stable, followed by the (102), (120) and  $(12\bar{1})$  faces. From the surface energies in vacuum  $\gamma_{vacuum}$ , only the (010), (102) and (100) faces showed up in the morphology prediction as can be seen in Figure 3, which not coherent with experimental observations. Emergence of (100) face in the morphology prediction may be surprising as it is less stable than other surfaces, which are not represented in Figure 3 (see Table 3). This raise the importance of the crystal shape in the Wulff prediction. Indeed, a shape may be preferred among others if its total surface is smaller.

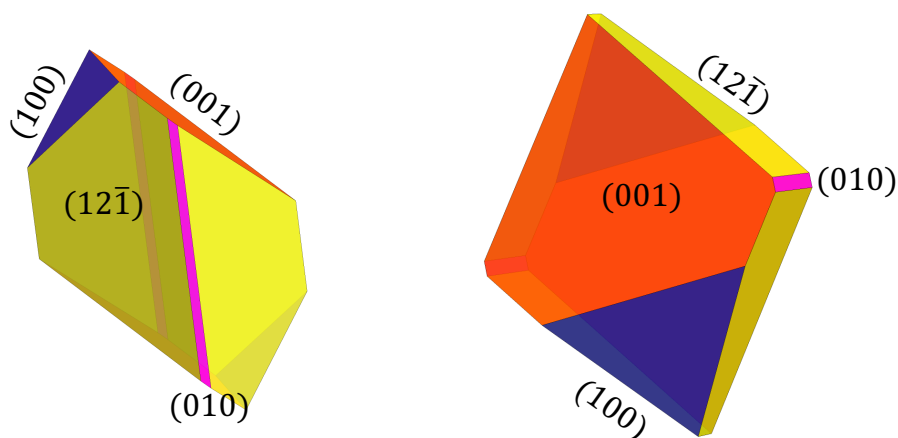
face	$\gamma_{vacuum}$	$\gamma_{H_2O}$
$(12\bar{1})$	0.50	0.33
(021)	0.62	0.39
(010)	0.28	0.39
(001)	0.63	0.42
(011)	0.76	0.46
(102)	0.52	0.51
(120)	0.55	0.51
(012)	0.65	0.59
(100)	0.65	0.68
(101)	1.19	0.91
(110)	0.58	1.04

***Table 3 : Surface Energies of Calcium Oxalate monohydrate (COM) in vacuum and with a bilayer of water. Surface energies in  $J.m^{-2}$ .***



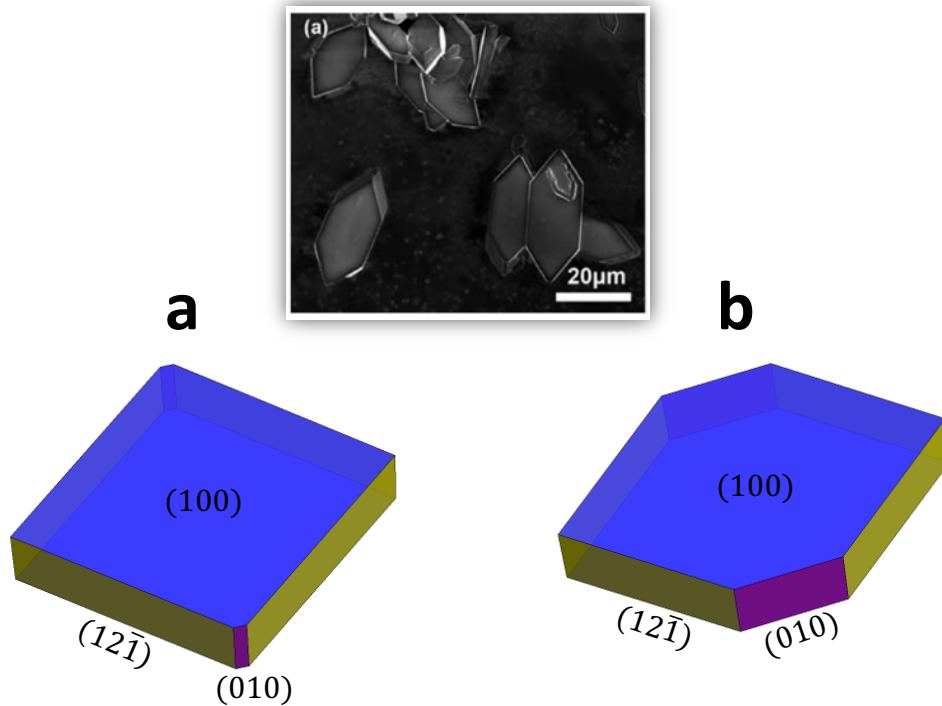
**Figure 3: Morphology prediction of COM crystal in vacuum and based on  $\gamma_{vacuum}$  values in Table 3.**

Adding a bilayer of water on the surfaces has a great impact on the surface energies (see Table 3). The  $(12\bar{1})$  and  $(021)$  faces are indeed strongly stabilized ( $-0.17 \text{ J.m}^{-2}$  and  $-0.23 \text{ J.m}^{-2}$ ), whereas the  $(010)$  face is slightly destabilized ( $+0.11 \text{ J.m}^{-2}$ ). As shown in Figure 4, solvation effect deeply affects the morphology, mainly due to the emergence of the  $(12\bar{1})$  face.



**Figure 4: Morphology prediction of COM crystal with water molecules (bilayer) and based on  $\gamma_{H_2O}$  values in Table 3.**

This morphology corresponds well to Wulff prediction obtained with force field, by Tommaso *et al.*<sup>25</sup>. Nevertheless, a disagreement stays for the  $(001)$  face, which is not observed in syntheses. As suggested by Milan<sup>16</sup>, the  $(001)$  face is not observed in syntheses due to kinetics effects. As can be seen on the left of Figure 5, removing the  $(001)$  face leads to a more satisfactory morphology.

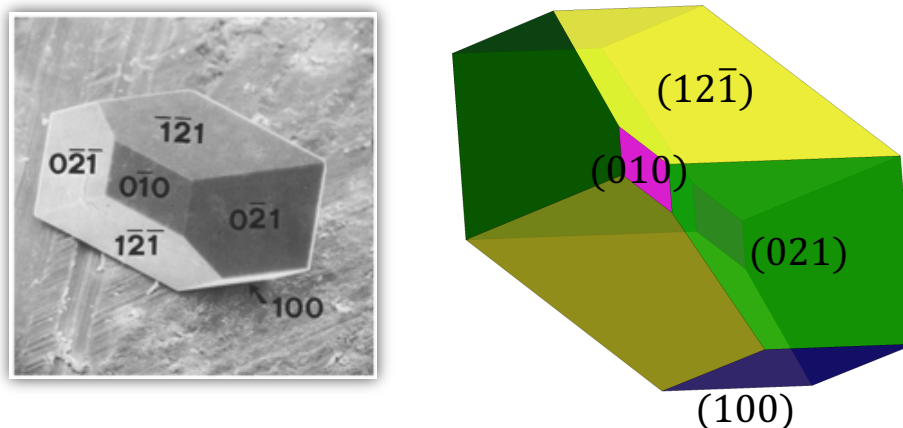


**Figure 5 : a) Morphology prediction of COM crystal with the water molecules, without (001) face. b) same but with a surface energy correction of  $-0.1 \text{ J.m}^{-2}$  on (010) face. SEM image is from Chen *et al.*<sup>19</sup>.**

Still, when compared to experimental data (Table 1a and b), (010) surface remained under expressed. As shown on Figure 5b, slightly decreasing the (010) surface energy ( $-0.1 \text{ J.m}^{-2}$ ) leads to a better prediction of the (010) surface exposure. Zucker *et al.* software<sup>59</sup> was not able to go further in the increase of the (010) face area.

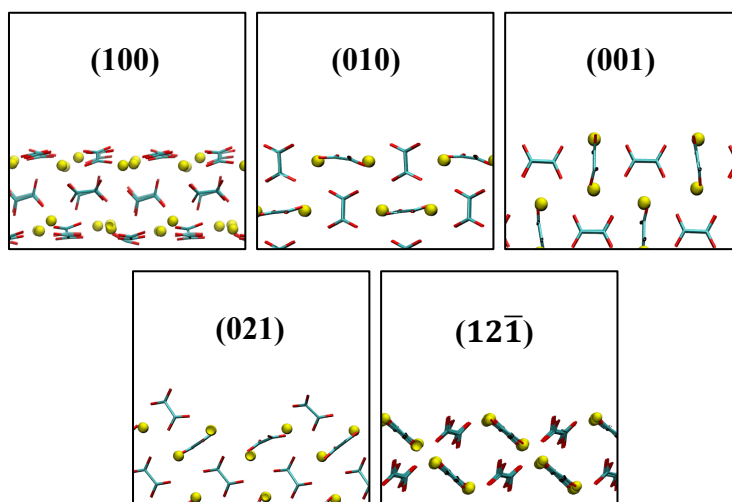
Furthermore, hydrated surface energies  $\gamma_{H_2O}$  of the  $(12\bar{1})$ , (010) and (021) faces are very close ( $0.33$ ,  $0.39$  and  $0.39 \text{ J.m}^{-2}$ , respectively). Expression of these faces in crystals are therefore very sensitive to external factors (temperature, additives, solvent, ...), and is in our error window. Thereby, as can be seen in Figure 6, a slight decrease of (010) and (021) surface energies ( $-0.03 \text{ J.m}^{-2}$ ) causes the emergence of (021) surface on the prediction.





**Figure 6 : Morphology prediction of COM crystal with water molecules, without (001) face and with energy corrections:  $-0.03 \text{ J.m}^{-2}$  for (021) and (010) surface energies. SEM image from Franchini-Angela et al.<sup>17,18</sup>**

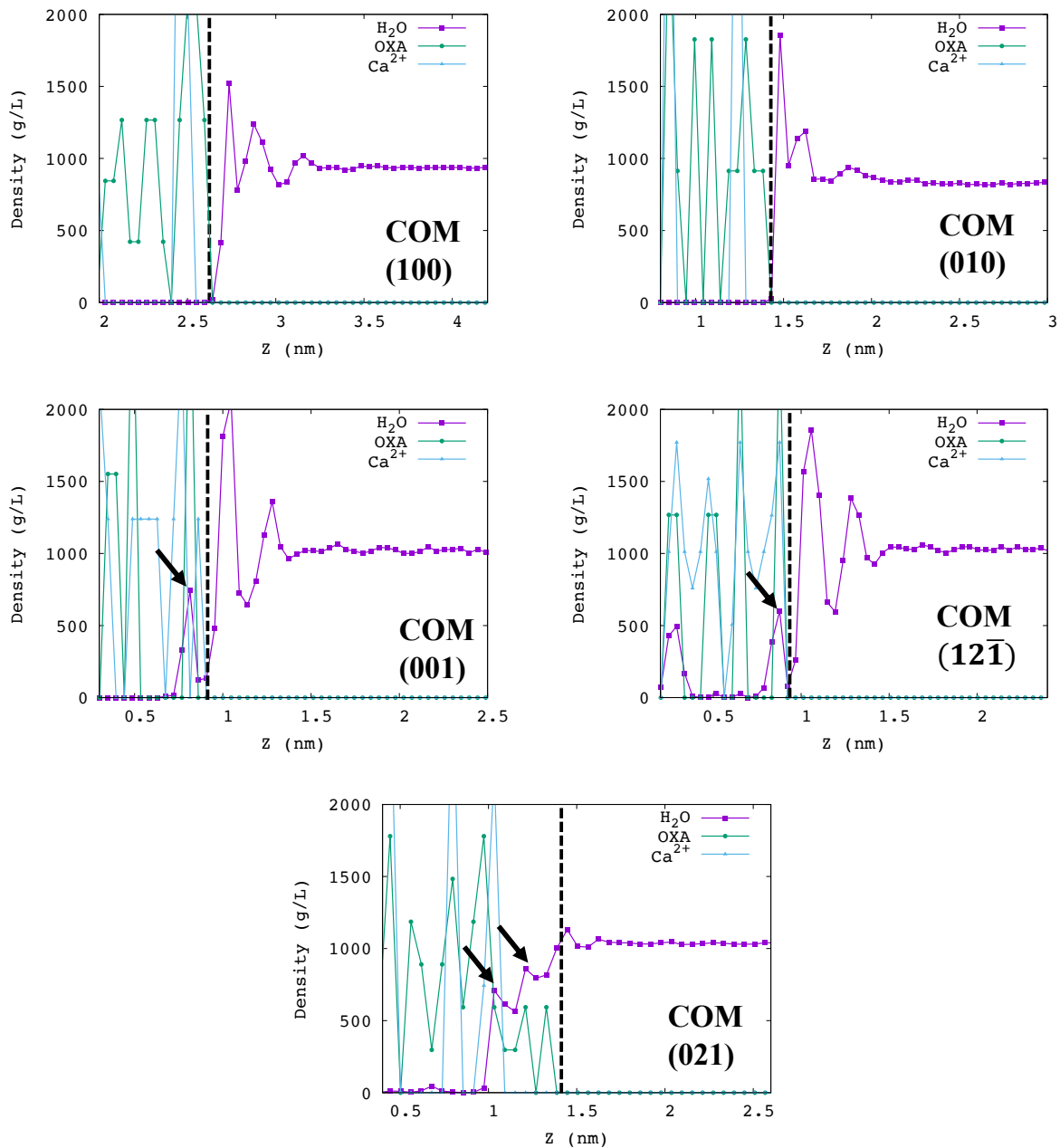
This morphology is in very good agreement with those observed in Franchini-Angela *et al.* syntheses (see Table 1c) and with one of the mineral crystal in Figure 1, which validates here again our calculation level. All relevant COM face are presented in Figure 7.



**Figure 7: predominant COM faces.**

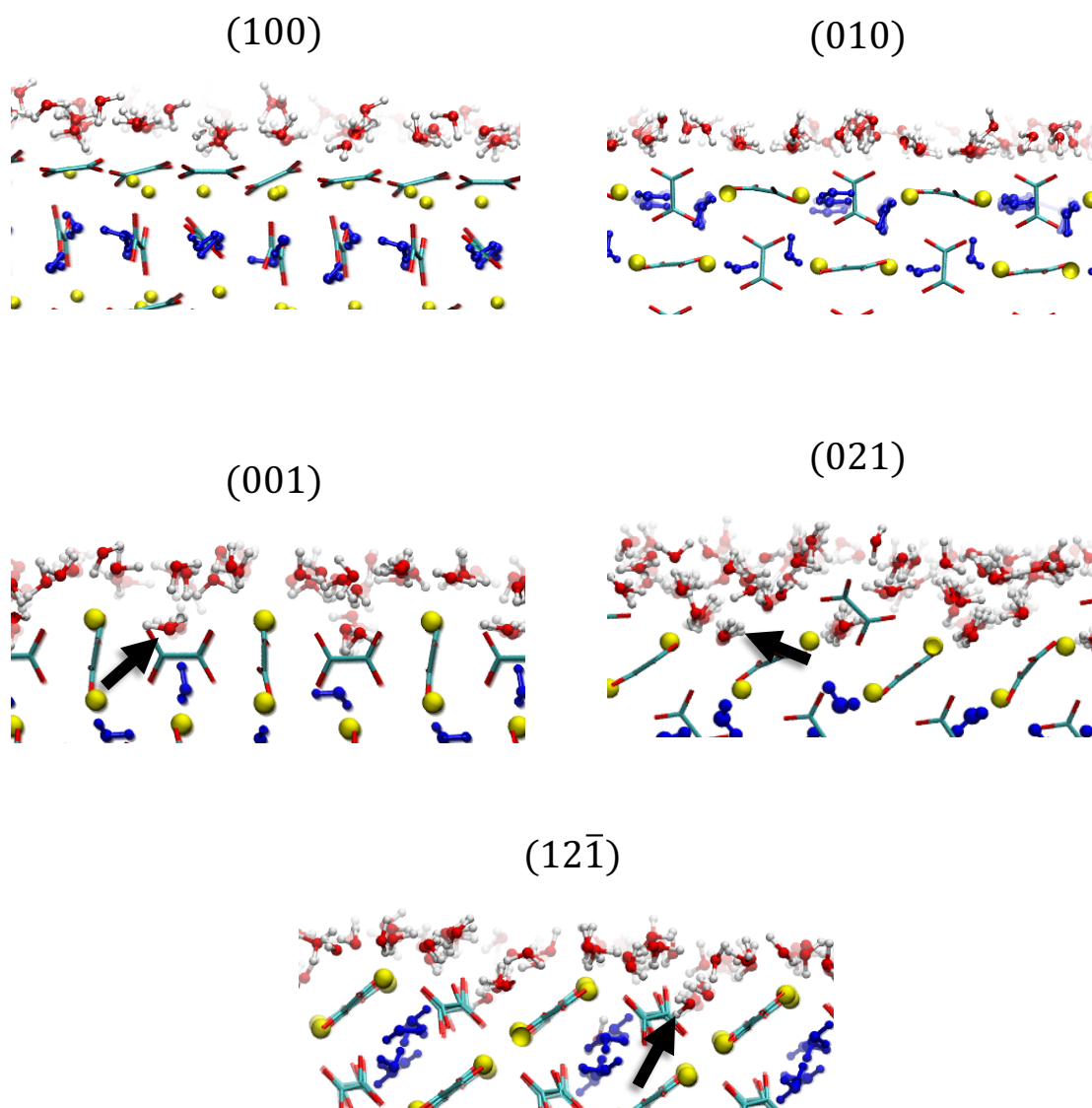
Interactions between COM surfaces and water were investigated by means of force field molecular dynamics simulations. Figure 8 represents the mean density profile of three entities: calcium, oxalate ions and water molecules. The density of the water molecules coming from the COM structure (structural water) is not plotted on these profiles. As said before, COM surface energies have a different behavior with water which has an impact on the crystal shape: the (021), (121) and (001) faces can be defined as “hydrophilic” ( $\gamma_{hydrated} - \gamma_{vacuum} < 0$ )

while the (010) and (100) faces can be defined as “hydrophobic” ( $\gamma_{hydrated} - \gamma_{vacuum} > 0$ ). The density profiles show also a difference between both families of surfaces. Indeed, in “hydrophilic” faces, water molecules are able to penetrate the surfaces in order to stabilize some under-coordinated ions. This effect leads to the appearance of water density peaks at comparable level than those of  $\text{Ca}^{2+}$  and oxalate ions density peaks. These particular  $\text{H}_2\text{O}$  peaks are labeled with black arrows in Figure 8 and are only present in the so-called “hydrophilic” faces.



**Figure 8: Density profile of water for COM (100), (010), (001) and (021) faces. Black arrows corresponds to water molecules inside COM structure. Z is the coordinate orthogonal to COM surface.**

The corresponding solvation layers are shown in Figure 9. For the (001), (021) and  $(12\bar{1})$  faces, water molecules labeled with black arrows are taking place inside the COM structure. Because of the model construction, some pseudo crystallographic sites where structural water was positioned, has been left empty at the surface. That is why, the water molecules from our solvation model are able to fill those these “holes” and stabilize the surface. For the (100) and the (010) faces there are no ‘holes’ inside the structure explaining why these faces are not stabilized by the use of a solvation model. By consequence, in order to correctly describe COM surfaces and to calculate accurate surface energies a solvation model is needed.



**Figure 9: water molecules (red and white) interacting with COM (100), (010), (001), (021) and  $(12\bar{1})$  faces (blue : crystallographic water, yellow : Ca<sup>2+</sup>, green and red : oxalate). Blacks arrows corresponds to water molecules “inside” a given face.**

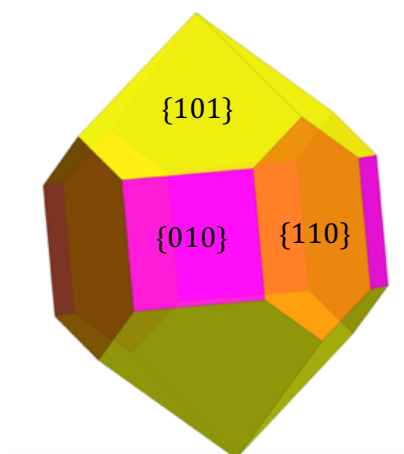
### *The COD case*

COD surface energies are reported in Table 4. The (101), (110), (010) and (011) faces are the most stable in vacuum. Their surface energies are lower than all other calculated surfaces. These results lead to a rather correct morphology prediction (see Figure 10), in overall agreement with theoretical morphology obtained with the attachment energies method<sup>31</sup>. However, in contrast with experimental observations (see Table 2) these four surfaces are equivalent.

face	$\gamma_{vacuum}$	$\gamma_{H_2O}$
(101)	0.29	0.24
(110)	0.29	0.28
(010)	0.29	0.35
(021)	0.46	0.39
(001)	0.47	0.41
(012)	0.42	0.42
(111)	0.59	0.42
(120)	0.41	0.42
(112)	0.46	0.44
(011)	0.30	0.46
(121)	0.48	0.46

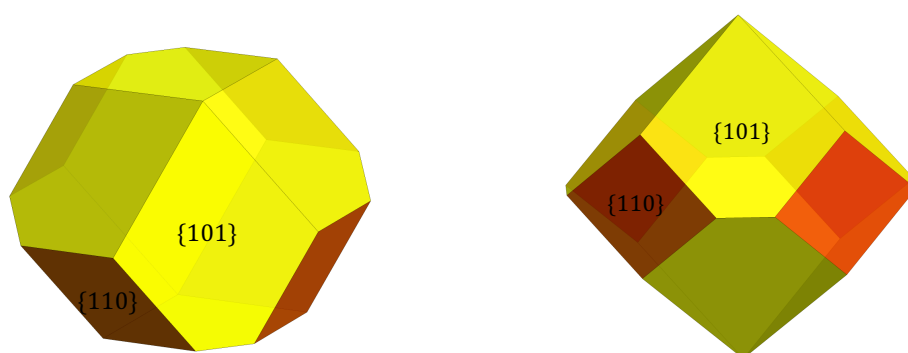
**Table 4 : Surface energies of COD in vacuum and with a bilayer of water molecules.  
Surface energies in  $J.m^{-2}$ .**

Adding a bilayer of water molecules allows to distinguish based on surface energies. Indeed, solvation leads to a slight stabilization of the (101) face and a slight destabilization of the (010) face causing the disappearance of the (010) face in the morphology prediction (See Figure 11).



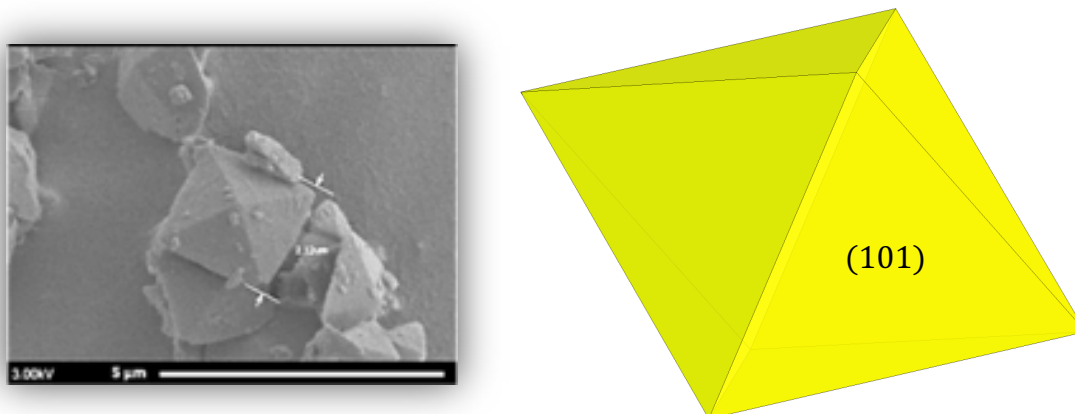
**Figure 10: Morphology prediction of COD in vacuum and based on  $\gamma_{vacuum}$  values in Table 4.**

This result is coherent with COD syntheses observations where the (010) face is observed only in presence of specific additives (see Table 2b).



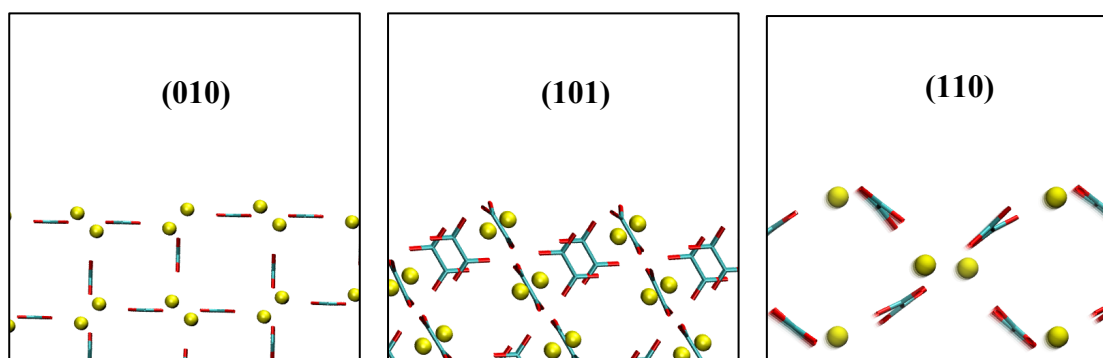
**Figure 11 : Morphology prediction of hydrated COD and based on  $\gamma_{H_2O}$  values in Table 4.**

However, this prediction is still not in agreement with experimental data (see Table 2a and Table 2c) where the (101) face is the only exposed face. Adding an energy correction of  $+0.2 \text{ J.m}^{-2}$  to (110) face is necessary in order to obtain the morphology described in Figure 12, and corresponding to the experimental observations. This energy is non-negligible suggesting that COD crystal may do not adopt the most stable morphology. The (110) face could be absent because of kinetics.



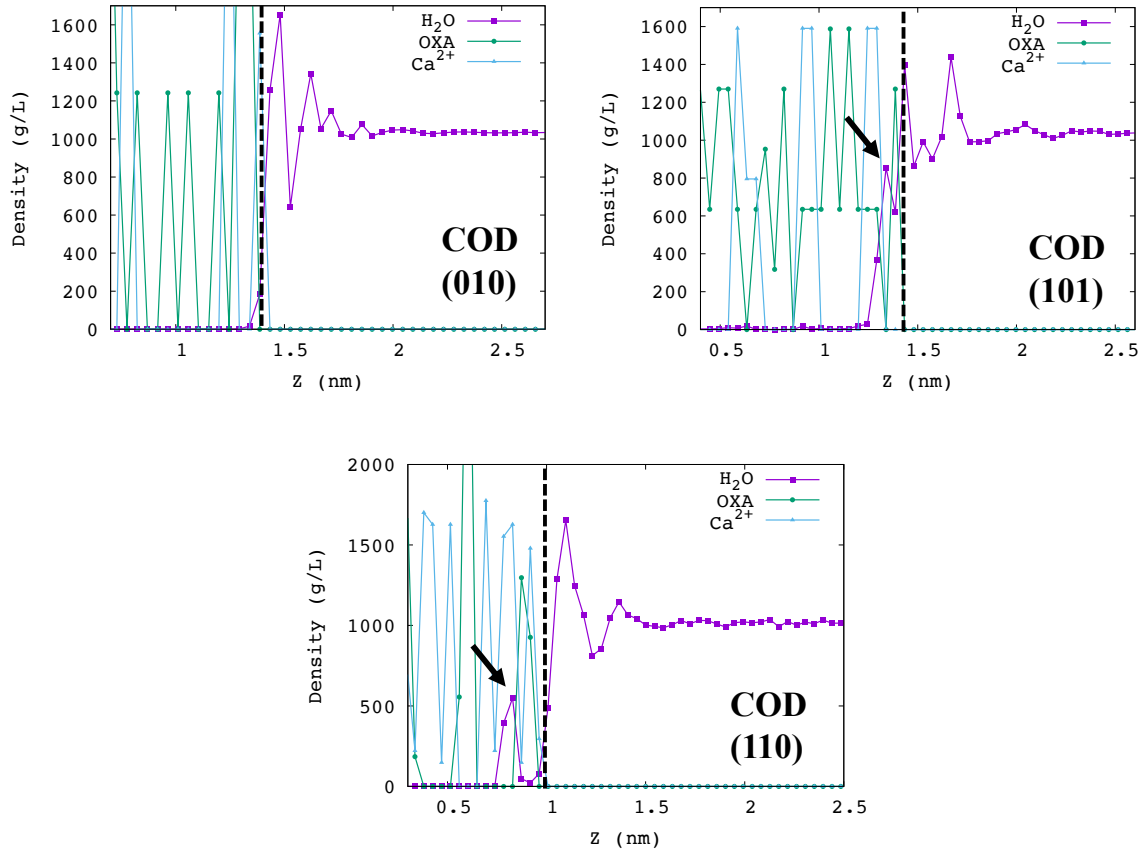
**Figure 12: Morphology prediction of hydrated COD with a surface energy correction of  $+0.2 \text{ J.m}^{-2}$  on (110) face. SEM image is from Leroy<sup>26</sup>.**

The three relevant COD faces are presented in Figure 13.



**Figure 13: predominant COD faces.**

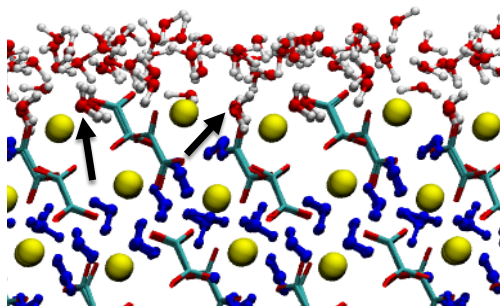
Similar to the “hydrophilic” COM surfaces, water molecules penetrated inside the (101) face structure (black arrow in Figure 14 and Figure 15) to stabilize the surface. This solvation effect is not observed on the (010) face which may explain why the face is not observed in syntheses and needs specific conditions to be expressed (*ie* additional molecules).



**Figure 14: Density profile of water for COD 010,101 and 110 surfaces. Blacks arrows corresponds to water molecules inside COM structure.**

A density peak was also found inside the structure of the (110) face (*cf* Figure 14) but no particular stabilization was observed.

**(101)**



**Figure 15: water molecules (red and white) filling empty water crystallographic sites at (101) interface (blue: crystallographic water, yellow:  $\text{Ca}^{2+}$ , green and red: oxalate). Blacks arrows corresponds to water molecules inside COM structure.**

## **Conclusion**

The scope of this contribution is to develop a theoretical approach being able to predict growth inhibition potential of targeted molecules' ability to inhibit calcium oxalate growth, in a drug design perspective. Reliable models for COM and COD surfaces were thus needed. In this study, COM and COD exposed faces from experimental observations were all predicted by means of *ab initio* total crystal surface energy minimization. A solvation model was crucial in order to obtain accurate surface energies. All morphologies were investigated and discussed. For most structures a very good agreement with experiment was obtained. Other, enables us to discuss the competition between thermodynamics and kinetics in the formation mechanism. In particular, the COM (001) face and the COD (110) face were predicted by our calculation while they are not observed experimentally. Removing those faces from the prediction was then essential to obtain a suitable morphology prediction. We therefore deduce that COM and COD crystals may not adopt the most thermodynamically stable morphology and should be subject to kinetic effects.

This study leading the way to the investigation of the COM/COD growth inhibition by specific molecules. In particular, the effect of a catechin content of green tea, will be investigated in the next chapter.



## **Acknowledgements**

This work was performed using HPC resources from GENCI- [CCRT/CINES/IDRIS] (Grant 2016-[x2016082022]) and the CCRE of Université Pierre et Marie Curie.

## **References**

1. Franceschi, V. R. & Nakata, P. A. Calcium Oxalates in Plants: Formation and Function. *Annu. Rev. Plant Biol.* **56**, 41–71 (2005).
2. Franceschi, V. R. & Horner, H. T. Calcium oxalate crystals in plants. *Bot. Rev.* **46**, 361–427 (1980).
3. Moe, O. W. Kidney stones: pathophysiology and medical management. *Lancet* **367**, 333–344 (2006).
4. Liu, Y. *et al.* Differentiating calcium oxalate and hydroxyapatite stones in vivo using dual-energy CT and urine supersaturation and pH values. *Acad. Radiol.* **20**, 1521–5 (2013).
5. Hajir, M., Graf, R. & Tremel, W. Stable amorphous calcium oxalate: synthesis and potential intermediate in biomineralization. *Chem. Commun.* **50**, 6534–6536 (2014).
6. Walton, R. C., Kavanagh, J. P., Heywood, B. R. & Rao, P. N. Calcium oxalates grown in human urine under different batch conditions. *J. Cryst. Growth* **284**, 517–529 (2005).
7. Goldschmidt, V. Atlas der Krystallformen (1913 edition) | Open Library. *Heidelberg : Winter*
8. Arnott, H. J., Pautard, F. G. E. & Steinfink, H. Structure of Calcium Oxalate Monohydrate. *Nature* **208**, 1197–1198 (1965).
9. Tazzoli, V. & Domeneghetti, C. The crystal structures of whewellite and weddellite: re-examination and comparison. *Am. Mineral.* **65**, 27–334 (1980).
10. Lieske, J. C., Toback, F. G. & Deganello, S. Sialic acid-containing glycoproteins on renal cells determine nucleation of calcium oxalate dihydrate crystals. *Kidney Int.* **60**, 1784–1791 (2001).
11. Deganello, S. & IUCr. The structure of whewellite, CaC<sub>2</sub>O<sub>4</sub>. H<sub>2</sub>O at 328 K. *Acta Crystallogr. Sect. B Struct. Crystallogr. Cryst. Chem.* **37**, 826–829 (1981).
12. Daudon, M. *et al.* Examination of whewellite kidney stones by scanning electron microscopy and powder neutron diffraction techniques. *J. Appl. Crystallogr.* **42**, 109–

- 115 (2009).
13. Sterling, C. & IUCr. Crystal structure analysis of weddellite,  $\text{CaC}_2\text{O}_4 \cdot (2+x)\text{H}_2\text{O}$ . *Acta Crystallogr.* **18**, 917–921 (1965).
  14. Izatulina, A., Gurzhiy, V. & Frank-Kamenetskaya, O. Weddellite from renal stones: Structure refinement and dependence of crystal chemical features on  $\text{H}_2\text{O}$  content. *Am. Mineral.* **99**, 2–7 (2014).
  15. Bouropoulos, N., Weiner, S. & Addadi, L. Calcium Oxalate Crystals in Tomato and Tobacco Plants: Morphology and in Vitro Interactions of Crystal-Associated Macromolecules. *Chem. Eur. J.* **7**, 1881–1888 (2001).
  16. Millan, A. Crystal Growth Shape of Whewellite Polymorphs: Influence of Structure Distortions on Crystal Shape. *Cryst. Growth Des.* **1**, 245–254 (2001).
  17. Aquilano, D. & Franchini-Angela, M. Twin laws of whewellite,  $\text{CaC}_2\text{O}_4 \cdot \text{H}_2\text{O}$ . A structural and growth approach. *Phys. Chem. Miner.* **7**, 124–129 (1981).
  18. Franchini-Angela, M. & Aquilano, D. Theoretical growth morphology of whewellite  $\text{CaC}_2\text{O}_4 \cdot \text{H}_2\text{O}$ . *Phys. Chem. Miner.* **10**, 114–120 (1984).
  19. Chen, Z., Wang, C., Zhou, H., Sang, L. & Li, X. Modulation of calcium oxalate crystallization by commonly consumed green tea. *Cryst. Eng. Comm* **12**, 845–852 (2010).
  20. Sandersius, S. & Rez, A. P. Morphology of crystals in calcium oxalate monohydrate kidney stones. doi:10.1007/s00240-007-0115-3
  21. Haussühl, S. P. Hartman (Ed.): Crystal Growth: An Introduction. North-Holland series in crystal growth Vol. 1. North-Holland Publishing Company, Amsterdam-London 1973. 531 S., Preis: \$ 21.10. *Berichte der Bunsengesellschaft für Phys. Chemie* **78**, 1276–1276 (1974).
  22. Docherty, R., Clydesdale, G., Roberts, K. J. & Bennema, P. Application of Bravais-Friedel-Donnay-Harker, attachment energy and Ising models to predicting and understanding the morphology of molecular crystals. *J. Phys. D. Appl. Phys.* **24**, 89–99 (1991).
  23. Rijpkema, J. J. M., Knops, H. J. F., Bennema, P. & van der Eerden, J. P. Determination of the Ising critical temperature of F slices with an application to garnet. *J. Cryst. Growth* **61**, 295–306 (1983).
  24. Hartman, P. & Bennema, P. The attachment energy as a habit controlling factor: I. Theoretical considerations. *J. Cryst. Growth* **49**, 145–156 (1980).

25. Tommaso, D. Di, Hernández, S. E. R., Du, Z. & Leeuw, N. H. de. Density functional theory and interatomic potential study of structural, mechanical and surface properties of calcium oxalate materials. *RSC Adv.* **2**, 4664 (2012).
26. Leroy, C. Oxalates de calcium et hydroxyapatite : des matériaux synthétiques et naturels étudiés par techniques RMN et DNP. (Université P. et M. Curie, 2016).
27. Giordani, P., Modenesi, P. & Tretiach, M. Determinant factors for the formation of the calcium oxalate minerals, weddellite and whewellite, on the surface of foliose lichens. *Lichenologist* **35**, 255–270 (2003).
28. Zhang, D., Qi, L., Ma, J. & Cheng, H. Morphological Control of Calcium Oxalate Dihydrate by a Double-Hydrophilic Block Copolymer. (2002). doi:10.1021/cm010768y
29. Leroy, C. *et al.* Hyperoxaluria is related to whewellite and hypercalciuria to weddellite: What happens when crystalline conversion occurs? *Comptes Rendus Chim.* **19**, 1492–1503 (2016).
30. Wesson, J. A. & Ward, M. D. Pathological Biomineralization of Kidney Stones. *Elements* **3**, 415–421 (2007).
31. Heijnen, W. M. M. & Van Duijneveldt, F. B. The theoretical growth morphology of calcium oxalate dihydrate. *J. Cryst. Growth* **67**, 324–336 (1984).
32. Qiu, S. R. *et al.* Molecular modulation of calcium oxalate crystallization by osteopontin and citrate. *Proc. Natl. Acad. Sci.* **101**, 1811–1815 (2004).
33. Chen, W.-C., Lin, H.-S., Chen, H.-Y., Shih, C.-H. & Li, C.-W. Effects of Tamm-Horsfall Protein and Albumin on Calcium Oxalate Crystallization and Importance of Sialic Acids. *Mol. Urol.* **5**, 1–5 (2001).
34. Ruiz-Agudo, E. *et al.* A non-classical view on calcium oxalate precipitation and the role of citrate. *Nat. Commun.* **8**, (2017).
35. Chen, Z., Wang, C., Zhou, H., Sang, L. & Li, X. Modulation of calcium oxalate crystallization by commonly consumed green tea. *CrystEngComm* **12**, 845 (2010).
36. Grases, F., Rodriguez, A. & Costa-Bauza, A. Efficacy of Mixtures of Magnesium, Citrate and Phytate as Calcium Oxalate Crystallization Inhibitors in Urine. *J. Urol.* **194**, 812–819 (2015).
37. Riley, J. M., Kim, H., Averch, T. D. & Kim, H. J. Effect of magnesium on calcium and oxalate ion binding. *J. Endourol.* **27**, 1487–92 (2013).
38. Alamani, B. G. & Rimer, J. D. Molecular modifiers of kidney stones. *Curr. Opin. Nephrol. Hypertens.* **26**, 256–265 (2017).

39. Kaleta, B. The role of osteopontin in kidney diseases. *Inflamm. Res.* **68**, 93–102 (2019).
40. Wilhelm Riemenschneider, M. T. No Title. in *Ullmann's Encyclopedia of Industrial Chemistry Oxalic Acid* (2004).
41. Ihli, J. *et al.* Precipitation of Amorphous Calcium Oxalate in Aqueous Solution. (2015). doi:10.1021/acs.chemmater.5b01642
42. VandeVondele, J. *et al.* Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Comput. Phys. Commun.* **167**, 103–128 (2005).
43. Lippert, G., Hutter, J. & Parrinello, M. The Gaussian and augmented-plane-wave density functional method for ab initio molecular dynamics simulations. *Theor. Chem. Accounts Theory, Comput. Model. (Theoretica Chim. Acta)* **103**, 124–140 (1999).
44. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
45. Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **38**, 3098–3100 (1988).
46. Klime, J., Bowler, D. R. & Michaelides, A. Van der Waals density functionals applied to solids. *Phys. Rev. B - Condens. Matter Mater. Phys.* **83**, 1–13 (2011).
47. Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate *ab initio* parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **132**, 154104 (2010).
48. Goedecker, S., Teter, M. & Hutter, J. Separable dual-space Gaussian pseudopotentials. *Phys. Rev. B* **54**, 1703–1710 (1996).
49. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
50. Reif, M. M., Hünenberger, P. H. & Oostenbrink, C. New Interaction Parameters for Charged Amino Acid Side Chains in the GROMOS Force Field. *J. Chem. Theory Comput.* **8**, 3705–3723 (2012).
51. Berendsen, H. J. C., Grigera, J. R. & Straatsma, T. P. The Missing Term in Effective Pair Potentials I. *J. Phys. Chem* **91**, 6269–6271 (1987).
52. Hess, B. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *J. Chem. Theory Comput.* **3**, 116–122 (2008).
53. Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **52**, 255–268 (1984).

54. Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **31**, 1695–1697 (1985).
55. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
56. Essmann, U. *et al.* A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577 (1995).
57. Gibbs, J. W. A. du texte. The collected works / of J. W. Gibbs,... (1928).
58. Wulff, G. Zur Frage der Geschwindigkeit des Wachstums und der Auflösung der Krystallflagen. *Zeitschrift für Kryst. und Mineral.* **34 (5/6)**, 449–530 (1901).
59. Zucker, R. V., Chatain, D., Dahmen, U., Hagège, S. & Carter, W. C. New software tools for the calculation and display of isolated and attached interfacial-energy minimizing particle shapes. *J. Mater. Sci.* **47**, 8290–8302 (2012).

## - Chapter 3 -

# Effect of Catechin on the Calcium Oxalate Crystal Growth: A Molecular Dynamics Study

*Théau Debroise<sup>1</sup>, César Leroy<sup>1</sup>, Christian Bonhomme<sup>1</sup>, Marialore Sulpizi<sup>2</sup> and Frederik Tielens<sup>3\*</sup>*

<sup>1</sup> Sorbonne Université, CNRS, Laboratoire Chimie de la Matière Condensée de Paris, LCMCP, 4 Place Jussieu, F-75005 Paris, France

<sup>2</sup> Johannes Gutenberg University Mainz, Staudinger Weg 7 55099 Mainz, Germany

<sup>3</sup> General Chemistry (ALGC), Vrije Universiteit Brussel (Free University Brussels-VUB), Pleinlaan 2, 1050 Brussel, Belgium

\*Author to whom correspondence should be addressed: frederik.tielens@vub.be

<b>Abstract .....</b>	<b>101</b>
<b>Introduction.....</b>	<b>102</b>
<b>Methods .....</b>	<b>105</b>
<b>Results and Discussion .....</b>	<b>110</b>
<b>Conclusion .....</b>	<b>119</b>
<b>Acknowledgements .....</b>	<b>120</b>
<b>References.....</b>	<b>120</b>

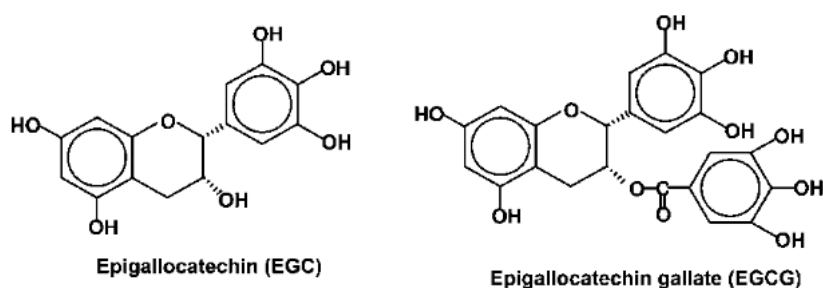
## **Abstract**

Green tea is known to have a positive impact on kidney stone disease. Its catechins constituents are suspected to strongly interact with calcium oxalate (CaOx) minerals. On one hand, it has been observed *in vitro* that green tea promotes calcium oxalate dihydrate (COD) against monohydrated phase (COM). On the other hand, structural modifications have been observed by TEM on COD surfaces synthesized in a green tea medium. Interactions between the main catechin of green tea, epigallocatechin gallate (EGCG) and calcium oxalate surfaces in water was investigated by the mean of Molecular Dynamics (MD). We present here a method to quantify the adsorption/desorption free energies and we show that sophisticated algorithms are needed to get free energy convergence. Replica Exchange Molecular Dynamics (REMD) coupled with metadynamics was finally the most adapted method to get a full picture of EGCG adsorption/desorption on calcium oxalate surfaces. We demonstrated that EGCG interact preferentially with COM (010) and (100) and COD (010) surfaces.

## Introduction

Tea is one of the most popular drinks in the world thanks to its flavors but also to its therapeutic value. Many epidemiological studies already identified effect of green tea on cancer prevention<sup>1</sup>. The main evoked reason is related to the presence of a family of molecules called catechin. Indeed, catechins can reach up to 30% of the mass of the tea leaf<sup>2</sup> and have probably antioxidative properties.

Their phenol groups are suspected to act on transcription factors and enzymatic activity<sup>3</sup> to reduce cell's oxidative stress. These results generated a certain interest for this family of molecules in the literature. In particular, EGCG ((-)- Epigallocatechin-3-gallate) is referenced as the major catechin in green tea (between 50 and 80% of catechins)<sup>4</sup> (See Figure 1).



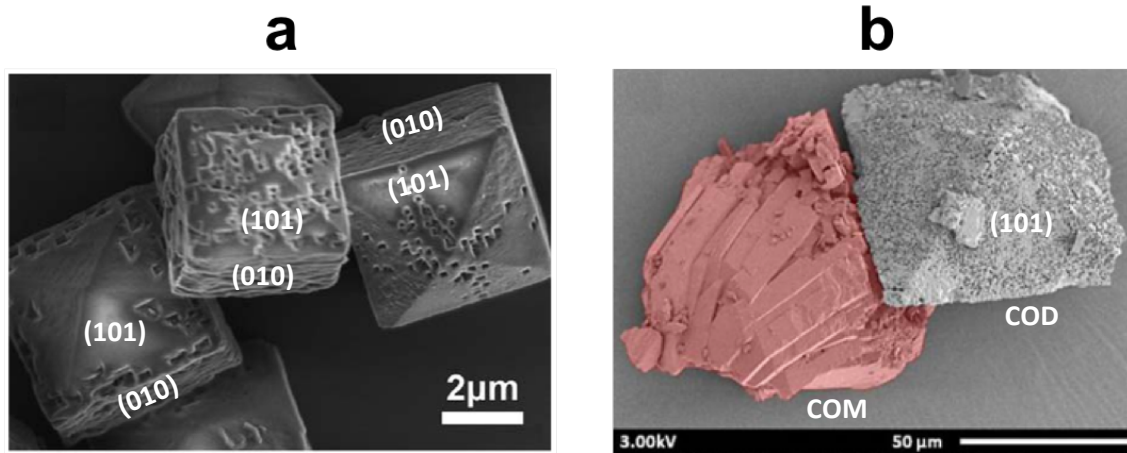
*Figure 1: Epigallocatechin (EGC) and (-)-Epigallocatechin-3-gallate (EGCG)*

In addition, clinical studies have demonstrated the positive impact of green tea on the formation of kidney stones<sup>5,6</sup>. The reason remains unclear. On the one hand, the antioxidative effect of catechins may have an impact on biological mechanisms that in turn may influence the regulation of calcification in the urinary system<sup>5,6</sup>. On the other hand, catechins are known to have a direct effect on calcium oxalates crystal growth *in vitro*<sup>7</sup>.

In previous studies, influence of green tea was evaluated in calcium oxalate syntheses<sup>7,8</sup> and in kidney stones (surgically retrieved)<sup>9</sup>. According to Leroy<sup>8</sup>, the yield of syntheses is not affected by the presence of green tea. In addition, the clinical study of Rode *et al.*<sup>9</sup> does not show a decrease in the amount of kidney stones in patients drinking green tea. However, it has been demonstrated that green tea is able to orient selectively the crystallization: the kinetic product COD is favored at the expense of COM<sup>7,8</sup>. This effect also tends to be observed in kidney stones. Indeed, tea drinking patients' kidney stones present a larger amount of COD when compared to non-drinking ones<sup>9</sup>. COD crystals are assumed to have lower adhesion than COM to kidney's biological tissues and would then be more effectively removed than COM in the urinary flow,



giving catechins a major therapeutic benefit<sup>10</sup>. Furthermore, profound changes are observed *in vitro* on SEM images of the obtained COD crystals<sup>7-9</sup> (see Figure 2).



**Figure 2: (a) SEM image of COD crystal in the presence of green tea. Adapted from Chen *et al.*<sup>7</sup> (b) SEM image of COM (red) and COD (grey) crystals synthesized in agar gel without green tea and then immersed in a green tea infusion. Adapted from Leroy.<sup>8</sup>**

In presence of green tea, cavities ranging from 200 to 500 nm emerge on the COD (010) surface and an elongation of the (100) face is described by Chen *et al.*<sup>11</sup>. Leroy<sup>8</sup> showed that COM is much less altered than COD when immersed in green tea. Indeed, cavities are clearly observed on COD crystals whereas COM agglomerates (in red in Figure 2b) are much less affected. Finally, it is known that COD (010) face exposition requires an additional stabilization by interacting molecules<sup>12</sup>. In Chen *et al.*<sup>11</sup> synthesis, green tea cause the emergence of COD (010) face (Figure 2**Erreur ! Source du renvoi introuvable.**a).

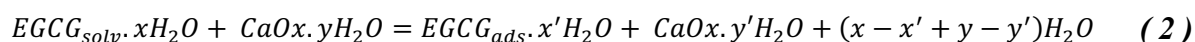
Impact of green tea is demonstrated experimentally but the role of EGCG is just an assumption proposed by various author. Consequently, our main goal is to find a method able to quantify EGCG adsorption/desorption on calcium oxalate surfaces to allow new insight on EGCG selectivity regarding to COM and COD surfaces.

In the canonical ensemble, free energy of adsorption ( $\Delta F_{\text{adsorption}}$ ) is the adequate thermodynamic variable to measure the feasibility of the adsorption process:

$$\Delta F_{\text{adsorption}} = F_{\text{adsorbed}} - F_{\text{solvated}} \quad (1)$$

Several contributions to  $\Delta F_{\text{adsorption}}$  can be inferred from chemical characteristics of EGCG (Figure 1), the various CaOx surfaces and water.

- EGCG could be involved in **short range interactions** between hydroxyl, COO and COC groups and Ca<sup>2+</sup> ions on COM/COD surfaces. Moreover, EGCG should be able to develop hydrogen bonds with oxalates ions by the means of the hydroxyl groups.
- **Long-range electrostatic interactions** may play a non-negligible role as well on the interaction between EGCG and the COM/COD surfaces.
- **Solvation effect:** adsorption of EGCG causes a decrease of the total solvent accessible surface area. Indeed, a new interface between EGCG and CaOx must involve the expulsion of some water molecules from the first layer of solvation of both EGCG and calcium oxalate surface as described in equation 2. This process has an entropic effect because water is more ordered in the solvation sphere than in the “bulk” water. It has an enthalpic effect as well because EGCG-water and surface-water interactions are replaced by EGCG-surface interactions.



- EGCG **steric constraints** should play a non-negligible role. When EGCG is adsorbed, it can adopt a particular geometry to optimize favorable interactions. This geometry can have an enthalpic cost because of constraint imposed to covalent bonds, bond angles and dihedrals. It can also have an entropic cost due to “freezing” EGCG in one specific geometry.
- Induced surface reconstruction/relaxation due to the interaction of EGCG with the surface: this last term is not investigated in present study.

MD is a valuable tool to take into account all these energetic terms. However, from our observations of standard MD, it is not sufficient to simulate the adsorption process reversibly and to explore all possible EGCG adsorbed configurations. As a consequence, we have developed an original methodology, based on Replica Exchange Molecular Dynamic (REMD) and metadynamics techniques which is able to quantify EGCG adsorption. First, we present a comparison of four different methods applied to EGCG adsorption on COM (010) surface. Then, we applied the best method to predict EGCG affinities for COM and COD predominant surfaces.

## **Methods**

### ***Surface models***

Three models were built for COM surfaces: (100), (010) and (12 $\bar{1}$ ) and two for COD: (100) and (101). This particular choice was made from the conclusions established in chapter 2. These surfaces correspond to predominant surfaces of COM and COD<sup>13,14</sup>. 3D structures of COM and COD were obtained by neutron diffraction<sup>15</sup> and X-ray diffraction studies<sup>16</sup>, respectively. The surfaces are then constructed by adding a vacuum space of 50 Å inside the bulk structure. Model sizes are chosen large enough to avoid a possible interaction between periodic images of EGCG. The constructed surfaces are thus composed of more than 6000 atoms. After water addition, each simulation box contains about 50000 atoms. EGCG final concentration is 5.10<sup>-3</sup> mol.L<sup>-1</sup>.

### ***Interatomic potential, force field***

Version 54a7 of GROMOS<sup>17</sup> force field implemented in GROMACS<sup>18,19</sup> software (2016.5 version) was chosen to describe oxalate and calcium ions, and EGCG. Force field parameters were taken from a previous force field study<sup>20</sup>.

The parameters of the EGCG are determined using the Automated force field Topology Builder (ATB) tool<sup>21</sup>. In this approach, force field description of the molecule is derived specifically for the GROMOS forcefield. It is based on quantum mechanical (QM) calculations and a knowledge-based approach to ensure the compatibility with other molecules described also by the GROMOS force field. We are aware that the interactions between catechin and calcium oxalate may require some force field adjustments. This was not performed due to the size and the complexity of catechin molecule involving heavy calculations. However, we believe that this description of the interactions will allow conclusions to be drawn about the selectivity of EGCG on different calcium oxalate surfaces.

Water is described by the TIP3P model<sup>22</sup>. All simulations were run in the canonical ensemble (NVT) with a Nosé-Hoover thermostat to ensure thermal stability<sup>23,24</sup>. Particle Mesh Ewald (PME) method was used to optimize calculation of long-range interactions<sup>25,26</sup> with a grid spacing of 0.12 nm and an interpolation of 4. A real space cutoff of 1 nm was used, and the Lennard-Jones interactions were truncated at the same distance. Calcium oxalate positions were fixed using the P-LINCS algorithm<sup>27</sup>.

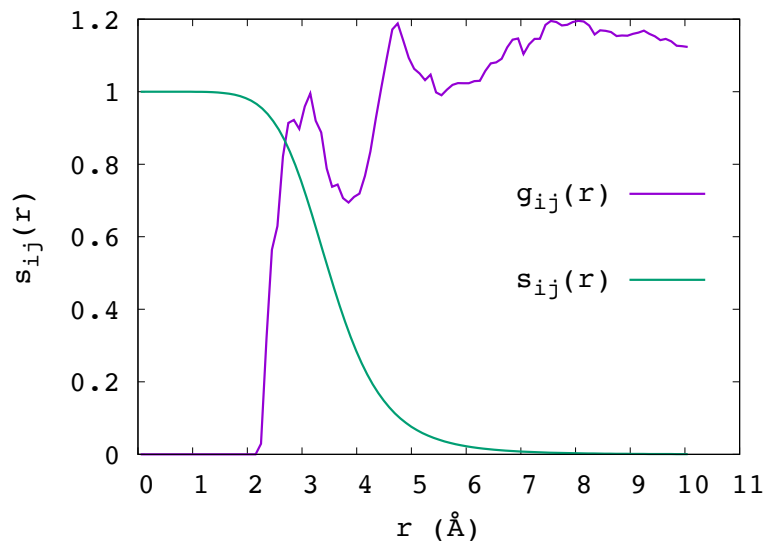
### Collective variables for adsorption

Two collective variables were chosen to describe the adsorption/desorption process of EGCG on CaOx surfaces:

- **D**: The **D**istance between the center of mass of EGCG and the surface of calcium oxalate
- **CN**: The **C**oordination **N**umber between EGCG oxygens atoms and all surface atoms of calcium oxalate<sup>28</sup>:

$$CN = \sum_i \sum_j s_{ij} \quad s_{ij} = \frac{1 - \left(\frac{r_{ij}}{r_0}\right)^n}{1 - \left(\frac{r_{ij}}{r_0}\right)^m} \quad (3)$$

Here,  $i$  represents the index of each EGCG *oxygen* atom and  $j$  represents each CaOx *surface atom*.  $s_{ij}$  is a switching function between 0 and 1 that describes the contact between atoms  $i$  and  $j$ .  $s_{ij} = 0$  corresponds to the absence of contact while  $s_{ij} = 1$  corresponds to a contact.  $r_0$  is the distance where the switching function is equal to 0.5.  $n$  and  $m$  exponents are two additional parameters.  $r_0$ ,  $n$  and  $m$  are fitted to take into account the first peak of the radial pair distribution function of atoms  $i$  and  $j$  as described in Figure 3. At the end,  $r_0 = 3.5 \text{ \AA}$ ,  $n = 7$  and  $m = 14$ .

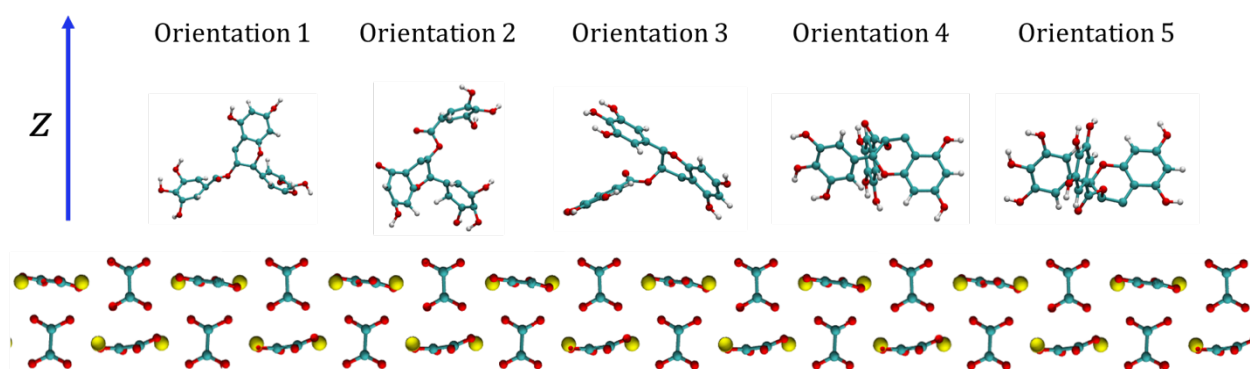


**Figure 3: Switching function  $s_{ij}(r)$  describing the contact between catechin's oxygen atoms ( $i$ ) and CaOx surface atoms ( $j$ ) and radial pair distribution function  $g_{ij}(r)$  of these two types of atoms**

### Free energy calculations

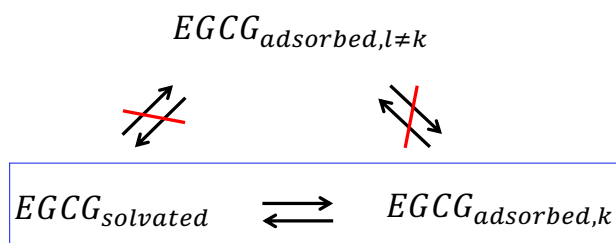
MD in the canonical ensemble is a technique of choice to describe adsorption processes. However, MD alone is not enough to reach ergodicity. Enhanced sampling is thus needed. *Four* different methods were implemented consequently and tested on COM (010) surface.

**In method 1**, the simplest description of the adsorption process is taken into account. First, EGCG molecule is simulated (NVT MD) in vacuum “above” the surfaces with different starting points corresponding to five different orientations (see Figure 4). During this first MD, EGCG is strongly attracted to CaOx surface and adopt quickly an adsorbed configuration.



**Figure 4:** Five different starting points of EGCG used to generate adsorption configurations above COM (010) surface. *Z* corresponds to the direction perpendicular to the (010) surface (*x* and *y* directions)

Water is then added to complete the vacuum of the simulation box. The system is again equilibrated in the NVT ensemble. During this step, EGCG is able to relax from its initial geometry. Finally, metadynamics<sup>29</sup> is implemented in order to reconstruct the free energy landscape of the adsorption process. The plumed v2.2-5 package was used<sup>30,31</sup>. An initial gaussian height of 2.5 kJ.mol<sup>-1</sup> was added every 0.4 ps, along both CVs **CN** and **D**. EGCG position was restrained in *x* and *y* directions with LINCS algorithm<sup>27</sup>. It follows that EGCG is only able to move perpendicular to the surface of interest (*Z* direction in Figure 4).

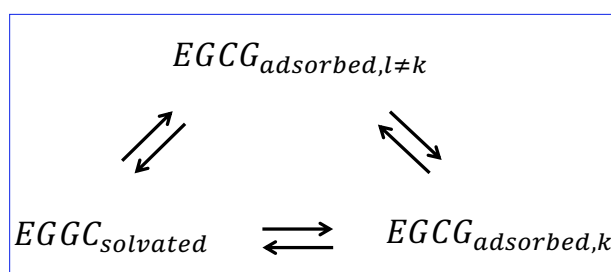


**Figure 5:** Adsorption process equation (method 1).

This constraint avoids EGCG to visit other adsorbed states on the surface (see Figure 5). Therefore, the free energy of a unique adsorption process is calculated.

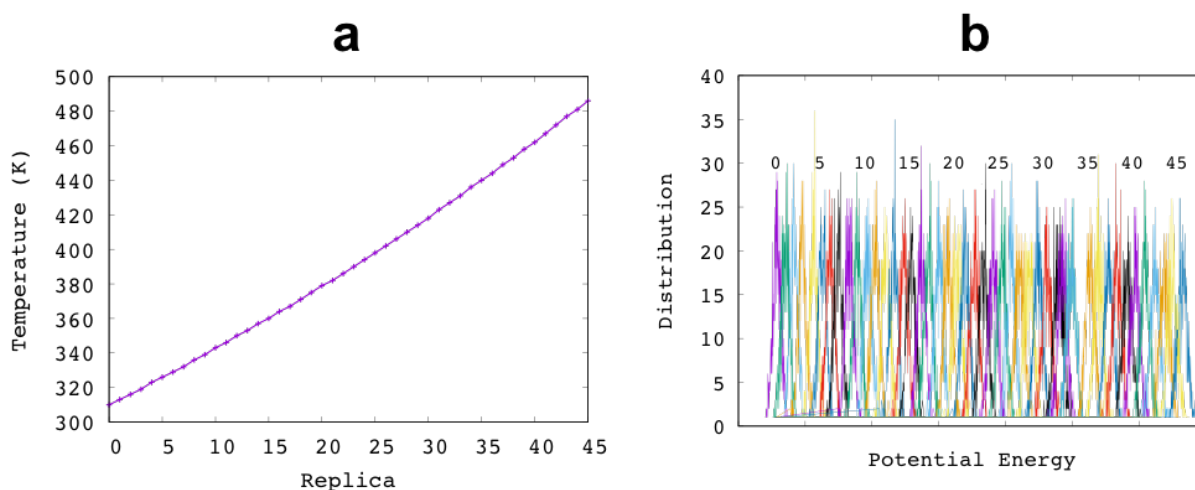
**In method 2**, EGCG position is not constraint and is able to visit any adsorbed configuration on the CaOx surface (see Figure 6).

REMD<sup>32</sup> was used to improve the sampling. In this technique, N molecular dynamics, named as *replicas*, are run in parallel, at different temperatures. Every x MD steps, adjacent replicas are allowed to exchange positions according to a probability depending on a metropolis criterion. Replicas with higher temperatures allow the system to escape local free energy wells and therefore should improve the ability of EGCG to visit all possible adsorbed configurations. REMD has been parameterized on COM (010) surface according to two criteria.



**Figure 6: Adsorption process equation (method 2)**

First, the highest replica's temperature must allow EGCG desorption. Such a process is observed above a temperature of 450K. Second, a recovery between replicas' potential energies is essential to observe a sufficient exchange rate between them.



**Figure 7: (a) Replicas' temperature distribution. (b) Potential energy distribution of the 46 replicas. One color corresponds to one replica.**

At the end, 46 replicas were chosen ranging from 310 to 486K as shown in Figure 7a. Potential energy recovery was verified by plotting the potential energy distribution of the 46 replicas during 46 NVT molecular dynamics of 1ns at the 46 temperatures. This distribution is presented in Figure 7b. One color corresponds to one replica. The higher the replicas' temperature, the higher the replica's potential energy. A good recovery is observed between replica's potential energies. Replica exchange was performed every 0.2 ps.

**In method 3**, only metadynamics is used to improve the sampling. In metadynamics, a gaussian bias is added to the potential energy, along the CVs every x steps. First, it improves sampling by discouraging the system from returning to previously visited configurations. Second, it allows the reconstruction of the free energy diagram. A gaussian height of 2.5 kJ.mol<sup>-1</sup> was added every 0.4 ps, along the two CVs, **CN** and **D**.

**In method 4**, metadynamics and REMD are applied in order to take full advantage of both approaches.

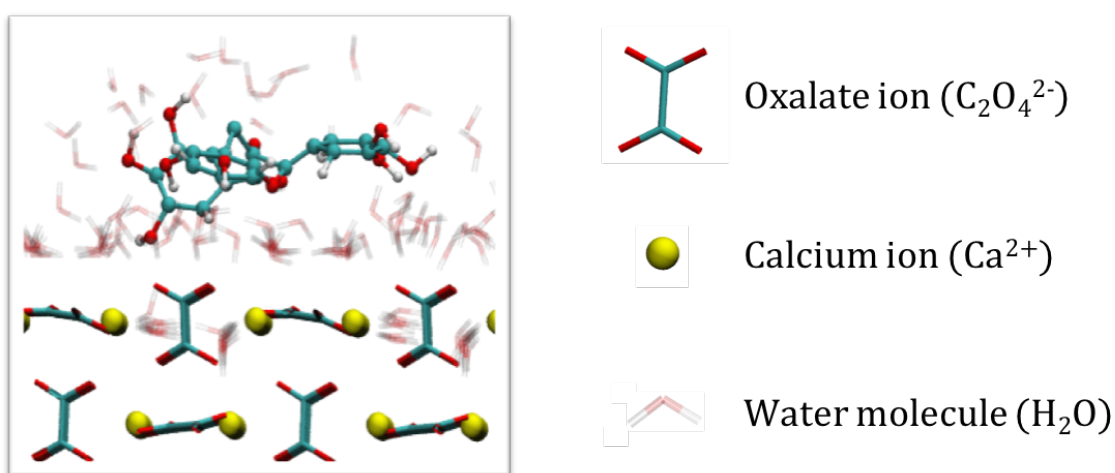
In **methods 2, 3 and 4**, the starting configuration (t = 0 ns) of MD corresponds to solvated EGCG, "far" from the given surface.

In all methods the CaOw atoms were not allowed to relax in order to focus on the adsorption sites and interaction energies. VMD<sup>33</sup> software was used in order to make the MD snapshots.

## Results and Discussion

### *Method 1*

As mentioned above, five EGCG orientations were tested above COM (010) surface leading to five different adsorbed configurations. From five metadynamics, no stable configuration was found for the adsorption of EGCG on COM (010) surface. As shown in Figure 8, water molecules are present inside the COM structure. These water molecules prevent EGCG to access the adsorbed state that was obtained without water. Consequently, **method 1** is not able to find correct adsorption configurations in the presence of water.

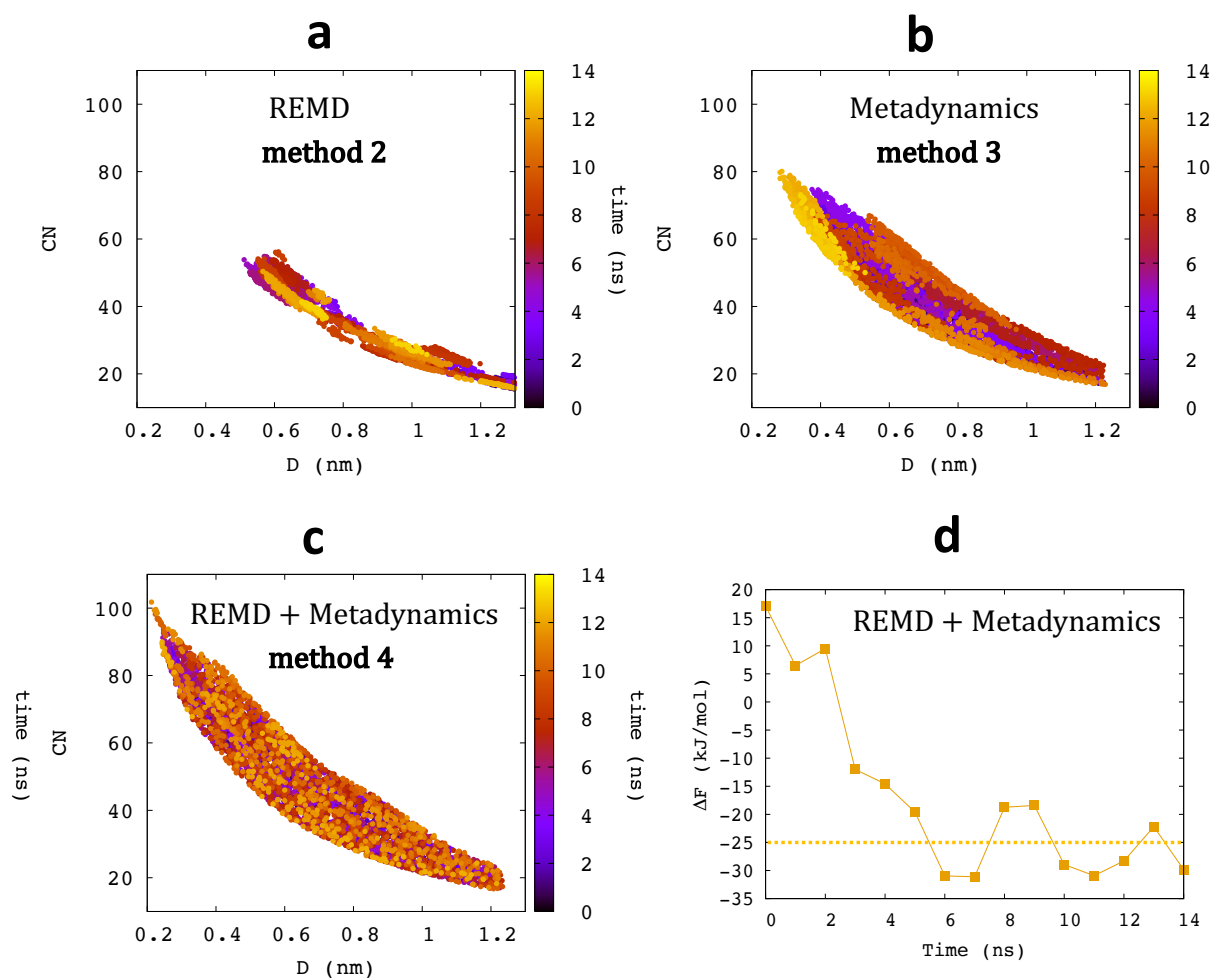


*Figure 8: EGCG above COM (010) surface CN = 52; D = 0.53 Yellow:  $Ca^{2+}$ , Red: O, Blue: C, White: H*

### *Method 2, 3 and 4*

REMD and metadynamics were thus used with the intention of increasing adsorbed configurations sampling. The three methods were compared in Figure 9 where collective variables **D** and **CN** were plotted with respect to time (color scale). As can be seen in Figure 9a, REMD generates configurations with a maximum **CN**  $\approx 55$ . From the trajectory observations, a **CN**  $\approx 55$  does not correspond to an adsorbed configuration but rather to EGCG interacting with the first layer of solvation of COM (010) surface. Contrary to REMD, some adsorbed configurations were found on COM (010) surface with metadynamics (Figure 9b, where configurations with **CN** between 60 and 80 are explored). We therefore concluded that metadynamics allow a better sampling compared to REMD alone.





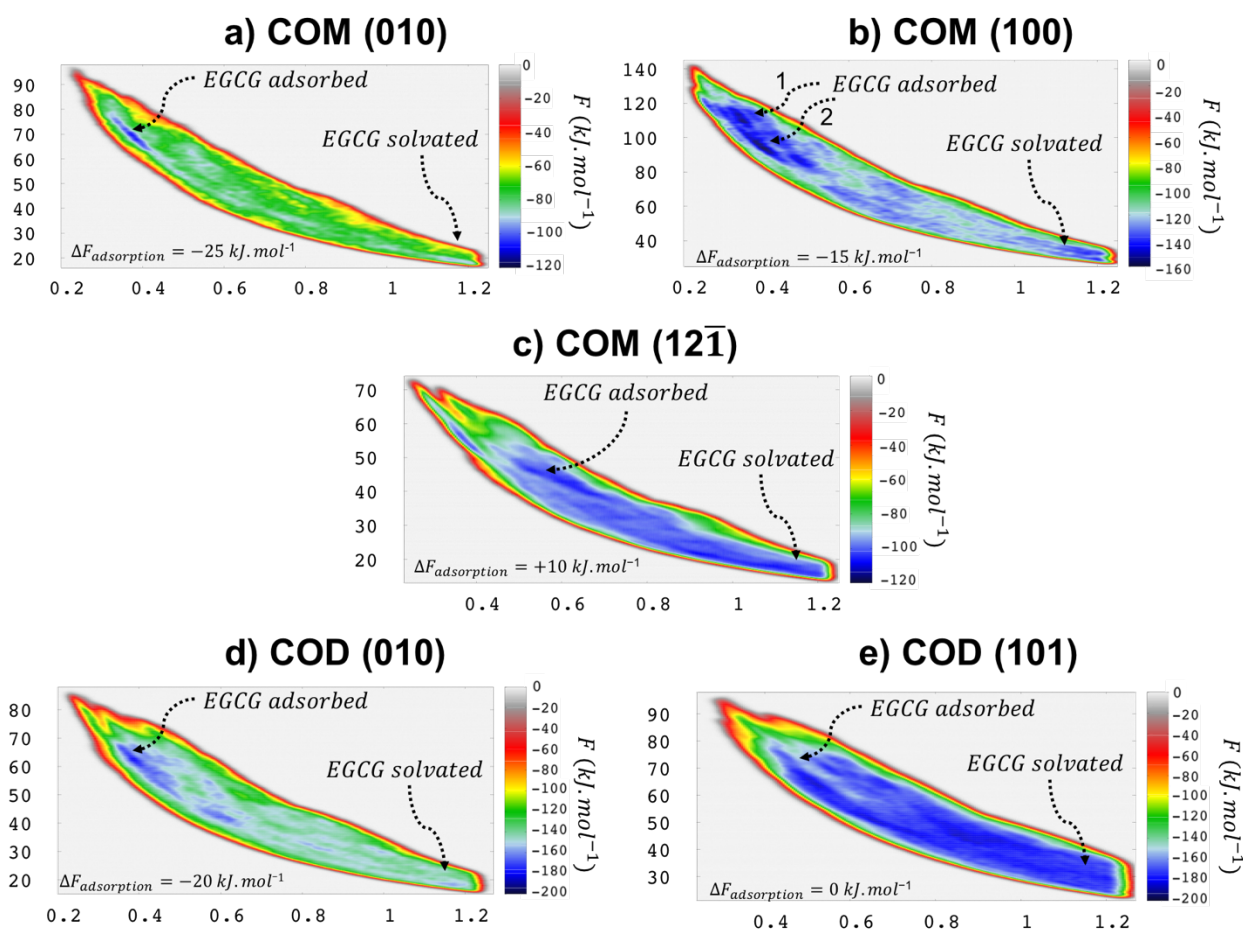
**Figure 9: EGCG adsorbed configurations exploration in the CVs 2D space ( $D$ ,  $CN$ ). MD of EGCG above COM (010) surface with (a) REMD (method 1) (b) metadynamics (method 2) (c) REMD + metadynamics (method 3). (d)  $\Delta F_{\text{adsorption}}$  free energy convergence against time.  $D$  is the Distance between the center of mass of EGCG and the surface of calcium oxalate.  $CN$  is the Coordination Number between EGCG oxygens atoms and all surface atoms of calcium oxalate**

Finally, as can be seen in Figure 9c REMD combined with metadynamics generate even more configurations. Furthermore, a better dispersion is obtained. Indeed, when  $t \approx 12-14$  ns (orange and yellow color), EGCG visits configurations covering all possible  $CN$  and  $D$  values. This means that all free energy wells are filled with the metadynamic bias allowing EGCG to freely move through all possible configurations (without any preference). This is a criterion for the free energy convergence. Figure 9-d confirms the convergence of the adsorption free energy. Starting from 6 ns, the adsorption free energy fluctuates between -30 and -20  $\text{kJ}\cdot\text{mol}^{-1}$ . Free energy of adsorption is therefore taken as -25  $\text{kJ}\cdot\text{mol}^{-1}$  with an uncertainty of  $\pm 5\text{kJ}\cdot\text{mol}^{-1}$ . REMD combine with metadynamics was the best alternative to get a good exploration of

adsorbed configurations and to reach free energy convergence. It was then used for the rest of the study.

### *Application of REMD + metadynamics to COM/COD surfaces*

The sum of gaussians bias along the CVs allowed us to reconstruct the free energy diagrams for EGCG adsorption on COM (010), (100), ( $12\bar{1}$ ) and COD (010), (101) surfaces. These free energy diagrams are presented in Figure 10.



**Figure 10 : free energy diagrams of EGCG adsorption on  $\text{CaO}_x$  surfaces.  $D$  is the Distance between the center of mass of EGCG and the surface of calcium oxalate.  $CN$  is the Coordination Number between EGCG oxygens atoms and all surface atoms of calcium oxalate**

Remarkably, major differences are found between COM/COD surfaces. For COM (010), (100) and COD (010) surfaces, a net difference is found between adsorbed state and solvated state, while for COM ( $12\bar{1}$ ) and COD (101) surfaces, the difference is much less pronounced. For them, free energy diagram colors suggest that EGCG has no clear preference to be in adsorbed, solvated or intermediate state. This observation is confirmed by adsorption free energies which

are calculated by taking the difference between minimum free energy extracted from areas labeled with “EGCG adsorbed” and “EGCG solvated”. Convergence of the adsorption energy is verified as described before for COM (010) surface. Convergence of  $\Delta F_{adsorption}$  against time can be found in the appendices, for each surface. Such adsorption free energies are comparable to those calculated by Parvaneh *et al.*<sup>34</sup> for acetate ( $\text{CH}_3\text{COO}^-$ ) adsorption on COD surfaces.

Indeed, EGCG free energy of adsorption on COD (010) surface is similar to acetate one ( $-20 \pm 5 \text{ kJ}\cdot\text{mol}^{-1}$  for EGCG against  $-23 \pm 3 \text{ kJ}\cdot\text{mol}^{-1}$  for acetate ion). At first glance, it can be surprising because acetate is an anion and interact strongly with calcium ions. This can be explained because the EGCG molecule has more coordination sites compared to acetate and is able to make more interactions with COD surface. However, a different behavior is found on the COD (101) surface ( $0 (\pm 5) \text{ kJ}\cdot\text{mol}^{-1}$  for EGCG against  $-12 (\pm 6) \text{ kJ}\cdot\text{mol}^{-1}$  for acetate ion). A reasonable explanation is that EGCG depends on other energetic terms. For example, EGCG is more likely to have more predominant steric effects compared to acetate. Furthermore, EGCG molecule is able to adopt many more configurations than acetate. The adsorption may thus be associated to an entropic cost due to the “freezing” in one specific geometry. Finally, EGCG could be associated to a larger solvation effect. Indeed, more water molecules may be involved in the creation of the EGCG/CaOx interface during adsorption than for acetate ion.

### ***COM (010) surface***

For COM (010) surface, the free energy diagram on Figure 10a shows a stable area for CN between 60 and 80 and D around 0.4 nm. Adsorbed geometries in the stable free energy area are described in Figure 11. In Figure 11a, EGCG is able to interact with COM (010) calcium ions by the mean of two of its hydroxyl groups. However, the geometry presented in Figure 11b is the most observed one: here, EGCG exhibits a short-range interaction with calcium ion by the mean of its COO group.

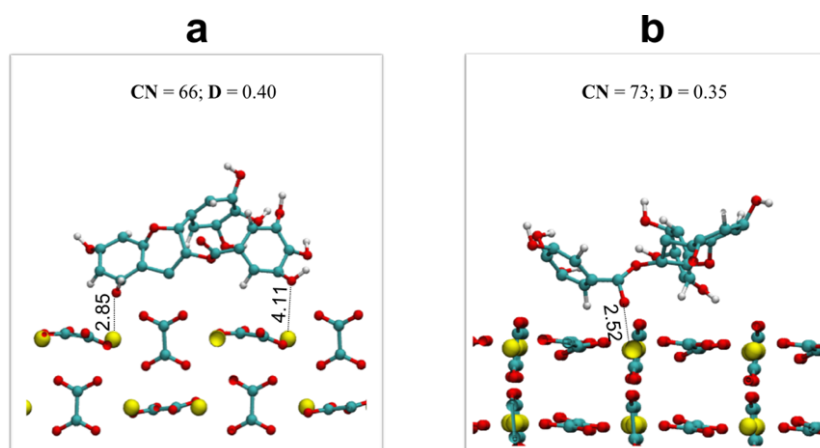


Figure 11: Most stable EGCG adsorbed geometries on COM (010) surface.

### COM (100) surface

A more complex landscape was found for COM (100) surface (see Figure 10b). First, a free energy well is found in the area of  $CN \approx 80$  and  $D \approx 0.55$ .

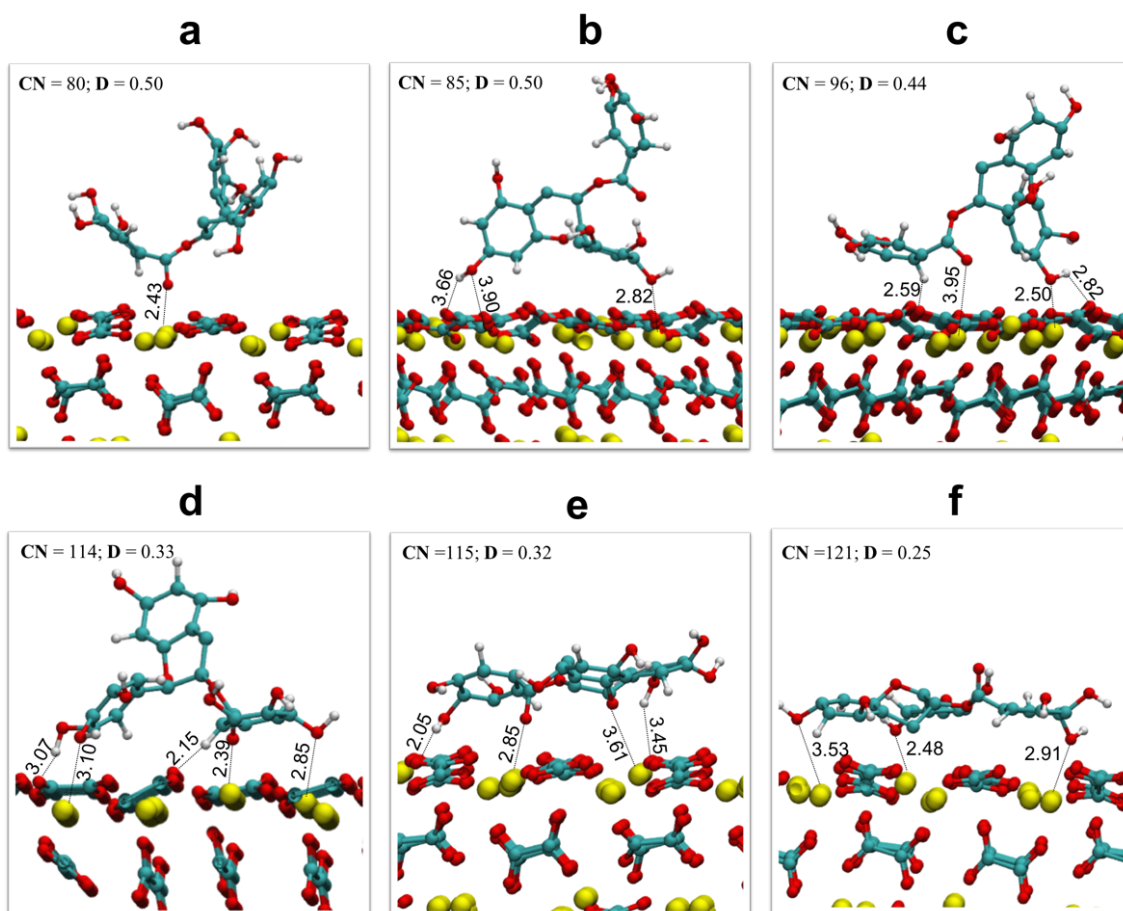


Figure 12: Most stable EGCG adsorbed geometries on COM (100) surface.

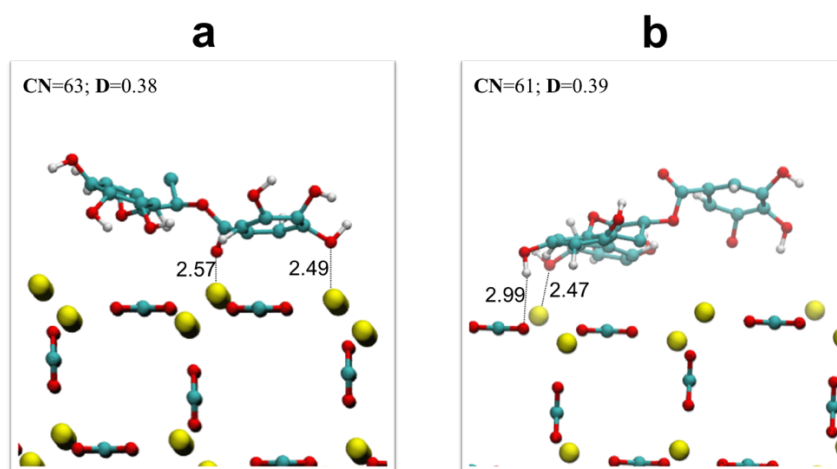
Adsorbed geometries presented in Figure 12a-c are similar to configurations observed on COM (010) surface. In these configurations EGCG interact with COM surface through its COO group. Figure 12b is the most stable adsorption geometry in the EGCG **adsorbed area 1** (Figure 10b). In this configuration, EGCG makes two stable interactions between hydroxyl groups and  $\text{Ca}^{2+}$  ions on COM (100) surface. Configurations with higher CN were found in the **adsorbed area 2** (Figure 10b). This is a specificity of COM (100) surface compared to (010) where no adsorbed configuration was found above  $\text{CN} = 80$ . In Figure 12d, EGCG is “orthogonal” to the surface establishing some contacts by the mean of both hydroxyl and COO groups. In Figure 12e,f EGCG is able to maximize interactions, by orienting itself along COM (100) surface in a ‘flat’ geometry. This highly coordinated geometry was only observed for COM (100) surface.

### **COM (12 $\bar{1}$ ) surface**

COM (12 $\bar{1}$ ) face was the only one with a positive EGCG adsorption free energy (see Figure 10c). In chapter 2, we have found that this surface is subject to a strong solvation effect ( $\gamma_{hydrated} - \gamma_{vacuum} = -0.17 \text{ J.m}^{-2}$ ) contrary to (010) and (100) surfaces ( $\gamma_{hydrated} - \gamma_{vacuum} = +0.11$  and  $+0.03 \text{ J.m}^{-2}$ , respectively). This specificity of COM (12 $\bar{1}$ ) surface may explain why EGCG adsorption is unfavorable.

### **COD (010) surface**

For COD (010) surface, a significant difference was found between solvated and adsorbed areas in the free energy diagram (see Figure 10d). Free energy difference between minima of these two areas gives a  $\Delta F_{adsorption} \simeq -20 (\pm 5) \text{ kJ/mol}$ . The most stable adsorbed configurations have a CN around 60 and a D around 0.4.

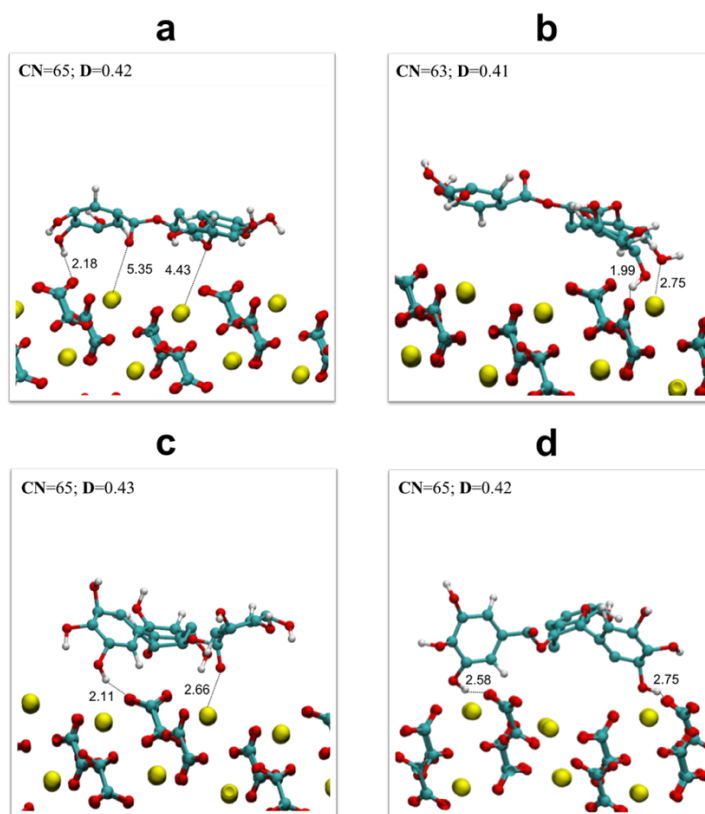


**Figure 13: most stable EGCG adsorbed geometries on COD (010) surface**

These configurations are presented in Figure 13. Figure 13a shows geometries similar to those already observed for COM surfaces. EGCG interacts with COD (010) surface with its COO group while in Figure 13b, EGCG makes one hydrogen bond and one Ca-O interaction through hydroxyl groups.

### *COD (101) surface*

Free energy diagram corresponding to COD (101) surface has a “flat” aspect (see Figure 10e). A free energy of adsorption  $\approx 0$  ( $\pm 5$ ) kJ/mol is found which is weaker than for COD (010) surface. Most stable configurations are in the area of CN  $\approx 65$  and D  $\approx 0.4$ . Adsorbed configurations in this area are presented in Figure 14. In Figure 14a,c, COO group stabilizes the adsorption state by interacting with calcium ions. In all adsorbed configurations some hydrogen bonds between EGCG OH groups and oxalate ions are observed, as well. These configurations are quite similar to the ones observed on the other faces. Therefore, the difference in free energy between COD (101) face and the other surfaces does not seem to depend on EGCG geometry in the adsorbed state. This suggests the importance of the solvation effect in the free energy contribution.



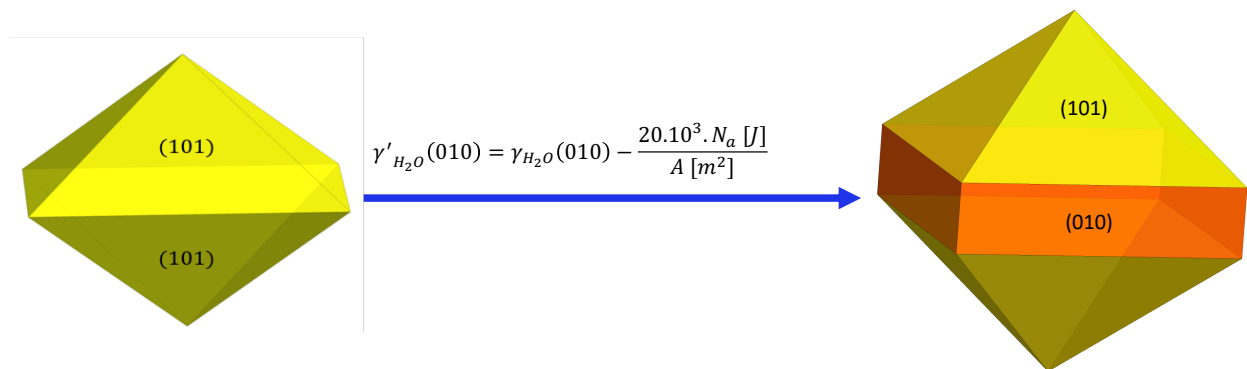
*Figure 14: most stable EGCG adsorbed geometries on COD (101) surface*

Table 1 summarizes all calculated free energies. Our results were able to demonstrate that the adsorption free energy is depending on the exposed COM or COD surface. As described by Millan<sup>13</sup>, COM (010) and (100) faces are the first surfaces to be observed in the early stage of crystallization. By strongly interacting with both surfaces, EGCG may be able to inhibit COM crystal growth in orthogonal directions that would explain green tea ability to orient crystallization toward COD instead of COM.

Structure (surface)	COM (010)	COM (100)	COM (12 $\bar{1}$ )	COD (010)	COD (101)
$\Delta F_{adsorption}$ (kJ.mol <sup>-1</sup> )	-25 ( $\pm 5$ )	1: -15 ( $\pm 5$ ) 2: -12 ( $\pm 5$ )	+10 ( $\pm 5$ )	-20 ( $\pm 5$ )	0 ( $\pm 5$ )

**Table 1: Free energies of EGCG adsorption on calcium oxalate surfaces obtained by REMD + metadynamics for COM and COD.**

In the case of COD, EGCG interacts preferentially with COD (010) surface. Indeed, the difference between both surface energies is  $20 \pm 10$  kJ.mol<sup>-1</sup>. This anisotropic stabilization of COD may explain why green tea is responsible of the emergence of COD (100) surface. Indeed, a Wulff construction<sup>35</sup> with a correction of -20 kJ.mol<sup>-1</sup> applied to (010) surface energy (obtained from **chapter 2**), is enough to lead to the emergence of (010) face in the morphology (see Figure 15).



**Figure 15 : effect of COD (010) surface stabilization by EGCG (-20 kJ.mol<sup>-1</sup>) on COD morphology. A [m<sup>2</sup>] is the area of the (010) face model (6.125×10<sup>-17</sup> m<sup>2</sup>).**

Finally, our results do not explain why EGCG degrade specifically COD (and not COM) surfaces as observed experimentally (Figure 2). To go further, COM and COD surface reconstructions in presence of EGCG should be investigated. For that, constraints on calcium

and oxalates ions should be removed. Furthermore, a collective variable is needed to take into account the dissolution process of calcium oxalate. In that way, we should be able to evaluate associated free energy barriers. Coordination number between  $\text{Ox}^{2-}/\text{Ca}^{2+}$  ions and water molecules should be a relevant collective variable to take into account this transformation.



## **Conclusion**

From experimental studies, green tea is known *(i)* to orient crystallization of calcium oxalate towards COD at the expense of COM, *(ii)* to promote COD (010) surface and *(iii)* to cause the emergence of cavities from 200 to 500 nm on COD (101) surface.

Our results suggest that EGCG is able to inhibit COM crystal growth by strongly interacting with COM (100) and (010) surfaces.

In the COD case, the interaction with EGCG is much weaker for (101) face than for COD (010) face. This anisotropic stabilization of COD would explain why this specific face emerge in syntheses in presence of green tea.

Finally, the relaxation of the surface has not been taken into account. Relaxation terms in the adsorption energy may play an important role, in the COD surface reconstruction and the experimental observation of cavities due to green tea.

Our methodology opens the way to the study of other organic molecules and should be useful to predict their potential interest in urolithiasis therapy. A next step will be to fully automatize the method in order to be able to predict the therapeutic interest of any molecule.

## Acknowledgements

This work was performed using HPC resources from GENCI- [CCRT/CINES/IDRIS] (Grant 2017-[x2017082022]) and the CCRE of Université Pierre et Marie Curie.

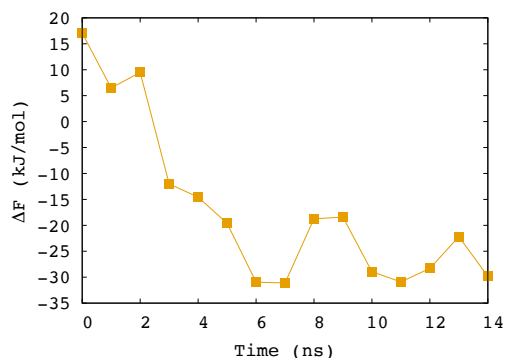
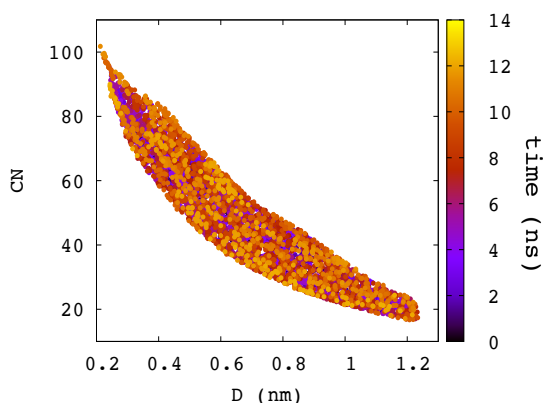
## References

1. Bettuzzi, S. *et al.* Chemoprevention of Human Prostate Cancer by Oral Administration of Green Tea Catechins in Volunteers with High-Grade Prostate Intraepithelial Neoplasia: A Preliminary Report from a One-Year Proof-of-Principle Study. *Cancer Res.* **66**, 1234–1240 (2006).
2. Graham, H. N. Green tea composition, consumption, and polyphenol chemistry. *Prev. Med. (Baltim).* **21**, 334–350 (1992).
3. Higdon, J. V. & Frei, B. Tea Catechins and Polyphenols: Health Effects, Metabolism, and Antioxidant Functions. *Crit. Rev. Food Sci. Nutr.* **43**, 89–143 (2003).
4. Singh, B. N., Shankar, S. & Srivastava, R. K. Green tea catechin, epigallocatechin-3-gallate (EGCG): Mechanisms, perspectives and clinical applications. *Biochem. Pharmacol.* **82**, 1807–1821 (2011).
5. Hirose, M. *et al.* Renal tubular epithelial cell injury and oxidative stress induce calcium oxalate crystal formation in mouse kidney. *Int. J. Urol.* **17**, 83–92 (2010).
6. Itoh, Y. *et al.* Preventive Effects Of Green Tea On Renal Stone Formation and the Role of Oxidative Stress in Nephrolithiasis. *J. Urol.* **173**, 271–275 (2005).
7. Chen, Z., Wang, C., Zhou, H., Sang, L. & Li, X. Modulation of calcium oxalate crystallization by commonly consumed green tea. *Cryst. Eng. Comm* **12**, 845–852 (2010).
8. Leroy, C. Oxalates de calcium et hydroxyapatite : des matériaux synthétiques et naturels étudiés par techniques RMN et DNP. (Université P. et M. Curie, 2016).
9. Julie, R. *et al.* Influence of daily green tea infusions on hypercalciuric renal stone patients: biological and morpho-constitutional analysis. **46**, 8.27 (2016).
10. Benzerara, Y., Bazin, D., Daudon, M., Haymann, J.-P. & Rode, J. Catechin for treating Renal Lithiasis. (2015).
11. Chen, Z., Wang, C., Zhou, H., Sang, L. & Li, X. Modulation of calcium oxalate crystallization by commonly consumed green tea. *CrystEngComm* **12**, 845 (2010).

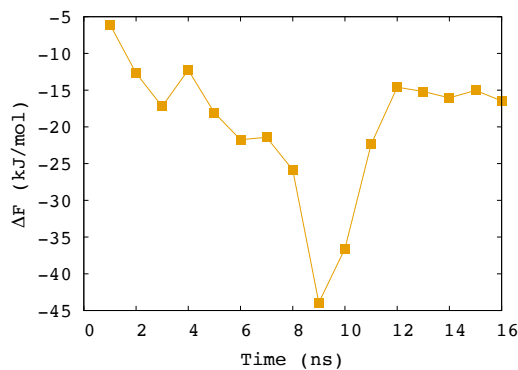
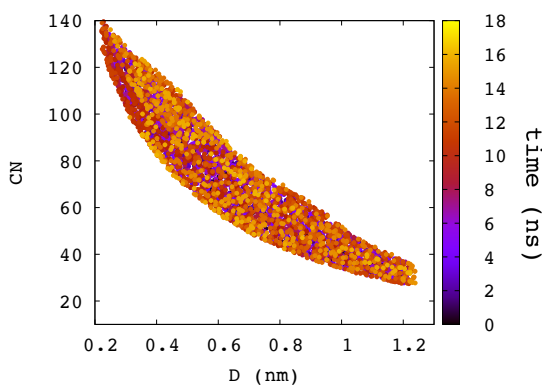
12. Zhang, D., Qi, L., Ma, J. & Cheng, H. Morphological Control of Calcium Oxalate Dihydrate by a Double-Hydrophilic Block Copolymer. (2002). doi:10.1021/cm010768y
13. Millan, A. Crystal Growth Shape of Whewellite Polymorphs: Influence of Structure Distortions on Crystal Shape. *Cryst. Growth Des.* **1**, 245–254 (2001).
14. Heijnen, W. M. M. & Van Duijneveldt, F. B. The theoretical growth morphology of calcium oxalate dihydrate. *J. Cryst. Growth* **67**, 324–336 (1984).
15. Daudon, M. *et al.* Examination of whewellite kidney stones by scanning electron microscopy and powder neutron diffraction techniques. *J. Appl. Crystallogr.* **42**, 109–115 (2009).
16. Tazzoli, V. & Domeneghetti, C. The crystal structures of whewellite and weddellite: re-examination and comparison. *Am. Mineral.* **65**, 27–334 (1980).
17. Schmid, N. *et al.* Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur. Biophys. J.* **40**, 843–856 (2011).
18. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
19. Páll, S., Abraham, M. J., Kutzner, C., Hess, B. & Lindahl, E. Tackling Exascale Software Challenges in Molecular Dynamics Simulations with GROMACS. in 3–27 (Springer, Cham, 2015). doi:10.1007/978-3-319-15976-8\_1
20. Kahlen, J., Salimi, L., Sulpizi, M., Peter, C. & Donadio, D. Interaction of Charged Amino-Acid Side Chains with Ions: An Optimization Strategy for Classical Force Fields. *J. Phys. Chem. B* **118**, 3960–3972 (2014).
21. Malde, A. K. *et al.* An Automated Force Field Topology Builder (ATB) and Repository: Version 1.0. *J. Chem. Theory Comput.* **7**, 4026–4037 (2011).
22. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926 (1998).
23. Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **52**, 255–268 (1984).
24. Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **31**, 1695–1697 (1985).
25. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).

26. Essmann, U. *et al.* A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577 (1995).
27. Hess, B. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *J. Chem. Theory Comput.* **3**, 116–122 (2008).
28. Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C. & Bussi, G. PLUMED 2: New feathers for an old bird. (2013). doi:10.1016/j.cpc.2013.09.018
29. Barducci, A., Bussi, G. & Parrinello, M. Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Phys. Rev. Lett.* **100**, 020603 (2008).
30. Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C. & Bussi, G. PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **185**, 604–613 (2014).
31. Laio, A. & Gervasio, F. L. Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Reports Prog. Phys.* **71**, 126601 (2008).
32. Rhee, Y. M. & Pande, V. S. Multiplexed-Replica Exchange Molecular Dynamics Method for Protein Folding Simulation. *Biophys. J.* **84**, 775–786 (2003).
33. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
34. Parvaneh, L. S., Donadio, D. & Sulpizi, M. Molecular Mechanism of Crystal Growth Inhibition at the Calcium Oxalate/Water Interfaces. *J. Phys. Chem. C* **120**, 4410–4417 (2016).
35. Zucker, R. V., Chatain, D., Dahmen, U., Hagège, S. & Carter, W. C. New software tools for the calculation and display of isolated and attached interfacial-energy minimizing particle shapes. *J. Mater. Sci.* **47**, 8290–8302 (2012).

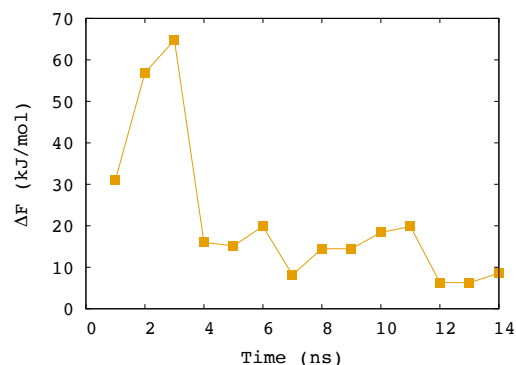
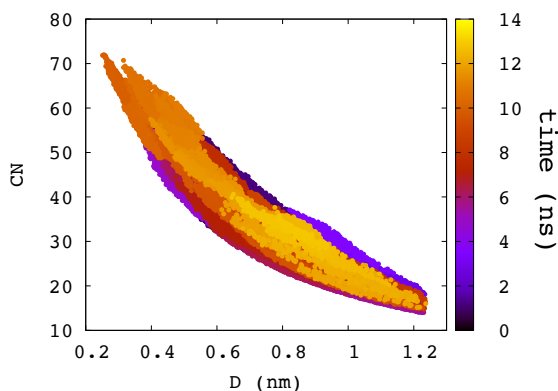
## Appendices



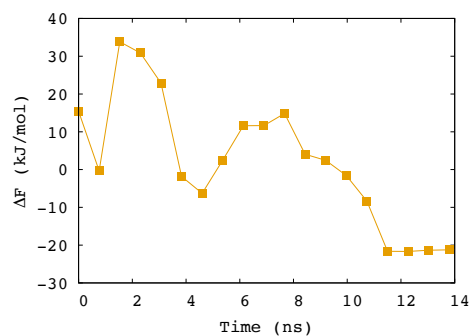
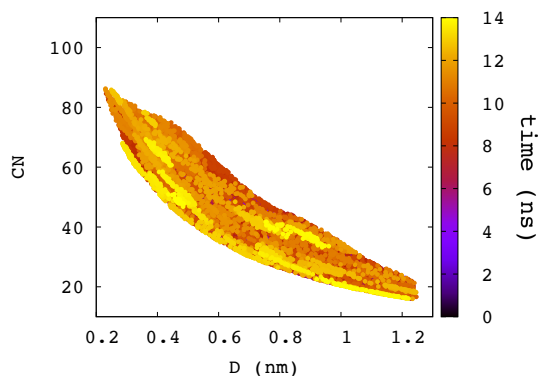
**Appendix 1 : (a) EGCG adsorbed configurations exploration in the CVs 2D space ( $D$ ,  $CN$ ). MD of EGCG above COM (010) surface with REMD + metadynamics. (b)  $\Delta F_{adsorption}$  free energy convergence against time.**



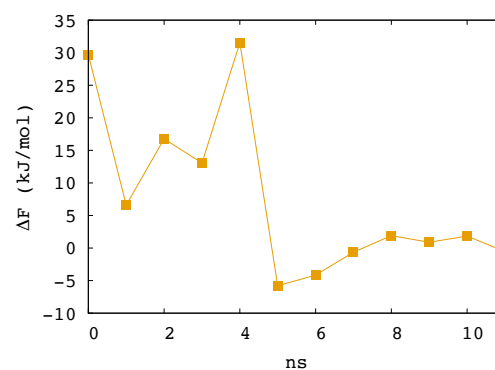
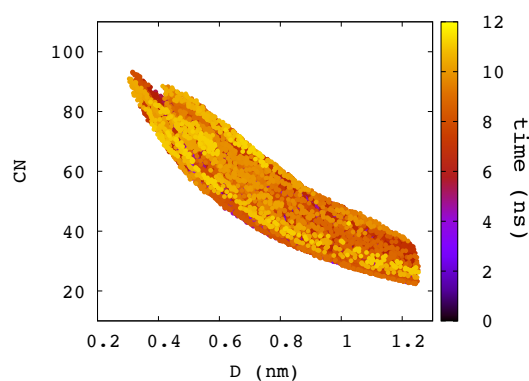
**Appendix 2 : (a) EGCG adsorbed configurations exploration in the CVs 2D space ( $D$ ,  $CN$ ). MD of EGCG above COM (100) surface with REMD + metadynamics. (b)  $\Delta F_{adsorption}$  free energy convergence against time.**



**Appendix 3 : (a) EGCG adsorbed configurations exploration in the CVs 2D space ( $D$ ,  $CN$ ). MD of EGCG above COM ( $12\bar{1}$ ) surface with REMD + metadynamics. (b)  $\Delta F_{adsorption}$  free energy convergence against time.**



**Appendix 4 : (a) EGCG adsorbed configurations exploration in the CVs 2D space ( $D$ ,  $CN$ ). MD of EGCG above COD (010) surface with REMD + metadynamics. (b)  $\Delta F_{adsorption}$  free energy convergence against time.**



**Appendix 5 : (a) EGCG adsorbed configurations exploration in the CVs 2D space ( $D$ ,  $CN$ ). MD of EGCG above COD (101) surface with REMD + metadynamics. (b)  $\Delta F_{adsorption}$  free energy convergence against time.**

## - Chapter 4 -

# Calcium Phosphate Nucleation: A Molecular Dynamics Study

*Théau Debroyse<sup>1</sup>, Thierry Azaïs<sup>1</sup>, Fabio Pietrucci<sup>2</sup>, Marco Saitta<sup>2</sup>, Christian Bonhomme<sup>1</sup> and Frederik Tielens<sup>3\*</sup>*

<sup>1</sup> Sorbonne Université, CNRS, Laboratoire Chimie de la Matière Condensée de Paris, LCMCP, 4 Place Jussieu, F-75005 Paris, France

<sup>2</sup> Sorbonne Université, CNRS, Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie, IMPMC, 4 Place Jussieu, F-75005 Paris, France

<sup>3</sup> General Chemistry (ALGC) Materials Modeling Group, Vrije Universiteit Brussel (Free University Brussels-VUB), Pleinlaan 2, 1050 Brussel, Belgium

\*Author to whom correspondence should be addressed: [frederik.tielens@vub.be](mailto:frederik.tielens@vub.be)

<b>Abstract.....</b>	<b>126</b>
<b>Introduction .....</b>	<b>127</b>
<b>Methods.....</b>	<b>131</b>
<b>Results and discussion .....</b>	<b>139</b>
<b>Conclusion.....</b>	<b>146</b>
<b>Acknowledgements .....</b>	<b>147</b>
<b>References .....</b>	<b>147</b>

## **Abstract**

The first steps of calcium phosphate precipitation, a crucial process for bone biomineralization in vertebrates, was investigated by the mean of molecular dynamics and metadynamics. It was possible to fully reconstruct the free energy diagram of the transformation going from a supersaturated solution to amorphous calcium phosphate nanoparticles of  $\approx 2$  nm diameter and shed some light on the mechanism. We confirmed that a non-classical nucleation pathway is more suitable to describe calcium phosphate nucleation. Furthermore, we supported the experimental observation of a metastable state composed of prenucleation clusters with  $\approx 0.9$  nm diameter.



## **Introduction**

### ***Bone biomineralization***

Bone is a complex hybrid material made of an organic matrix with collagen fibrils (22 wt%) containing an embedded mineral phase made of hydroxyapatite (64 wt%). Other components are water (9 wt%), non-collagen proteins (3 wt%) and cells (2 weight%)<sup>1</sup>. Hydroxyapatite (HAP) phase ( $\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$ ) which is the most stable calcium phosphate phase<sup>2,3</sup> increases bone's stiffness because of its high Young's modulus (80-120 GPa<sup>4</sup>). Furthermore, it is a dynamic material, interacting with its biological environment and acting as an ionic reservoir for the living organism. Hydroxyapatite is also found in the dental tissue or in pathological calcifications such as kidney stones or Randall's plaque<sup>5</sup>. Bone's biomineralization, the process used by the living organism to induce mineral formation, has been extensively studied in the literature<sup>6-14</sup>.

Various theories have been proposed to understand the mechanisms involved in its growth process. The major difficulty is to understand how the living organism is able to specifically induce calcium phosphate precipitation inside the bone or dental tissues and not in non-mineralized tissues (like blood vessels) though that they are located in a similar extra cellular medium<sup>15</sup>. Understanding key factors that induce mineralization is also crucial to understand the formation of pathological calcifications.

Cells inside the bone's structure named as osteoblasts are known to be involved in the formation of new bone. First assumptions made in the 1920s argued that calcium phosphate nucleation occurs in extracellular vesicles coming from osteoblasts. Precipitation is triggered by the supersaturation concentration induced inside these vesicles. Phosphate concentration is increased by organic phosphate metabolism<sup>6,7</sup>, while calcium concentration is increased by the means of ionic transporters<sup>8</sup>. However, microscopy measurements have highlighted the presence of apatite crystals in specific zones inside the collagen matrix, named the "gap region"<sup>9-11</sup> where nucleation would take place. Furthermore, HAP crystal growth is oriented by collagen fibers along the c axis. A mechanism of heterogeneous nucleation has been proposed. According to this theory, nucleation would be induced and oriented by the collagen itself<sup>12,13</sup>. Non-collagen proteins present inside the bone structure are also suspected to play an important role in the mineral growth and in its inhibition. SIBLINGs proteins for example interact with calcium ions by the mean of negatively charged amino acids (glutamate,

aspartate and phosphoserine) and are able to increase nucleation rate *in vitro* and *in vivo*<sup>14</sup>. There is currently no technique allowing the observation of the nucleation process *in vivo*, that is why it is difficult to conclude where the first steps of nucleation are taking place inside the bone structure. Nevertheless, the common strategy that seems to be adopted by the living organism is to increase locally the concentration of calcium and phosphate ions to reach supersaturation and trigger the precipitation.

### ***Calcium Phosphate Nucleation***

Beyond the complexity of the bone material, calcium phosphate nucleation is not fully understood, and many models coexist. First, classical nucleation theory (CNT) assumes that crystal nucleation occurs when the cluster - having the same atomic structure as the bulk crystalline phase - has reached a size large enough to overcome a free energy barrier. This barrier depends on the competition between the negative free energy of bulk formation ( $V\Delta g$ ) and the positive surface free energy ( $S\sigma$ ) related to the formation of a new liquid / solid interface<sup>16</sup>. The bulk free energy is proportional to the cluster volume  $V$  while the surface free energy is proportional to the surface  $S$ . The total free energy is then written:

$$\Delta G = V\Delta g + S\sigma \quad (1)$$

With a spherical shape approximation, the following equation is obtained:

$$\Delta G = \frac{4}{3}\Pi r^3 \Delta g + 4\Pi r^2 \sigma \quad (2)$$

Thus, for a small size ( $r \ll r_c$ ), the unfavorable surface energy term is preponderant, while the favorable bulk energy term becomes important with the increase of the size of the cluster. The cluster must then exceed a certain size for its growth to become energetically favorable. Thus, to overcome this barrier the system must have a sufficient density fluctuation to allow it to exceed a certain critical size from which growth process becomes energetically favorable:

$$r_c = \frac{-2\sigma}{\Delta g} \quad (3)$$

Over the last decade, several experimental studies have hypothesized the presence of prenucleation clusters in supersaturated ionic solutions<sup>17-19</sup>. These studies argued that there is

another more complex nucleation pathway than in the classical theory, where the supersaturated solution passes through a metastable state consisting of complexes and clusters that aggregate together to form amorphous calcium phosphate (ACP) and then turn into a crystalline phase.

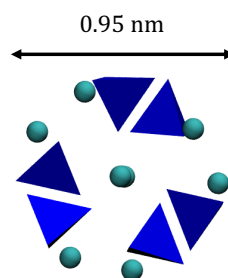
Furthermore, ACP has been observed in several forms of biological materials (tooth, bone, shells)<sup>20</sup> suggesting the passage through ACP before a transition to HAP crystalline phase. Literature described many different ACP types, depending mainly on Ca/P ratio. Indeed, depending on the pH,  $\text{PO}_4^{3-}$  ions can be protonated to  $\text{HPO}_4^{2-}$  and/or  $\text{H}_2\text{PO}_4^{2-}$  ions modifying the Ca/P ratio. Amorphous Tricalcium Phosphate (ATCP) is described in the literature as a common precipitate in basic solution (pH range 9-11)<sup>21</sup>. Indeed, in these conditions,  $\text{PO}_4^{3-}$  ions are not protonated and a Ca/P ratio of 1.5 is needed to ensure charge neutrality.

### ***Prenucleation complexes***

The role of pre-nucleation complexes has also been described in Habraken *et al.*<sup>22</sup> study, the brute formula  $[\text{Ca}(\text{HPO}_4)_3]^{4-}$  was established from titration measurements during the first stages of nucleation. The presence of a balance between these clusters and  $[\text{Ca}(\text{HPO}_4)]$  ion pair. *Ab initio* calculations confirmed the stability of  $[\text{Ca}(\text{HPO}_4)_3]^{4-}$  compared to  $[\text{Ca}(\text{HPO}_4)_2]^{2-}$ ,  $[\text{Ca}(\text{HPO}_4)_2]$  and free ions<sup>23</sup>.

### ***Prenucleation clusters***

Calcium phosphates nucleation processes were studied *in vitro* by Dey *et al.*<sup>18</sup> by CryoTEM at 37°C. (in a solution of SBF, Simulated Body Fluid, containing solid arachidic acid monolayers acting as a nucleation trigger). Clusters of size  $0.87 \pm 0.2$  nm were observed in the SBF solution even without arachidic acid monolayers. They were therefore named pre-nucleation clusters.



***Figure 1: Posner Cluster***

Their characteristic size corresponds to the so-called Posner clusters diameter (see Figure 1) supporting the hypothesis that ACP would be formed before crystallization of calcium phosphate. Indeed, it has been shown by Posner that ACP is composed of subunits of the formula  $\text{Ca}_9(\text{PO}_4)_6$  which have a diameter around 1 nm<sup>24</sup>. Adding the arachidic monolayers cause those clusters to densify above the surface forming spherical domains of 50 nm size. Amorphous nanoparticle formation of diameter ranging from 40 to 80 nm is then observed. At the end of the nucleation, those amorphous nanoparticles transformed to HAP crystalline phase. A critical radius of 50 nm for the amorphous nanoparticle was measured for this last transformation. Posner-like clusters have been detected in shell-model molecular dynamics simulations<sup>25</sup> of a supersaturated ionic solution.

### ***Goals and novelty***

Until now, it is still not clear if prenucleation clusters have a role to play in calcium phosphate nucleation. For the first time we have simulated the transformation from a calcium phosphate supersaturated solution to an amorphous nanoparticle, *reversibly*. The free energy diagram of the transformation has been fully reconstructed which allowed us to investigate the mechanism of calcium phosphate nucleation. The method opens the ways to study the influence of key parameters such as temperature, concentrations, additives, etc. on calcium phosphate nucleation.

## **Methods**

### ***Force Field***

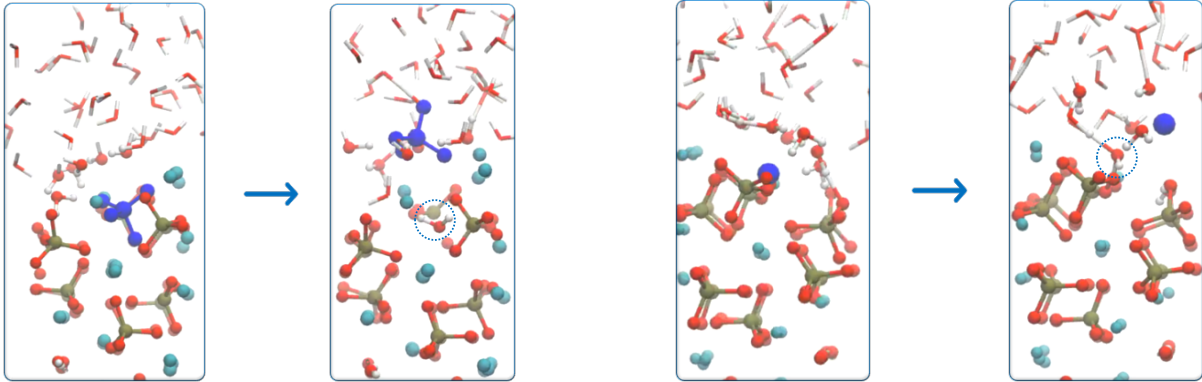
In order to validate Lin *et al.*<sup>26</sup> force field, activation free energies corresponding to (i) solvation of a  $\text{PO}_4^{3-}$  and (ii) solvation of  $\text{Ca}^{2+}$  ion from HAP (100) surface, were calculated exploiting the metadynamics technique. Both transformations were investigated with DFT-based Born–Oppenheimer molecular dynamics (BOMD) and Lin *et al.*<sup>26</sup> force field in order to compare the corresponding free energies coming from these two levels of theory. The initial configuration was deduced from positions resolved by X-Ray diffraction<sup>27</sup>. HAP (100) surface model was built by introducing a slab of bulk water between (100) planes ( $1\text{kg.L}^{-1}$ ).

Kinetic barrier involved in this phenomenon is probably too large for an efficient exploration of the configuration space within the available timescale. One approach to improve the exploration consists in using enhanced sampling techniques like metadynamics<sup>28</sup> with the help of low-dimensional collective variables (CV). Metadynamics was performed with the plumed package<sup>29</sup>.  $\text{PO}_4^{3-}$  coordinate or  $\text{Ca}^{2+}$  coordinate orthogonal to (100) plane was used as the collective variable (CV). Repulsive Gaussian bias functions of height 1 kJ/mol and width 0.015 nm were added every 0.02 ps.

Lin *et al.*<sup>26</sup> have parametrized the force field in order to reproduce hydroxyapatite properties and surface energies in water. It was developed with TIP3P water model<sup>30</sup>. We used the Particle Mesh Ewald (PME) method to optimize the calculation of long-range interactions<sup>31,32</sup> with a grid spacing of 0.12 nm and an interpolation of 4. A real space cutoff of 1 nm was used, and the Lennard-Jones interactions were truncated at the same distance. A time step of 1 fs was used and a temperature of 310K.

BOMD simulations were performed using the generalized gradient approximation exchange-correlation functional BLYP<sup>33–35</sup> with a hybrid gaussian and plane wave method<sup>36</sup>, implemented in the CP2K/Quickstep package<sup>37</sup>. Grimme (D3) corrections were included for van der Waals interactions<sup>38</sup>. The interaction of the valence electrons with the cores is described by analytic Goedecker, Teter, and Hutter (GTH) pseudopotentials<sup>39</sup>, a triple- $\zeta$  valence polarized (TZVP) Gaussian basis set was chosen for the real space representation of the valence electrons. Plane waves up to a cutoff energy of 300 Ry were used to expand the density in reciprocal space.

BOMD and force field simulations are run in the canonical ensemble (NVT) with a Nosé-Hoover thermostat to ensure thermal stability<sup>40,41</sup> at physiological temperature (310K). Equations of motion were integrated using a time step of 0.5 fs.



**Figure 2:**  $PO_4^{3-}$  (left) and  $Ca^{2+}$  (right) ion going from HAP surface to solvated state

The observed mechanism consists of an exchange between a surface ion and a water molecule. Initial and final configurations of  $PO_4^{3-}$  and  $Ca^{2+}$  surface ions are presented in blue in Figure 2. The bias imposed by metadynamics on  $PO_4^{3-}$  or  $Ca^{2+}$  ion (in blue on Figure 2) accelerates the diffusion of the ion from the HAP surface to the solvated state. A water molecule indicated by dotted circle in Figure 2 stabilized the surface defect (vacancy) caused by the transformation. The event corresponding to the exchange between the water molecule and the ion was observed in both force field and DFT-based molecular dynamics. The corresponding free energies of activation, estimated from the bias accumulated until the observation of ion solvation, are reported in Table 1. (In the case of the classical force field, the metadynamics simulations were repeated several times with different Gaussian heights and deposition frequencies to confirm the robustness of the estimated barriers.)

	Lin <i>et al.</i> force field <sup>42</sup>	AIMD
$F^\ddagger(PO_4^{3-} \leftrightarrow H_2O)$	395 kJ/mol	407 kJ/mol
$F^\ddagger(Ca^{2+} \leftrightarrow H_2O)$	321 kJ/mol	288 kJ/mol

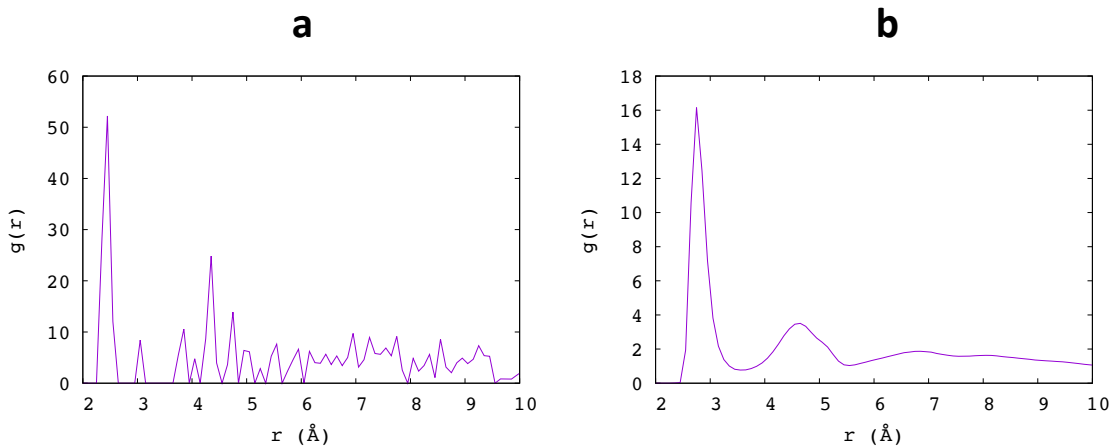
**Table 1:** Transition free energies of ion going from HAP surface to solvated state

Force field energies are very close to AIMD energies. Indeed, the relative differences between force field and AIMD free energies are equal to 1.7% and 11.5% of the AIMD free energy of activation for  $PO_4^{3-} \leftrightarrow H_2O$  and  $Ca^{2+} \leftrightarrow H_2O$  processes, respectively. This result

suggest that Lin *et al.* force field is suitable to be used in the rest of the study to access the much larger size- and time scales needed to observe nucleation.

### ***Nanoparticle model***

A crystalline nanoparticle model was built from  $\beta$  - Tri Calcium phosphate (TCP) positions resolved by neutron crystallography<sup>43</sup>. The choice of this phase has been done to ensure neutrality of the nanoparticle. A spherical nanoparticle of diameter 2nm composed of 78  $\text{Ca}^{2+}$  and 52  $\text{PO}_4^{3-}$  was built with CrystalMaker software. The cluster was then equilibrated with classical molecular dynamics in an  $8 \times 8 \times 8 \text{ nm}^3$  box which contain 16885 water molecules corresponding to a density of  $1 \text{ kg.L}^{-1}$ , during 20 ns at 310K. Figure 3a and 3b represent radial pair distribution function between  $\text{Ca}^{2+}$  ions and oxygens in  $\text{PO}_4^{3-}$ , before (a) and after (b) NVT equilibration. Initials TCP positions were deeply modified by the equilibration.



***Figure 3: radial pair distribution function between nanoparticle's  $\text{Ca}^{2+}$  ions and oxygen atoms in  $\text{PO}_4^{3-}$  before (a) and after a 20 ns equilibration (b)***

Our results correspond to a spontaneous amorphization of the initially crystalline nanoparticle, on a rapid time scale. Indeed, after 1 ns,  $g(r)$  does not evolve anymore. This result is in coherence with experimental observations, where an amorphous structure is observed for nanoparticles below  $50 \text{ nm}^{18}$ .

### ***Solution model***

A model of supersaturated solution was obtained by dissolving the nanoparticle using a metadynamics simulation. The total coordination number between all ions and water molecules was taken as the CV. A bias of  $10 \text{ kJ.mol}^{-1}$  was deposited every 0.5 ps along this

CV, causing the system to evolve towards a configuration where all ions were solvated. Box dimension and composition are the same than for the nanoparticle model and corresponds to following concentrations:  $[\text{Ca}^{2+}] = 0.25 \text{ mol.L}^{-1}$  and  $[\text{PO}_4^{3-}] = 0.17 \text{ mol.L}^{-1}$ . Lower ions concentrations in our models would imply either the use of a bigger simulation box which has a computational cost, or a decrease of the number of ions that would imply a smaller nanoparticle model. The best compromise was chosen. This final configuration was then equilibrated in an unbiased simulation during 2 ns. The spontaneous formation of some complexes could already be observed in this free molecular dynamics. ATCP solubility product constant are in the 23-25 range at 25°C<sup>21</sup>. Our concentrations correspond to a pK equal to 3.5 placing us in supersaturated conditions. These concentrations are very far from physiological conditions ( $[\text{Ca}^{2+}] = 1.23 \text{ mmol.L}^{-1}$  in blood)<sup>44</sup>, but as said before, a common strategy in biomineralization is to induce supersaturated conditions locally. Our model corresponds to such a situation.

### ***Path Permutation Invariant Vector (path-PIV) metadynamics***

Our goal is to describe the nucleation process starting from the supersaturated solution and ending to amorphous nanoparticle (see Figure 4). Molecular dynamics based on realistic force field potentials is in principle able to track such a transition. The choice of the collective variable is not trivial: it strongly depends on the transformation under investigation, and it strongly affects the efficiency of the enhanced sampling simulation. The CV must be able to distinguish between initial and final configurations and between all important intermediates we do not necessarily know. An effective metric for the comparison of condensed matter structures is found in the Permutation Invariant Vector (PIV) introduced in Gallet *et al.*<sup>45</sup> and Pipolo *et al.*<sup>46</sup>. Each component of the vector is obtained according to equation 4.

$$v_{kk'}^{\beta\beta'} = c_{kk'} f \left( \sqrt[3]{\frac{\Omega_0}{\Omega}} \|\mathbf{R}_{\beta k} - \mathbf{R}_{\beta' k'}\| \right) \quad (4)$$

$v_{kk'}^{\beta\beta'}$  is the component corresponding to the pair of the  $\beta$ -th atom of type  $k$  and the  $\beta'$ -th atom of type  $k'$ .  $\|\mathbf{R}_{\beta k} - \mathbf{R}_{\beta' k'}\|$  is the distance between these two atoms and  $f$  is a switching function monotonically decreasing from one to zero with increasing distance. The constant  $c_{kk'}$  can be chosen as 0 or 1 to switch on/off some atom pairs in the PIV calculation.  $\Omega$  and



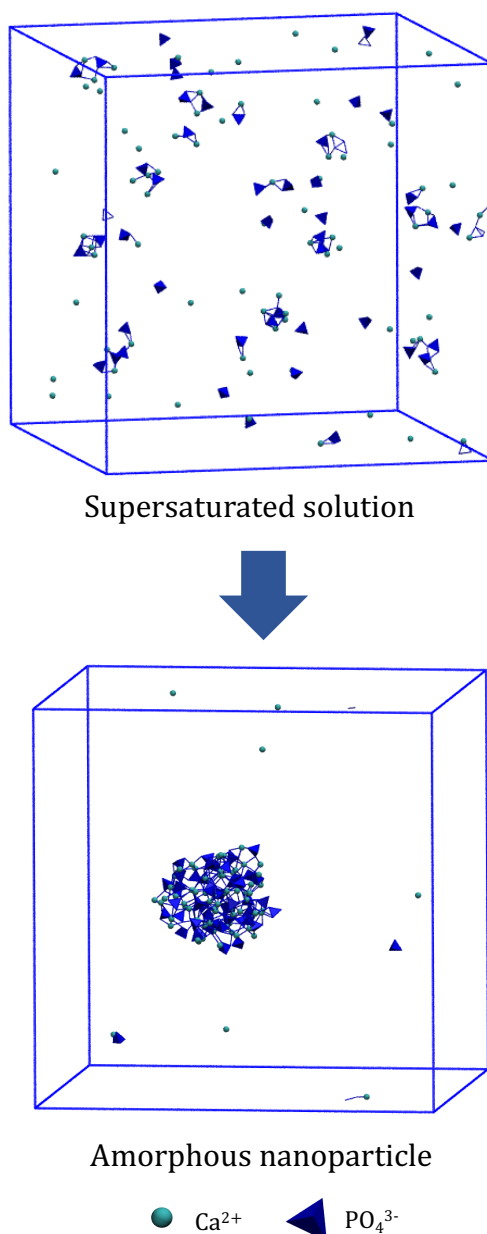
$\Omega_0$  are the volume of the simulation box and a reference volume.  $\sqrt[3]{\frac{\Omega_0}{\Omega}}$  is a scaling factor, present to avoid high fluctuations of cell parameters during metadynamics. As we are in the NVT ensemble, this parameter is not relevant and is set to 1. To obtain the final PIV, the vector is separated in  $v_{kk'}$  blocks corresponding to all components of one pair type. Each  $v_{kk'}$  block is sorted in an ascending order to obtain the final PIV vector. This sorting ensures the invariance of the vector with respect to the exchange of two atoms of the same type. It is then possible to quantify difference between two configurations by simply taking the Euclidian distance between their respective PIV. In this study, phosphorus and calcium atom types have been chosen to calculate the PIV.

The PIV metric is combined with path collective variables (PCV)<sup>47</sup> in order to define the final set of CVs. In this approach, two collective variables are defined as follow:

$$spath(t) = \frac{e^{-\lambda D(\mathbf{R}(t), \mathbf{R}_i)} + 2e^{-\lambda D(\mathbf{R}(t), \mathbf{R}_f)}}{e^{-\lambda D(\mathbf{R}(t), \mathbf{R}_i)} + e^{-\lambda D(\mathbf{R}(t), \mathbf{R}_f)}} \quad (5)$$

$$zpath(t) = -\frac{1}{\lambda} \log(e^{-\lambda D(\mathbf{R}(t), \mathbf{R}_i)} + e^{-\lambda D(\mathbf{R}(t), \mathbf{R}_f)}) \quad (6)$$

$\mathbf{R}_i$  is the configuration of the initial state of the transformation, *i.e* the supersaturated solution model and  $\mathbf{R}_f$  is the configuration of the final state, *i.e* the nanoparticle model.  $\lambda$  is a scaling parameter.  $D(\mathbf{R}(t), \mathbf{R}_{i/f})$  is the Euclidian distance between configurations' PIV at time t and the initial/final configuration's PIV. *spath* measures the progress from the initial configuration to the final configuration. It is around 1 when the system corresponds to the supersaturated solution ( $D(\mathbf{R}(t), \mathbf{R}_i) = 0$  and  $\lambda D(\mathbf{R}(t), \mathbf{R}_f) \approx 1$ ) and around 2 when the system is in the equilibrated amorphous nanoparticle configuration. Adding bias along *spath* CV generates a force acting on the system movement between the initial state and the final state. On the contrary, adding bias along *zpath* CV generates a force to get away from both initial and final configurations. It therefore allows to differentiate between different pathways and to track transitions to states different from the targeted initial and final configurations.



**Figure 4: top: calcium phosphate supersaturated solution, bottom: Amorphous Tricalcium Phosphate (ATCP) nanoparticle. Simulation box measure 8x8x8 nm. Nanoparticle has a diameter of 2.2n.  $[Ca^{2+}] = 0.25 \text{ mol.L}^{-1}$ ,  $[PO_4^{3-}] = 0.17 \text{ mol.L}^{-1}$**

A major advantage of the present formulation of *path-PIV* is that only the initial and final coordinates of the transformation are used to determine the CVs, without making any assumptions on the path connecting the two configurations. In this way, it is possible to discover prejudice-free transition pathways. The use of *path-PIV* coordinates has proven itself in the study of several challenging transitions across the water phase diagram<sup>48</sup>. Here *path-PIV* has been applied for the first time to a precipitation process from solution, specifically the calcium phosphate nucleation problem.

### ***Molecular dynamics***

NVT simulations were performed at two different temperatures (298 K and 310 K) with the same parameters as those used in the force field section (see above). The metadynamics bias was added along *spath* and *zpath* collective variables every 4 ps. A Gaussian height of 2 kJ.mol<sup>-1</sup> was chosen with  $\sigma_{spath} = 0.015$  and  $\sigma_{zpath} = 1$ . A multiple walker<sup>49</sup> setup consisting of 16 walkers was used in order to improve convergence by parallelizing the simulation, with each walker evolving independently at the same temperature, except for sharing a unique bias potential with all the other walkers.

### ***Nucleation entities nomenclature and characterization***

Many nucleation intermediates were observed in this study. To clarify the discussion, we explain here how we defined them and which vocabular has been chosen to describe them. First, a distance cutoff of 3.2 Å has been chosen between Ca<sup>2+</sup> and PO<sub>4</sub><sup>3-</sup> oxygens to define the contact between both ions. This cutoff has been chosen to fully take into account the first sphere of coordination of Ca<sup>2+</sup> ions as described by the first peak in the radial pair distribution function (see Figure 3-b). Free ions correspond to Ca<sup>2+</sup> or PO<sub>4</sub><sup>3-</sup> in a solvated state. Complexes correspond to entities, with a number of ions higher than 1. Clusters correspond to entities with a number of ions greater or equal to 10 and with radius below 1 nm. The threshold of 10 ions has been chosen as it corresponds to entities with diameter above 0.85 nm which is the size of prenucleation clusters observed by Dey *et al.*<sup>18</sup> Finally, nanoparticle corresponds to an entity with a radius of gyration equal to  $1 \pm 0.1$ . It corresponds to the final state of the simulation.

Entities were characterized along the trajectory by the mean of different measures. Their radius of gyration was calculated according to equations 7 and 8.

$$r_G = \left( \frac{\sum_i^n m_i |r_i - r_{COM}|^2}{\sum_i^n m_i} \right)^{1/2} \quad (7)$$

$$r_{COM} = \frac{\sum_i^n r_i m_i}{\sum_i^n m_i} \quad (8)$$

Instead of end-end distance, the radius of gyration is more meaningful as it gives a sense of the size of the cluster.

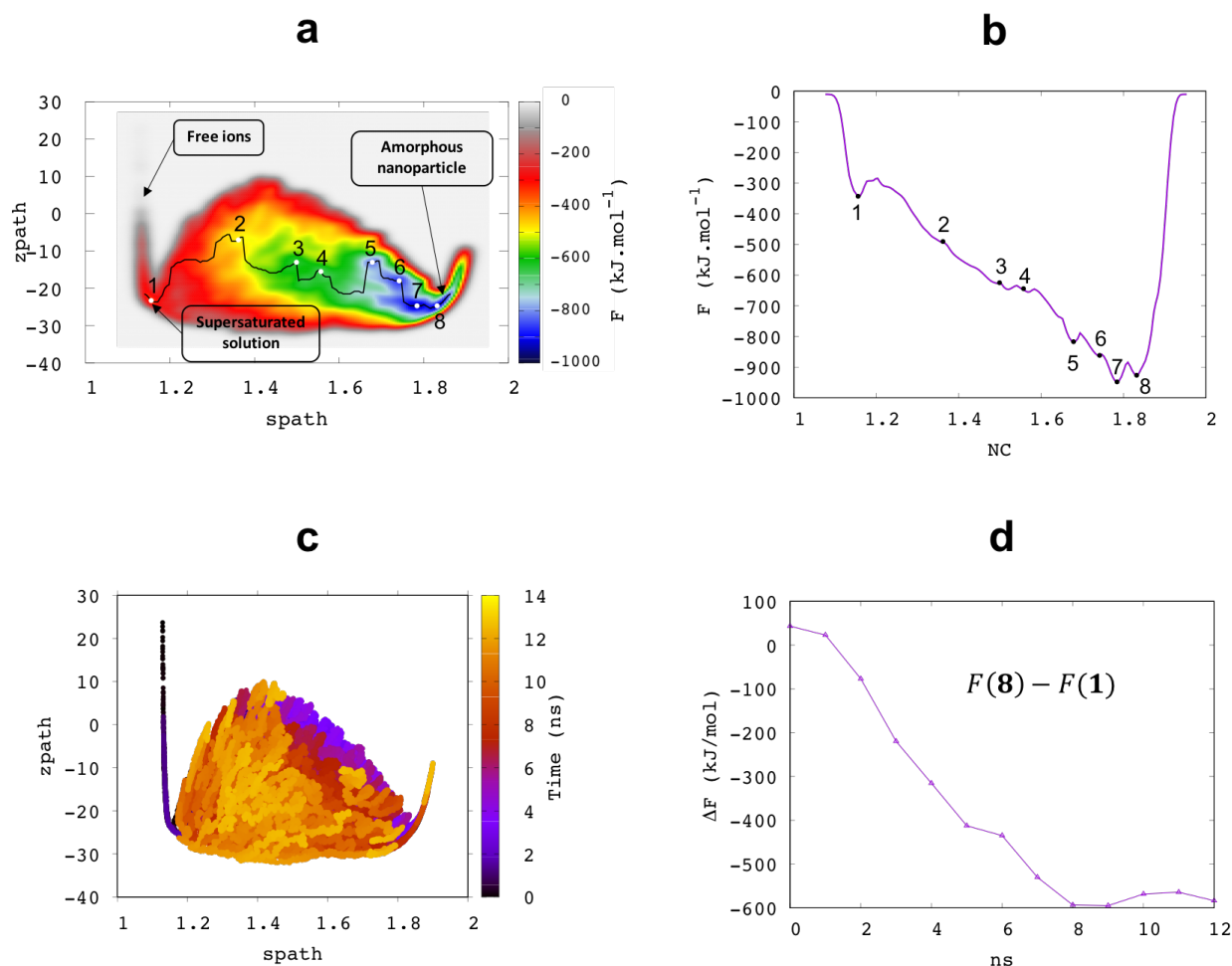
Radius of gyration was then used to calculate the cluster density assuming a spherical shape:

$$\rho_{cluster} = \frac{(N_{Ca^{2+}} + M_{Ca^{2+}} + N_{PO_4^{3-}} - M_{PO_4^{3-}})}{N a \frac{4}{3} \pi r_G^3} \quad d_{cluster} = \frac{\rho_{cluster}}{\rho_{water}} \quad (9)$$

## Results and discussion

### Calcium phosphate nucleation free energy diagram

Free energy diagram corresponding to the transformation from supersaturated solution to the nanoparticle is presented on Figure 5a.



**Figure 5:** (a) Free energy diagram of calcium phosphate nucleation from supersaturated solution (point 1) to amorphous nanoparticle model (point 8) ( $T = 310\text{K}$ ). (b) Minimum free energies project on  $\text{spath}$  (Nucleation Coordinate  $\text{NC}$ ) (c) Exploration of configurations in the  $(\text{spath}, \text{zpath})$  2D space. (d) Evolution of the free energy difference between supersaturated solution (point 1) to amorphous nanoparticle model (point 8), with respect to the time.

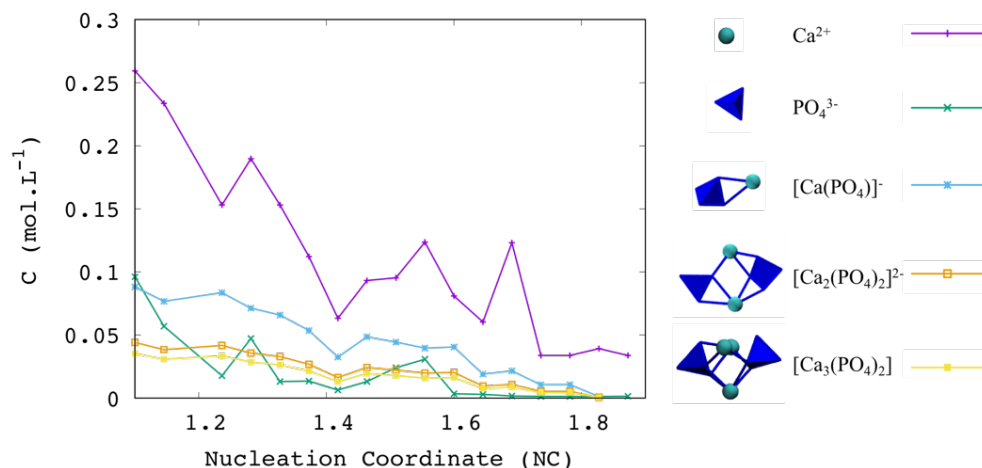
As can be seen in Figure 5d, free energy difference between supersaturated solution and nanoparticle model seems to converge. However, the system does not go to configurations in the area of point 1 and 8 during the last 2 ns (Figure 5c). This indicates that free energy wells are not completely filled. Convergence of the free energy diagram is still under investigation. That is why, free energy barriers must be taken with a grain of salt. Nevertheless, from

multiple simulations with different Gaussian heights, it appears that the path chosen by the system to perform the nucleation is well reproducible. This path is named as Nucleation Coordinate (NC) and is going through the lowest free energies along *s*path.

### ***Role of complexes and free ions***

On the free energy diagram (Figure 5a), the area labelled with point **1** corresponds to the initial supersaturated solution, located in a free energy well. The composition of the solution in this area consists of  $\text{Ca}^{2+}$  and  $\text{PO}_4^{3-}$  free ions mixed with  $[\text{Ca}(\text{PO}_4)]^-$ ,  $[\text{Ca}_2(\text{PO}_4)_2]^{2-}$  and  $[\text{Ca}_3(\text{PO}_4)_2]$  complexes. Conversely, the free energy diagram area with *s*path around 1.15 and *z*path  $\approx 0$  corresponds to a solution with free ions only. The free energy in this area is larger by a non-negligible amount ( $\approx 200$  kJ/mol). This demonstrates that a solution composed of free ions and complexes is more stable compared to a solution composed of free ions only, which is in agreement with experimental observations. Indeed, prenucleation complexes are described in Habraken *et al.*<sup>50</sup>. However, complexes found in our simulation are very different from those described in Habraken *et al* experiment because of our basic conditions.

Evolution of complexes and free ions concentrations during the nucleation process are plotted in Figure 6. The nucleation is linked to a constant uptake of free  $\text{Ca}^{2+}$  ions, which is not the case for  $\text{PO}_4^{3-}$  ions whose concentration drops to 0 at NC=1.6. This result shows that nucleation intermediates are calcium-deficient and evolve toward the neutral nanoparticle by a constant calcium uptake process. Indeed, Figure 10b shows the evolution of the Ca/P ratio of the biggest entity. At the starting point of the nucleation, the Ca/P ratio is around 1 and it evolves continuously to 1.5 along the nucleation coordinate. This observation is in agreement with Habraken *et al.*<sup>50</sup> where a calcium uptake mechanism is described. However, complexes found in our simulation are not the same due to our specific conditions. In Habraken *et al.*, pH is maintained around 7.4. This implies the presence of  $\text{HPO}_4^{2-}$  ions ( $\text{pK}_a = 12.67$  at 298K) in the solution. Our solution model is composed of only  $\text{PO}_4^{3-}$  and corresponds therefore to basic conditions.

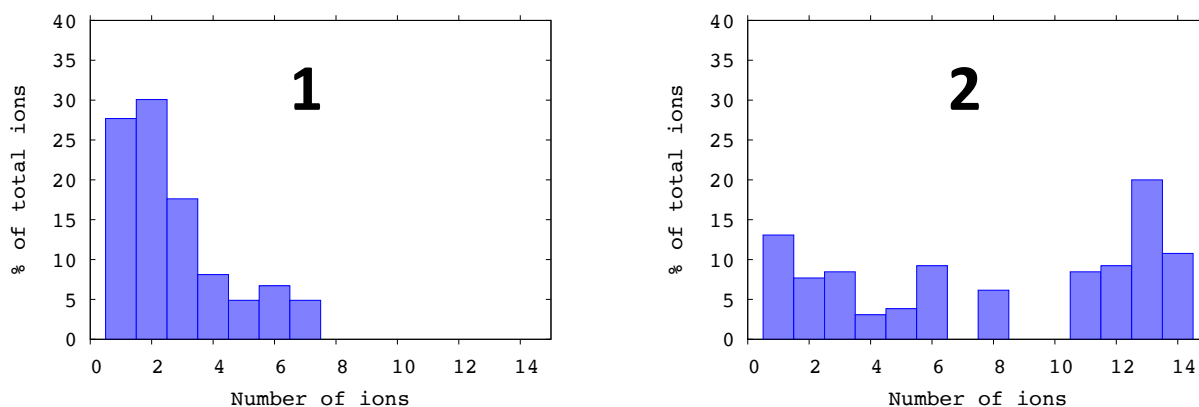


**Figure 6: Evolution of free ions and complexes concentration along nucleation coordinate.**

In those conditions, complexes of brute formula  $[\text{Ca}(\text{PO}_4)]^-$ ,  $[\text{Ca}_2(\text{PO}_4)_2]^{2-}$  and  $[\text{Ca}_3(\text{PO}_4)_2]$  play an important role in the nucleation process. All three complexes are always present along the nucleation pathway and constantly decrease. This result shows that those complexes participate actively to the growth throughout nucleation. Interestingly, the  $[\text{Ca}_3(\text{PO}_4)_2]$  geometry corresponds to a Posner cluster's subunit.

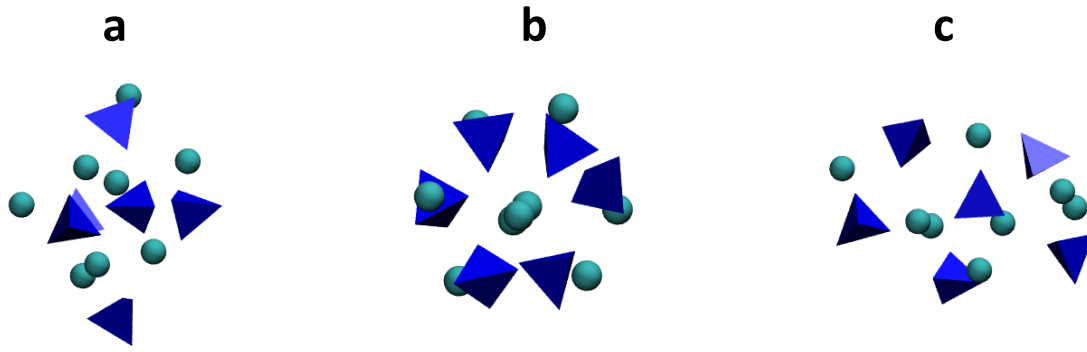
### Role of clusters

As can be seen on Figure 5b, a first barrier needs to be overcome in order to go from point **1** to **2**. During this step, the concentration of free ions and  $\text{Ca}_3(\text{PO}_4)_2$ ,  $[\text{Ca}_2(\text{PO}_4)_2]^{2-}$ ,  $[\text{Ca}(\text{PO}_4)]^-$  complexes decrease (see Figure 6). Figure 7 represent the entities' number of ions distribution. Point **1** correspond mainly to entities composed of 1, 2 and 3 ions. Transition to point **2** goes with an increase of entities composed of  $13 \pm 1$  ions.



**Figure 7: Distribution of nucleation entities' size for point 1 and 2 (see Figure 5) corresponding to  $s_{\text{path}} = 1.16; 1.36$  along NC coordinate.**

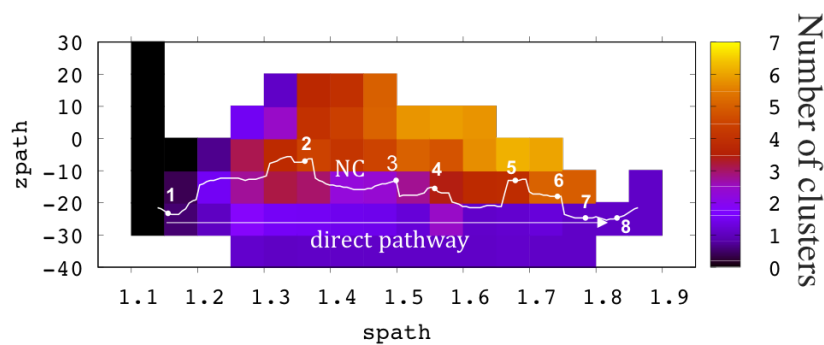
These last clusters have an average gyration radius of 4.3 Å ( $d = 0.86$  nm) that correspond to prenucleation clusters' diameter as described in Habraken *et al.*<sup>50</sup> and Dey *et al.*<sup>18</sup> ( $d = 0.87 \pm 0.2$  nm). After a careful observation of configurations in the area of point **2**, we were able to find Posner like clusters structures (Figure 8b), but also entities having a lower symmetry (Figure 8a,c).



*Figure 8: clusters found in configurations corresponding to point 2.*

In our conditions, we did not find a metastable state where these clusters are predominant. Indeed, there is no free energy well around point **2** in Figure 5b. This may be due to our supersaturated conditions. That is why, a study about the influence of the concentration on the presence of metastable states will need to be done in the future.

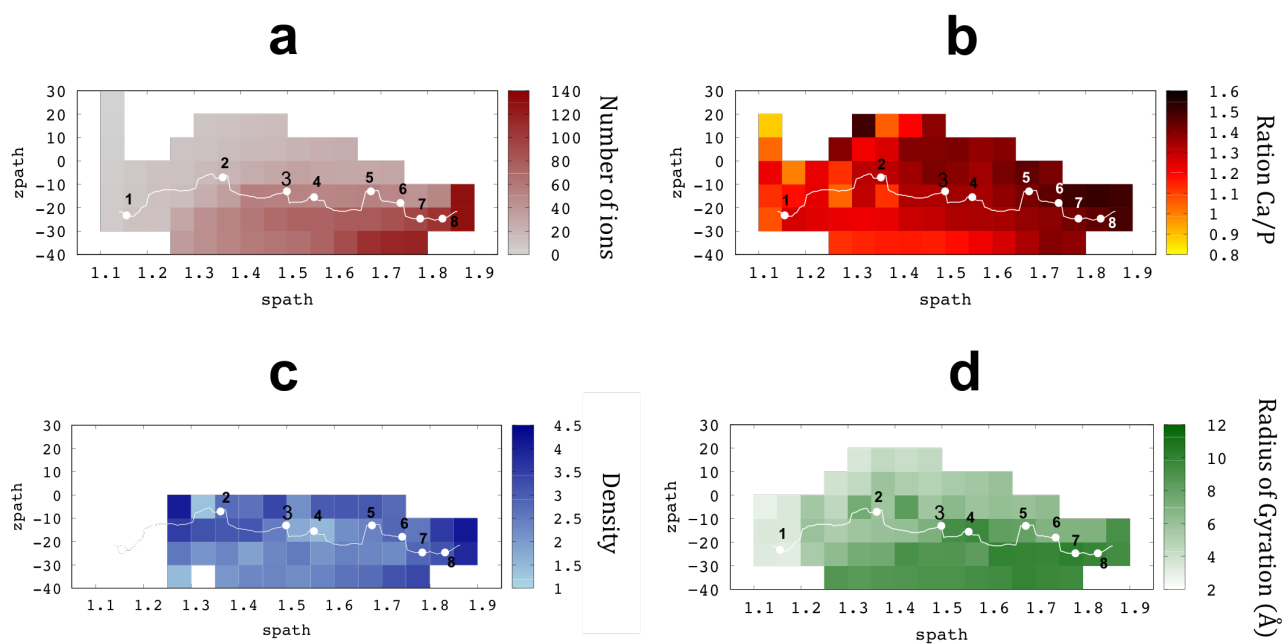
Figure 9 represent the average number of clusters in the *spath*, *zpath* space. The direct pathway consists of configurations with a unique entity along *spath*.



*Figure 9 : Number of clusters in the spath/zpath space. A cutoff of 10 ions was chosen to define a cluster. White line corresponds to the nucleation coordinate (NC).*

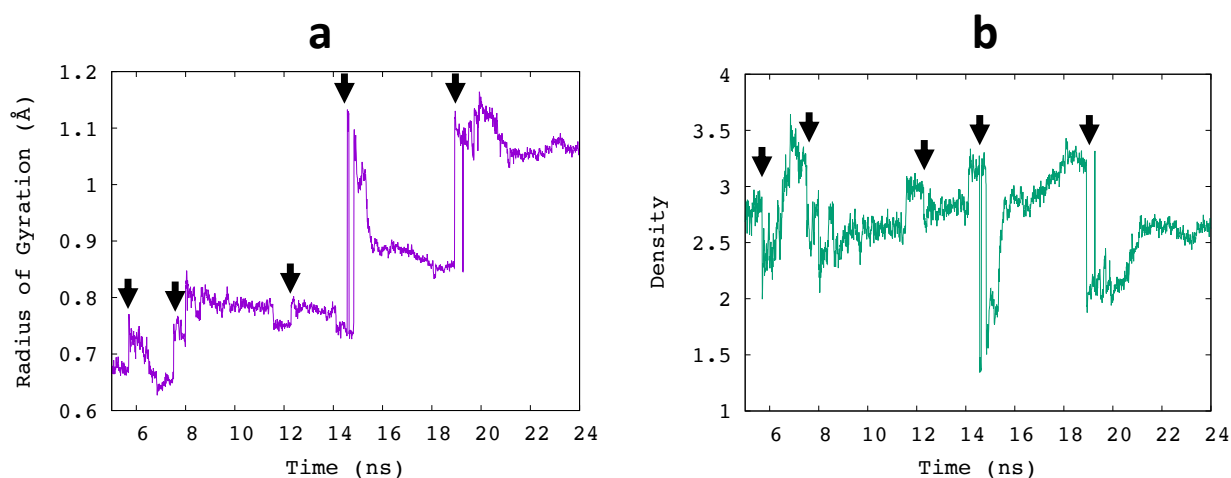
Furthermore, as can be seen, in Figure 10a, the direct pathway is associated with a continuous growth of this unique entity. However, direct pathway is not favored, and NC go rather through configurations composed of 2-6 clusters.





**Figure 10: Characterization of the biggest entity in the spath zpath space. The biggest entity corresponds to the entity with the largest number of ions. Each square of data corresponds to an average value calculated on the biggest entity of all molecular dynamic configurations corresponding to the area. (a) Average number of ions (b) Average gyration radius (c) Average density assuming a spherical shape with gyration radius (d) Average Ca/P ratio.**

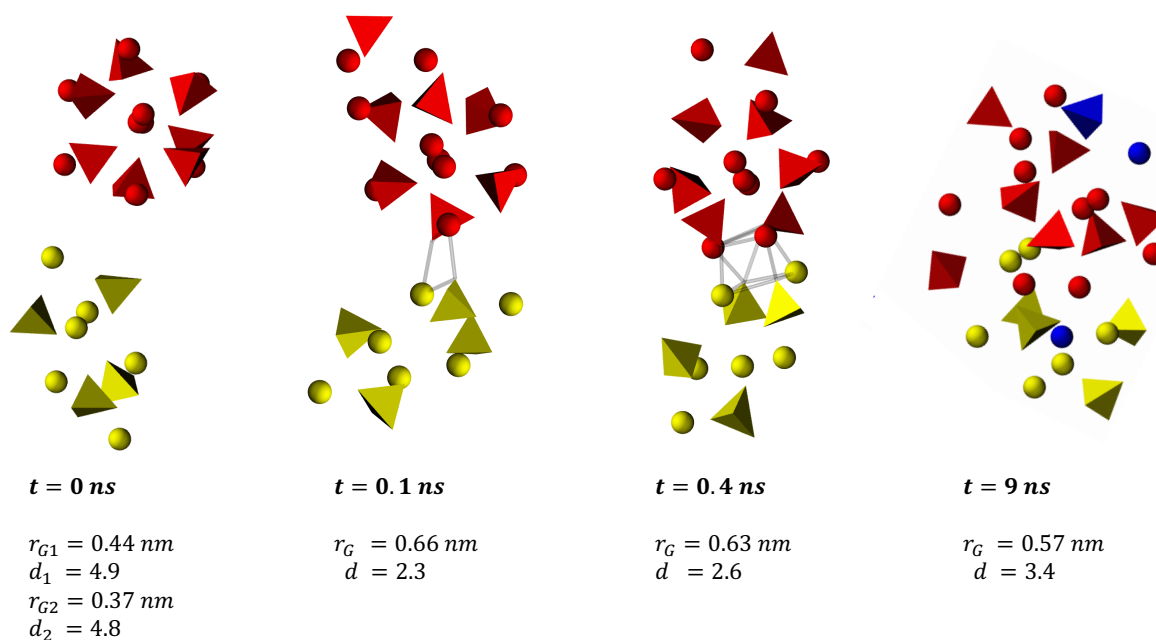
As can be seen in Figure 10d, the radius of gyration does not increase monotonously along NC but rather fluctuate from higher to lower values. This behavior is due to a cluster aggregation mechanism. Indeed, when two clusters aggregates, the size of the particle increase brutally. This assumption is verified in Figure 11a where the radius of gyration of a unique growing entity is plot with respect to the time.



**Figure 11 : Evolution of radius of gyration and density of a unique growing entity against time. Black arrows correspond to an aggregation between two clusters. Calculation of density is done according to equations 7-9**

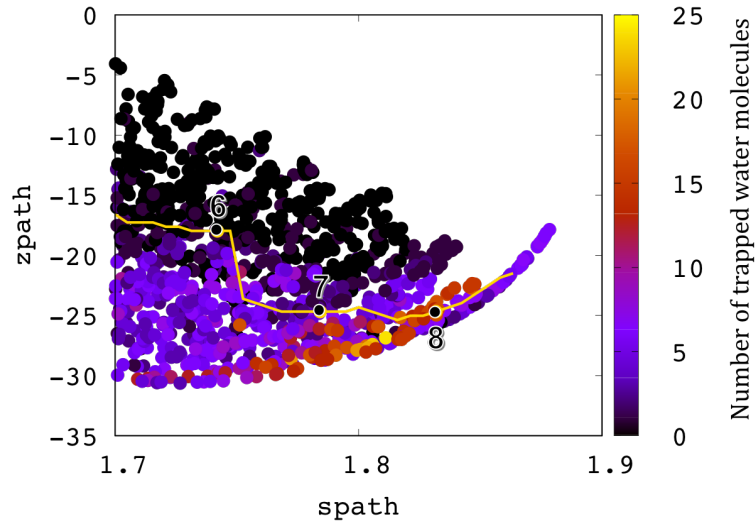
Black arrows correspond to a brutal augmentation of the radius of gyration due to a contact between two clusters.

Furthermore, according to Figure 10c, clusters are subjected to density fluctuations along the nucleation pathway. In particular, Figure 11b shows that the aggregation of two clusters (labelled with black arrow) is associated with a drastic fall of the calculated density followed by a progressive raise. These observations are rather due to a failure of our cluster density calculation (see equation 9) where we have made the assumption of a spherical shape for clusters. In our trajectory observation, clusters' shapes are often far away from a sphere, especially when two clusters aggregate. An aggregation between two clusters is described in Figure 12. As expected, a contact between two clusters cause a brutal raise of the gyration radius and a fall of the calculated density. After 9 ns the density increase and the radius of gyration decrease because of a rearrangement between the two clusters.



*Figure 12 : Example of an aggregation between two prenucleation clusters*

Point 8 in the free energy diagram (see Figure 5a,b) corresponds to the amorphous nanoparticle model that we have built. As shown, in Figure 13, going from point 7 to point 8 causes some water molecules to be trapped inside the cluster. This water effect is probably contributing to the free energy barrier between point 7 and point 8.



*Figure 13: Number of water molecules trapped inside a calcium phosphate entity. According to our observations, a water molecule is trapped inside a cluster when its coordination number is above 4.*

This result suggests also that more work needs to be performed on the nanoparticle model in order to find a more stable final state. Indeed, it would be interesting in the future to investigate different phases for the nanoparticle construction. In particular, OCP phase of brut formula  $\text{Ca}_8\text{H}_2(\text{PO}_4)_6 \cdot 5\text{H}_2\text{O}$  which is known to be a stable intermediate of the hydroxyapatite phase. Simulations from a supersaturated solution to an OCP-like nanoparticle is currently under investigation. This simulation will have more realistic pH conditions and will allow the study of the role of  $\text{HPO}_4^{2-}$  ions on the stabilization of some prenucleation complexes and clusters.

## **Conclusion**

Calcium phosphate nucleation was successfully studied by the *path-PIV* approach proving it a valuable tool to study such phase transition. As suggested in literature, a non-classical pathway was preferred for nucleation:

- (i) A spontaneous amorphization is observed in our nanoparticle model
- (ii) We have described a mechanism where clusters are able to aggregate between them and rearrange.
- (iii) This aggregation mechanism generates elongated shapes which are far away from a sphere.
- (iv) Clusters and complexes are calcium deficient and absorb calcium ions during the nucleation.
- (v) Some water molecules get trapped inside the structure of the cluster.

This method opens the way to the study of key parameters that can influence calcium phosphate nucleation. Indeed, it will be possible to study pH and temperature effect in the nucleation process. Furthermore, changing the target phase in the nanoparticle model will allow to shed some light on the calcium phosphate preferential phase at the early stage of nucleation. Finally, the method could also be adapted to study phase transitions between crystalline phases and reconstruct the phase diagram of calcium phosphates.

It would be also interesting to study the influence of other molecules on the nucleation free energy barriers. The method should be suitable for the study of heterogeneous nucleation as well. In particular, it will be interesting to study the nucleation in presence of a surface composed of collagen proteins.

## **Acknowledgements**

This work was performed using HPC resources from GENCI- [CCRT/CINES/IDRIS] (Grant 2015-[x2015082022]) and the CCRE of Université Pierre et Marie Curie.

## **References**

1. Zapanta LeGeros, R. Apatites in biological systems. *Prog. Cryst. Growth Charact.* **4**, 1–45 (1981).
2. Wang, K. *et al.* Theoretical analysis of protein effects on calcium phosphate precipitation in simulated body fluid. *CrystEngComm* **14**, 5870–5878 (2012).
3. Zawacki, S. J., Nancollas, G. H., Lore, M., Richardson, C. & Perez, L. Mineral phases of calcium phosphate. *Anat. Rec.* **224**, 234–241 (2005).
4. Amaral, M., Lopes, M. A., Silva, R. F. & Santos, J. D. Densification route and mechanical properties of Si<sub>3</sub>N<sub>4</sub>-bioglass biocomposites. *Biomaterials* **23**, 857–62 (2002).
5. Bazin, D. & Daudon, M. Nouvelles Méthodes d'étude des calculs et plaques de Randall. *Actual. Néphrologiques Jhamburg. Maison Chim.* 75–98 (2010).
6. Robison, R. The Possible Significance of Hexosephosphoric Esters in Ossification. *Biochem. J.* **17**, 286–93 (1923).
7. Robison, R. & Soames, K. M. The Possible Significance of Hexosephosphoric Esters in Ossification. *Biochem. J.* **18**, 740–754 (1924).
8. Genge, B. R., Wu, L. N. & Wuthier, R. E. Differential fractionation of matrix vesicle proteins. Further characterization of the acidic phospholipid-dependent Ca<sup>2+</sup>-binding proteins. *J. Biol. Chem.* **265**, 4703–10 (1990).
9. ROBINSON, R. A. An electron-microscopic study of the crystalline inorganic component of bone and its relationship to the organic matrix. *J. Bone Joint Surg. Am.* **34-A**, 389–435; passim (1952).
10. Robinson, R. A. & Watson, M. L. Collagen-crystal relationships in bone as seen in the electron microscope. *Anat. Rec.* **114**, 383–409 (1952).
11. GLIMCHER, M. J. Molecular Biology of Mineralized Tissues with Particular Reference to Bone. *Rev. Mod. Phys.* **31**, 359–393 (1959).

12. Freeman Rutgers, J. W., Landis, W. J., Silver, F. H. & Freeman, J. W. Collagen as a scaffold for biomimetic mineralization of vertebrate tissues. *Artic. J. Mater. Chem.* (2006). doi:10.1039/b505706j
13. Wang, Y. *et al.* The predominant role of collagen in the nucleation, growth, structure and orientation of bone apatite. *Nat. Mater.* **11**, 724–733 (2012).
14. Glimcher, M. J. Mechanism of calcification: Role of collagen fibrils and collagen-phosphoprotein complexes in vitro and in vivo. *Anat. Rec.* **224**, 139–153 (1989).
15. Glimcher, M. J. Bone: Nature of the Calcium Phosphate Crystals and Cellular, Structural, and Physical Chemical Mechanisms in Their Formation. *Rev. Mineral. Geochemistry* **64**, 223–282 (2006).
16. Gibbs, J. W. A. du texte. The collected works / of J. W. Gibbs,... (1928).
17. Pouget, E. M. *et al.* CaCO<sub>3</sub> Formation Revealed by Cryo-TEM. *Science (80-. )*. **323**, 1455–1458 (2009).
18. Dey, A. *et al.* The role of prenucleation clusters in surface-induced calcium phosphate crystallization. *Nat. Mater.* **9**, 1010–1014 (2010).
19. Gebauer, D., Kellermeier, M., Gale, J. D., Bergström, L. & Cölfen, H. Pre-nucleation clusters as solute precursors in crystallisation. *Chem. Soc. Rev.* **43**, 2348–2371 (2014).
20. Beniash, E., Metzler, R. A., Lam, R. S. K. & Gilbert, P. U. P. A. Transient amorphous calcium phosphate in forming enamel. *J. Struct. Biol.* **166**, 133–143 (2009).
21. Combes, C. & Rey, C. Amorphous calcium phosphates: Synthesis, properties and uses in biomaterials. *Acta Biomater.* **6**, 3362–3378
22. Habraken, W. J. E. M. *et al.* Ion-association complexes unite classical and non-classical theories for the biomimetic nucleation of calcium phosphate. *Nat. Commun.* **4**, 1507–1512 (2013).
23. Mancardi, G., Terranova, U. & de Leeuw, N. H. Calcium Phosphate Prenucleation Complexes in Water by Means of ab Initio Molecular Dynamics Simulations. *Cryst. Growth Des.* **16**, 3353–3358 (2016).
24. Betts, F. & Posner, A. S. An X-ray radial distribution study of amorphous calcium phosphate. *Mater. Res. Bull.* **9**, 353–360 (1974).
25. Mancardi, G., Hernandez Tamargo, C. E., Di Tommaso, D. & de Leeuw, N. H. Detection of Posner's clusters during calcium phosphate nucleation: a molecular dynamics study. *J. Mater. Chem. B* **5**, 7274–7284 (2017).
26. Lin, T.-J. Force Field Parameters and Atomistic Surface Models for Hydroxyapatite and Analysis of Biomolecular Adsorption At Aqueous Interfaces. (2013).

27. KAY, M. I., YOUNG, R. A. & POSNER, A. S. Crystal Structure of Hydroxyapatite. *Nature* **204**, 1050–1052 (1964).
28. Barducci, A., Bonomi, M. & Parrinello, M. Metadynamics. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **1**, 826–843 (2011).
29. Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C. & Bussi, G. PLUMED 2: New feathers for an old bird. (2013). doi:10.1016/j.cpc.2013.09.018
30. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926 (1998).
31. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
32. Essmann, U. *et al.* A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577 (1995).
33. Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **38**, 3098–3100 (1988).
34. Lee, C., Yang, W. & Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **37**, 785–789 (1988).
35. Miehlich, B., Savin, A., Stoll, H. & Preuss, H. Results obtained with the correlation energy density functionals of Becke and Lee, Yang and Parr. *Chem. Phys. Lett.* **157**, 200–206 (1989).
36. Lippert, G., Hutter, J. & Parrinello, M. The Gaussian and augmented-plane-wave density functional method for ab initio molecular dynamics simulations. *Theor. Chem. Accounts Theory, Comput. Model. (Theoretica Chim. Acta)* **103**, 124–140 (1999).
37. VandeVondele, J. *et al.* Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Comput. Phys. Commun.* **167**, 103–128 (2005).
38. Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate *ab initio* parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **132**, 154104 (2010).
39. Goedecker, Teter & Hutter. Separable dual-space Gaussian pseudopotentials. *Phys. Rev. B. Condens. Matter* **54**, 1703–1710 (1996).
40. Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **52**, 255–268 (1984).

41. Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **31**, 1695–1697 (1985).
42. Tzu-JenTzu-Jen, & Heinz, H. (2015). Accurate Force Field Parameters and pH Resolved Surface Models for Hydroxyapatite to Understand Structure, Mechanics, Hydration, and B. I. <http://doi.org/10.1021/acs.jpcc.5b12504>. & Heinz, H. Accurate Force Field Parameters and pH Resolved Surface Models for Hydroxyapatite to Understand Structure, Mechanics, Hydration, and Biological Interfaces. (2015). doi:10.1021/acs.jpcc.5b12504
43. Yashima, M., Sakai, A., Kamiyama, T. & Hoshikawa, A. Crystal structure analysis of  $\beta$ -tricalcium phosphate  $\text{Ca}_3(\text{PO}_4)_2$  by neutron powder diffraction. *J. Solid State Chem.* **175**, 272–277 (2003).
44. Larsson, L. & Ohman, S. Serum ionized calcium and corrected total calcium in borderline hyperparathyroidism. *Clin. Chem.* **24**, (1978).
45. Gallet, G. A. & Pietrucci, F. Structural cluster analysis of chemical reactions in solution. *J. Chem. Phys.* **139**, 074101 (2013).
46. Pipolo, S. *et al.* Navigating at Will on the Water Phase Diagram. (2017). doi:10.1103/PhysRevLett.119.245701
47. Branduardi, D., Gervasio, F. L. & Parrinello, M. From A to B in free energy space. *J. Chem. Phys.* **126**, 054103 (2007).
48. Pipolo, S. *et al.* Navigating at Will on the Water Phase Diagram. (2017).
49. Paolo Raiteri, \*,‡, Alessandro Laio, ‡, Francesco Luigi Gervasio, ‡, Cristian Micheletti, § and & Parrinello‡, M. Efficient Reconstruction of Complex Free Energy Landscapes by Multiple Walkers Metadynamics†. (2005). doi:10.1021/JP054359R
50. Habraken, W. J. E. M. *et al.* Ion-association complexes unite classical and non-classical theories for the biomimetic nucleation of calcium phosphate. *Nat. Commun.* **4**, 1507–1512 (2013).



## - Conclusion Générale -

Au cours de ce Doctorat, nous avons répondu à des problématiques liées à la biominéralisation et aux calcifications pathologiques, en les abordant sous l'angle de la modélisation chimique. Ces problématiques sont de véritables challenges pour le chimiste théoricien car elles sont associées à une complexité particulière en termes de chimie et de physico-chimie. Ainsi, l'oxalate de calcium qui constitue la majorité des calculs rénaux (une pathologie majeure actuellement), se forme dans un environnement biologique complexe impliquant des interactions avec un solvant (l'eau), des diverses molécules organiques et d'autres ions.

Ainsi, nous avons montré dans le **chapitre 2** que l'eau solvant jouait un rôle essentiel dans la stabilisation de certaines surfaces de l'oxalate de calcium, expliquant ainsi les morphologies cristallines observées dans les calculs rénaux (*in vivo*) et dans les synthèses (*in vitro*).

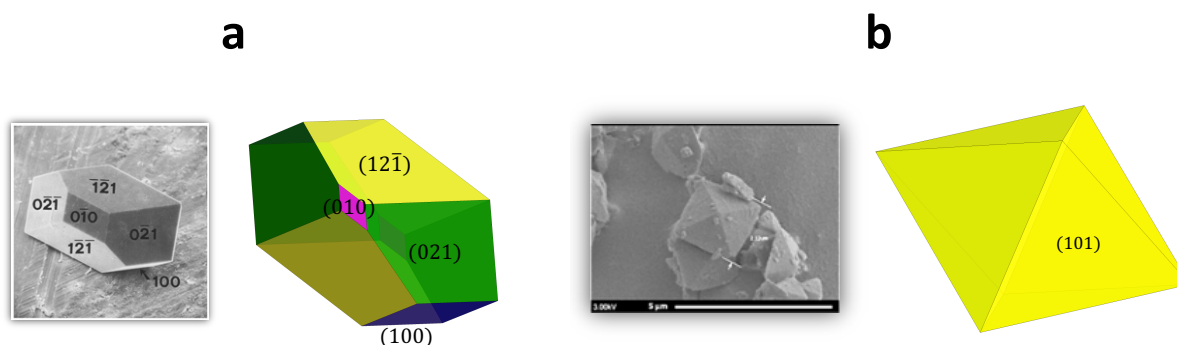


Figure 1 : prédiction des morphologies cristallines et images MEB<sup>1,2</sup> (a) du COM (Calcium Oxalate Monohydrated,  $\text{CaC}_2\text{O}_4\cdot\text{H}_2\text{O}$ , whewellite) (b) et du COD (Calcium Oxalate Dihydrated,  $\text{CaC}_2\text{O}_4\cdot 2\text{H}_2\text{O}$ , weddellite).

Ces résultats ont été obtenus par le calcul DFT des énergies de surfaces du COM (Calcium Oxalate Monohydrated,  $\text{CaC}_2\text{O}_4\cdot\text{H}_2\text{O}$ , whewellite) et du COD (Calcium Oxalate Dihydrated,  $\text{CaC}_2\text{O}_4\cdot 2\text{H}_2\text{O}$ , weddellite) en présence ou non d'un modèle de solvation (bicouche de molécules d'eau). Nous avons ainsi confirmé la prévalence des surfaces (100), (010), ( $12\bar{1}$ ) et (021) pour le COM et de la surface (101) pour le COD.

La prise en compte de l'énergie de solvation des surfaces a été essentielle pour prédire les morphologies observées lors des synthèses *in vitro*. En revanche, la prise en compte de la solvation n'a pas été suffisante. Les prédictions de morphologie du COM (Figure 1a) et du COD (Figure 1b) n'ont été rendues possibles que par l'"extinction" des surfaces (001) du COM et (110) du COD suggérant que celles-ci sont sujettes à des effets cinétiques.

Dans les synthèses, le COM présente deux autres morphologies où une seule des deux surface (021) ou (12 $\bar{1}$ ) est exprimée. En faisant varier les énergies de surfaces dans notre fenêtre d'erreur, il a été possible de prédire ce comportement expliquant alors l'observation de ces morphologies.

Au cours du **chapitre 3**, nous avons donné une base théorique aux résultats expérimentaux de l'effet du thé vert sur la cristallisation des oxalates de calcium. Dans la littérature, une molécule particulière présente dans le thé vert (l'épigallocatechine gallate, EGCG) est en effet suspectée d'influencer la croissance cristalline en orientant la cristallisation vers le COD (au lieu du COM), et ce, en promouvant la surface (010) du COD aux dépens de la surface (101).

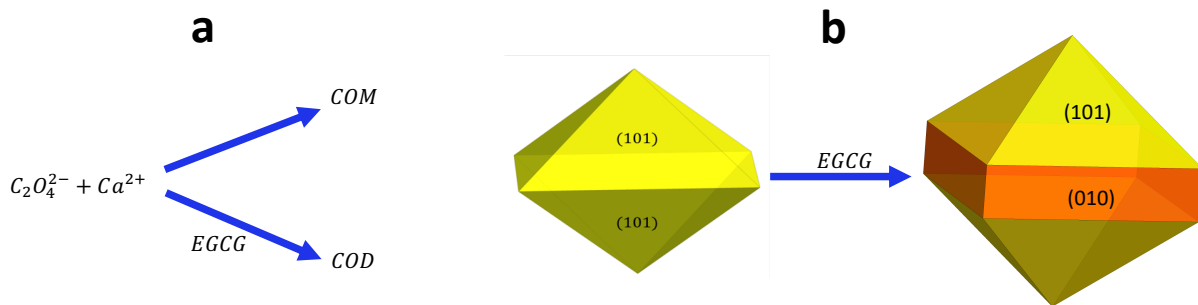


Figure 2 : Effets suspectés de l'EGCG sur la cristallisation de l'oxalate de calcium. (a) sélectivité vers la phase COD aux dépens de la phase COM (b) stabilisation de la face (010) du COD

Par des calculs DM/champs de force couplés avec des méthodes améliorant l'échantillonnage des configurations (Replica Exchange Molecular Dynamics et Metadynamics) nous avons montré que l'EGCG interagit spécifiquement avec les surfaces (001) et (010) du COM et la surface (010) du COD. Dans le cas du COM, l'interaction de l'EGCG avec des surfaces orthogonales expliquerait pourquoi la molécule empêche sa croissance cristalline. Dans le cas du COD, la stabilisation des faces équivalentes (010) et (100) (sans pour autant inhiber la croissance de la face (101)), peut expliquer l'apparition des faces (010) et (100) dans les synthèses *in vitro* en présence de thé vert. Le chapitre 3 a permis la mise en place d'une méthodologie de calculs ouvrant la voie à l'étude d'autres molécules organiques, pouvant présenter un intérêt thérapeutique dans la pathologie des calculs rénaux. La prochaine étape sera d'automatiser complètement la procédure afin de proposer un logiciel accessible permettant des études de drug design en amont des essais médicaux.

Enfin, dans le **chapitre 4**, nous avons étudié les premiers instants de la nucléation des phosphates de calcium par la métadynamique et plus particulièrement par la méthode *path-PIV*.

Ainsi, nous avons été capables de reconstruire le chemin d'énergie libre correspondant à la transition entre une solution supersaturée et une nanoparticule amorphe de phosphate de calcium. Nous avons montré, que dans les conditions de simulation choisies, des complexes de prénucléation se forment spontanément.

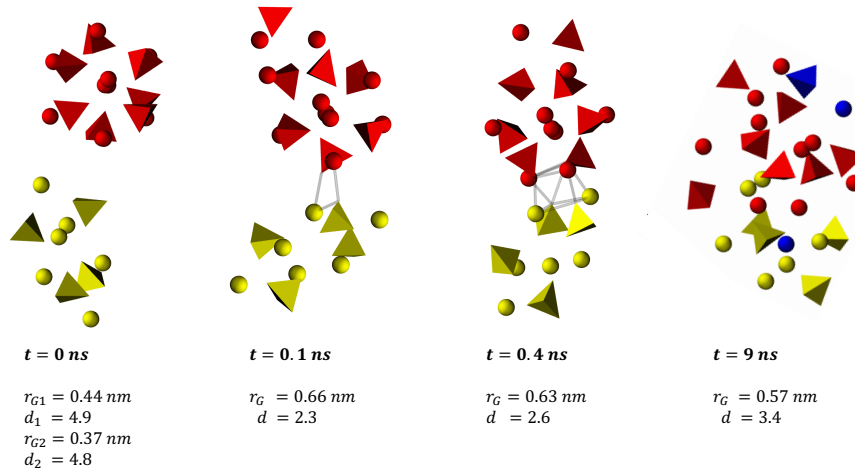


Figure 3 : Exemple de l'agrégation entre deux clusters, observée au cours d'une DM.  $r_G$  correspond au rayon de giration du cluster et  $d$  à sa densité (en assumant une forme sphérique)

Nous avons également mis en évidence plusieurs points montrant que la nucléation des phosphates de calcium s'écarte de la théorie classique de la nucléation :

- (i) Une amorphisation spontanée de notre modèle de nanoparticule est observée lors de son équilibration en DM.
- (ii) La croissance des clusters ne se fait pas seulement par addition d'ions ou de complexes de prénucléation mais également par l'agrégation de clusters entre eux (Figure 3).
- (iii) Ce mécanisme génère des clusters de forme allongée, loin de la forme sphérique prévue par la théorie classique de la nucléation
- (iv) des clusters chargés négativement (car déficients en calcium) se forment et un processus d'adsorption des ions calciums par les clusters s'effectue au cours de la nucléation
- (v) pendant le processus de croissance, des molécules d'eau se retrouvent piégée à l'intérieur du cluster

Les travaux effectués dans le **chapitre 4** ouvrent la voie à l'étude des paramètres clés pouvant influencer la nucléation des phosphates de calcium, comme la température, le pH, la supersaturation, ou l'effet de molécules supplémentaires présente dans le milieu réactionnel. Une prochaine étape consistera à simuler le processus de nucléation dans des conditions de pH plus réalistes. Nous pourrions ainsi étudier l'effet des ions  $\text{HPO}_4^{2-}$  sur les barrières d'énergie

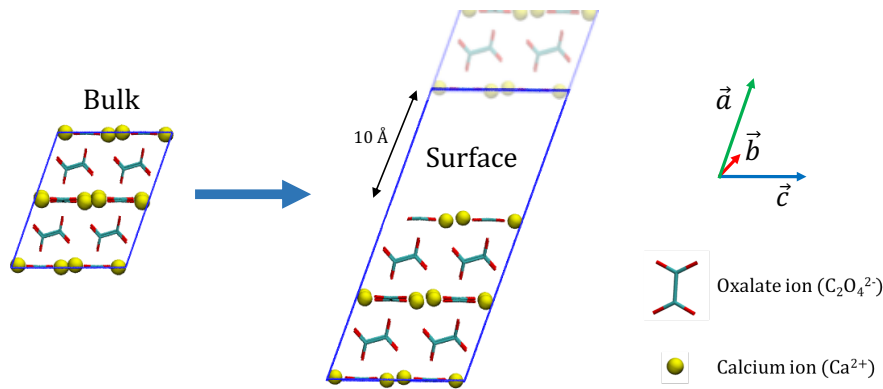
libre de la nucléation. Enfin, il sera également intéressant d'étudier la nucléation hétérogène pour comprendre l'effet que peuvent avoir certaines surfaces sur le processus, et en particulier l'effet du collagène connu comme étant un déclencheur de la nucléation dans l'os. Il sera pour cela important de construire un modèle représentatif de la zone de «gap» formée par le collagène et décrite dans la littérature comme étant le lieu de la nucléation des phosphates de calcium dans l'os<sup>3</sup>.

1. Leroy, C. Oxalates de calcium et hydroxyapatite : des matériaux synthétiques et naturels étudiés par techniques RMN et DNP. (Université P. et M. Curie, 2016).
2. Franchini-Angela, M. & Aquilano, D. Theoretical growth morphology of whewellite  $\text{CaC}_2\text{O}_4 \cdot \text{H}_2\text{O}$ . *Phys. Chem. Miner.* **10**, 114–120 (1984).
3. Weiner, S. & Wagner, H. D. THE MATERIAL BONE: Structure-Mechanical Function Relations. *Annu. Rev. Mater. Sci.* **28**, 271–298 (2002).

## - Résumé -

Cette thèse s'articule autour de la thématique de la **biominéralisation**, le processus par lequel les êtres vivants parviennent à orienter la précipitation de minéraux présents dans leur environnement pour confectionner des matériaux utiles à la survie de l'espèce considérée. La biominéralisation est un phénomène extrêmement complexe mettant en jeu des régulations génétiques et l'implication de protéines permettant d'enclencher la **minéralisation**, à des localisations bien précises. Il arrive cependant que celle-ci échappe au contrôle de l'organisme vivant conduisant alors à diverses pathologies. C'est le cas par exemple, quand des calcifications appelées **calculs rénaux**, se forment à l'intérieur des reins. A partir d'une certaine taille, un calcul rénal devient pathologique et nécessite une prise en charge hospitalière. La pathologie est communément appelée **lithiase urinaire**. Il est établi que les oxalates de calcium, et plus particulièrement les phases monohydratée (Calcium Oxalate Monohydrate, COM ou whewellite) et dihydratée (Calcium Oxalate Dihydrate, COD ou weddellite), composent la majorité des calculs rénaux<sup>1,2</sup>. L'environnement, et tout particulièrement l'alimentation, est un facteur clé influençant la formation et la composition des calculs rénaux. Actuellement, les traitements consistent principalement à une injection d'antalgiques et à l'éjection du calcul par les voies naturelles. Si nécessaire, une intervention chirurgicale est mise en oeuvre. Les médecins sont à la recherche de médicaments/traitements potentiellement capables de prévenir la formation des calculs rénaux.

Dans cette thèse, nous avons choisi d'aborder cette problématique par le biais de la chimie théorique et plus particulièrement de la **dynamique moléculaire**, une méthode prédictive d'importance permettant d'étudier le comportement d'un modèle atomique au cours du temps. Dans le **chapitre 2**, nous avons étudié les faces cristallines du COM et du COD dans l'objectif de construire des modèles moléculaires cohérents capables de décrire l'interface des cristaux d'oxalates de calcium avec l'eau considérée comme solvant. A partir des structures cristallographiques du COM et du COD résolues par diffraction des neutrons et des rayons-X<sup>3,4</sup>, nous avons généré de nombreux modèles de surfaces, en insérant un espace vide le long des plans d'indice de Miller variés, comme schématisé sur la Figure 1 pour la surface (100) du COM (à titre d'exemple). Au total, plus d'une dizaine de modèles de surfaces, avec des indices de Miller variés, ont été construits pour le COM et le COD.

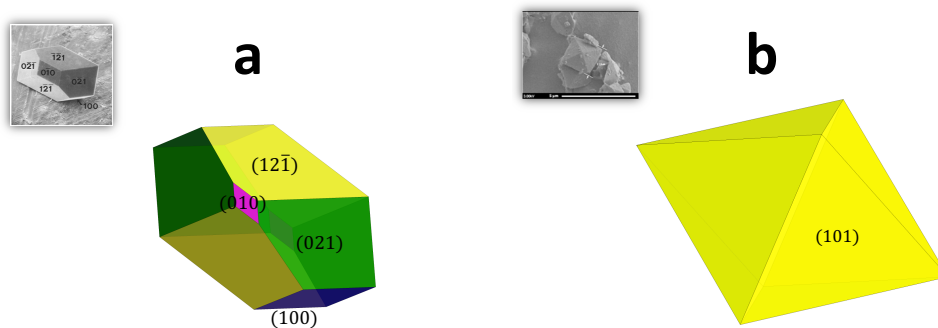


**Figure 1 : construction de la surface (100) du COM. Un espace de vide est situé entre deux surfaces consécutives.**

Nous avons ensuite calculé les énergies de surface des différents modèles par une approche DFT (PBE<sup>5</sup>) comme décrit dans l'équations ( 1 ).  $\gamma_{vide}$  correspond à l'énergie de surface dans le vide, sans prendre en compte la solvataion de l'eau.  $\gamma_{H_2O}$  correspond à l'énergie de surface hydratée par une bicouche de molécules d'eau.

$$\gamma_{vide} = \frac{E_{surface} - E_{bulk}}{2A} \quad \gamma_{H_2O} = \frac{E_{surface,H_2O} - E_{bulk} - N \cdot E_{H_2O}}{2A} \quad (1)$$

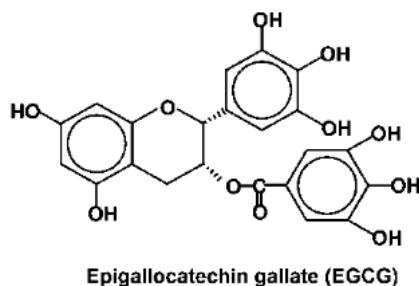
Nous avons montré qu'un modèle de solvataion est essentiel afin de prédire les morphologies cristallines observées expérimentalement. Néanmoins, ce modèle élaboré n'a pas été suffisant pour expliquer l'absence de certaines faces. C'est le cas pour la face (001) du COM et pour la face (110) du COD : leur expression a dû être supprimée afin d'obtenir les morphologies cristallines présentées sur la Figure 2. Ce résultat suggère que la croissance cristalline du COM (et du COD) est soumise à des effets cinétiques qui pourraient expliquer l'absence de telles faces.



**Figure 2 : prédiction des morphologies cristallines du COM (image MEB provenant de Franchini-Angela et coll.<sup>6</sup>) (a) et du COD (image MEB provenant de Leroy<sup>7</sup>) (b).**

En faisant varier les énergies de surfaces du COM dans la fenêtre d'erreur, nous avons été capables de prédire toutes les morphologies cristallines observées dans les synthèses en milieu aqueux, rendant ainsi compte de la diversité morphologique de cette phase particulière.

Dans le **chapitre 3**, nous avons étudié l'interaction des surfaces dominantes du COM et du COD avec une molécule du thé vert suspectée d'avoir un intérêt thérapeutique dans le cadre de la lithiase urinaire<sup>8</sup> : l'épigallocatechine gallate (EGCG) (voir Figure 3).



*Figure 3 : épigallocatechin (EGC) and (-)-épigallocatechin-3-gallate (EGCG)*

L'impact du thé vert sur la cristallisation du COM et du COD a été étudié par Chen *et coll.*<sup>9</sup> et Leroy<sup>7</sup>. Il en ressort que : **(i)** le thé vert est capable d'orienter sélectivement la cristallisation vers la phase COD, aux dépens du COM, **(ii)** en présence de thé vert, la face (010) du COD est bien en évidence, **(iii)** le thé vert est responsable de la formation de cavités sur la face (101) du COD. L'EGCG est l'un des composants du thé vert et est fortement suspectée d'interagir directement avec le COM et le COD. L'étude de l'EGCG avec ces deux phases par dynamique moléculaire a été menée dans l'objectif d'apporter des arguments supplémentaires sur le rôle spécifique de l'EGCG.

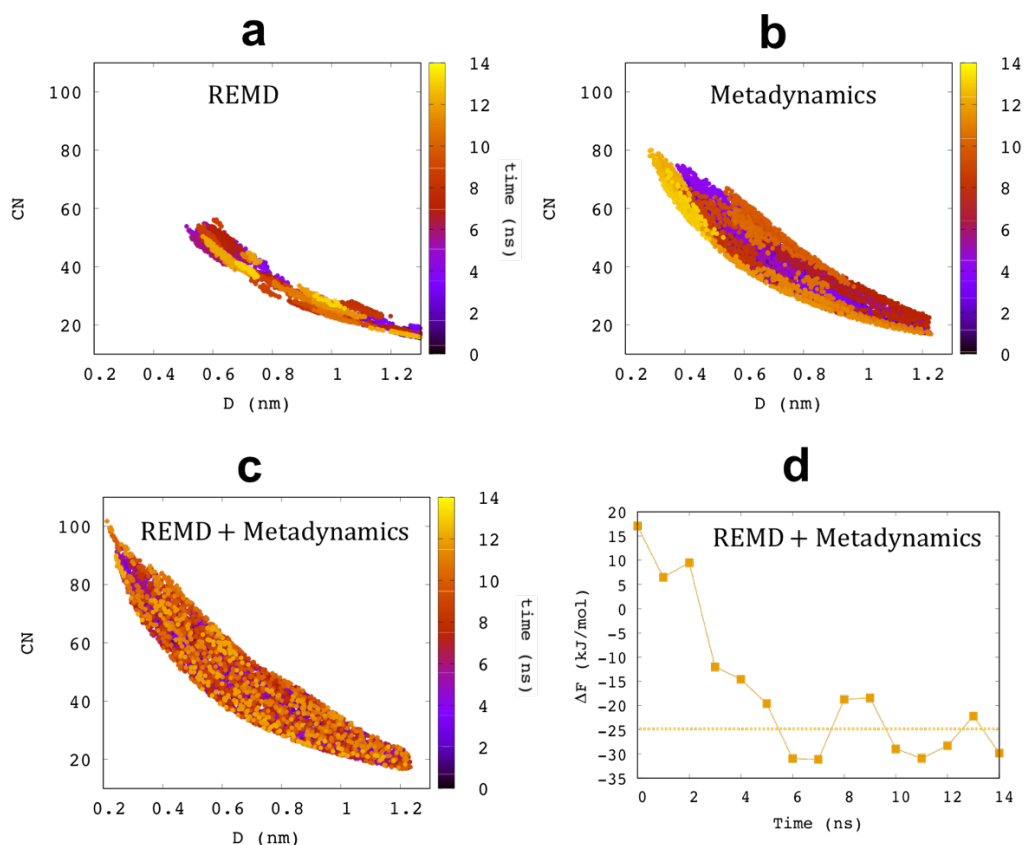
Une approche en champ de forces nous a permis de construire des modèles plus réalistes que dans le **chapitre 2**, avec une surface suffisamment grande afin d'éviter les images périodiques de l'EGCG et en mettant en œuvre explicitement un modèle de solvatation rejoignant la densité expérimentale de l'eau.

La simulation de la molécule d'EGCG solvatée « au-dessus » de la surface par une dynamique moléculaire « standard » ne permet pas d'observer le phénomène d'adsorption de manière ergodique. C'est pourquoi nous avons utilisé des techniques permettant d'améliorer l'échantillonnage des configurations. Pour cela, nous avons comparé trois approches sur la surface (100) du COM :

- (i) Dans la dynamique moléculaire par échange de réplique (Replica Exchange Molecular Dynamics, **REMD**)<sup>10</sup>, N DM sont lancées à N températures différentes. Chacune des simulations est appelée réplique. Tous les x étapes de la DM, deux répliques juxtaposées ont une certaine probabilité d'échanger leurs positions selon un critère de Metropolis-Hastings, dépendant de leur différence d'énergie potentielle. Cela permet d'explorer un plus grand nombre de configurations, en dépassant certaines barrières énergétiques grâce à l'énergie thermique.
- (ii) La **métadynamique**<sup>11</sup> constitue une stratégie complémentaire. Dans cette méthode, un biais de forme gaussienne est déposé à intervalle de temps régulier, le long d'une coordonnée appelée **variable collective** (Collective Variable, CV) correspondant à la transformation voulant être étudiée. Dans le cas de l'étude de l'adsorption de l'EGCG sur une surface d'oxalate de calcium, nous avons choisi deux CV : la distance entre le centre de masse de l'EGCG et la surface de l'oxalate de calcium considérée (**D**) et le nombre de coordination entre les atomes d'oxygène de l'EGCG et les atomes de surface de l'oxalate de calcium (**CN**). La métadynamique permet alors d'améliorer l'échantillonnage des configurations le long de la CV en « décourageant » le système à retourner vers les configurations déjà visitées. D'autre part, la métadynamique permet de reconstruire l'énergie libre associée à la CV.
- (iii) La combinaison des deux techniques de REMD et de métadynamique (**REMD-MD**).

Sur la Figure 4, nous avons représenté l'échantillonnage des configurations obtenues par ces trois techniques. La métadynamique (Figure 4b) a permis un meilleur échantillonnage des configurations que la REMD prise isolément (Figure 4a). La combinaison de la REMD avec la métadynamique a permis une convergence plus rapide de l'énergie libre. Comme montré en Figure 4c, la convergence de l'énergie libre permet une dispersion optimale de l'échantillonnage des configurations. En effet, quand tous les puits d'énergie libre sont remplis par les biais, le système évolue sans préférence entre toutes les configurations, dans l'espace des CV.





*Figure 4: Exploration des configurations dans l'espace 2D des CVs (D, CN). DM de l'EGCG aux abords de la surface (010) par les techniques de (a) REMD (b) métadynamique (c) REMD + métadynamique. (d) Convergence de  $\Delta F_{adsorption}$  par rapport au temps de simulation*

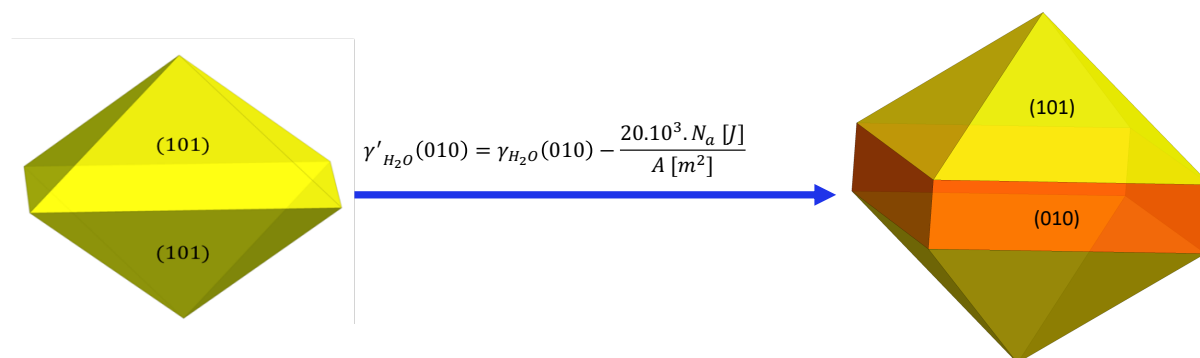
Nous avons alors appliqué cette dernière méthode à l'adsorption de l'EGCG sur les autres surfaces du COM et du COD. Les résultats des calculs d'énergies libres sont présentés dans le Tableau 1.

Structure (surface)	COM (010)	COM (100)	COM (12 $\bar{1}$ )	COD (010)	COD (101)
$\Delta F_{adsorption}$ (kJ.mol $^{-1}$ )	-25 ( $\pm 5$ )	-15 ( $\pm 5$ )	+10 ( $\pm 5$ )	-20 ( $\pm 5$ )	0 ( $\pm 5$ )

*Tableau 1 : Énergie libre d'adsorption de l'EGCG sur les surfaces du COM/COD.*

Dans le cas du COM, nous avons calculé des énergies libres d'adsorption négatives pour les surfaces (100) et (010). Ces surfaces sont décrite par Millan<sup>12</sup> comme étant prédominantes, en particulier lors des premiers instants de la cristallisation. Une interaction forte avec l'EGCG

expliquerait alors pourquoi le thé vert est capable d'inhiber sélectivement la croissance cristalline du COM, comme décrit dans les expériences<sup>7,9</sup>. Enfin, la stabilisation par l'EGCG de la surface (010) du COD expliquerait l'apparition de celle-ci lors de synthèses *in vitro* en présence de thé vert. En effet, comme présenté sur la Figure 5, une stabilisation de  $-20\text{kJ}\cdot\text{mol}^{-1}$  et rapportée à la surface du modèle est suffisante pour provoquer l'apparition de la face (010).



*Figure 5 : Effet de la stabilisation anisotrope de l'EGCG sur la morphologie du COD.*

Enfin, l'observation expérimentale de l'effet du thé vert sur la formation de cavités sur la face (101) du COD n'a pas été expliquée par nos calculs. Dans cette étude, nous avons fixé volontairement les positions des atomes de surface afin de simplifier les simulations. Dans le futur, il faudra prendre en compte les possibles restructurations de surfaces induites par l'EGCG. Ces résultats sur l'EGCG sont prometteurs et ouvrent la voie à l'étude du pouvoir inhibiteur d'autres molécules pouvant présenter un intérêt thérapeutique, vis-à-vis de la croissance de l'oxalate de calcium.

Une prochaine étape sera d'automatiser complètement la procédure afin de proposer une méthode simple et efficace permettant une approche de *drug screening*. Pour cela, la convergence de l'énergie libre peut être encore améliorée et un travail doit être fait sur le choix des CVs décrivant le phénomène d'adsorption. En effet, les CVs choisies devront être capables de distinguer de manière efficace les configurations solvatées et adsorbées pour tous les types de surfaces et de molécules étudiées.

Dans le **chapitre 4**, nous nous sommes intéressés aux premiers instants de la cristallisation, *i.e.* la nucléation, dans le cadre des phosphates de calcium. Nous aurions pu continuer à étudier l'oxalate de calcium mais nous avons préféré nous intéresser à ces composés pour plusieurs raisons : **(i)** la nucléation des phosphates de calcium est plus étudiée dans la littérature, **(ii)** il s'agit d'une thématique de recherche largement développée dans l'équipe SMILES du LCMCP

à laquelle j'appartiens, et (iii) nous avons pu bénéficier d'un champs de forces spécialement développé pour reproduire les données expérimentales de l'hydroxyapatite, et en particulier les énergies d'hydratation de surface. Néanmoins, la méthodologie développée dans ce **chapitre 4** sera applicable dans le futur aux oxalates de calcium, sans difficultés particulières.

Nous avons appliqué la méthode de la métadynamique avec une *CV path-PIV*<sup>13,14</sup> afin de simuler la transition d'une solution supersaturée de phosphate de calcium vers une nanoparticule amorphe. Nous avons ainsi été capables de reconstruire le diagramme d'énergie libre de la transition et de discuter des mécanismes impliqués. Selon la CNT (Classical Nucleation Theory), la croissance du cristal devient favorable lorsque sa taille est suffisamment grande pour dépasser une barrière d'énergie libre. Cette barrière dépend d'une part de l'énergie libre négative (favorable) liée à la formation du « bulk » ( $V\Delta g$ ), et d'autre part, de l'énergie libre positive (défavorable)  $S\sigma$ , liée à la formation d'une nouvelle interface liquide/solide<sup>15</sup>. L'énergie libre de « bulk » est proportionnelle au volume  $V$  du cluster, alors que l'énergie libre de surface est proportionnelle à la surface  $S$ . Ainsi, pour une taille faible ( $r \ll r_c$ ), le terme défavorable de l'énergie de surface est prépondérant tandis que le terme favorable d'énergie de bulk prend de l'importance avec l'augmentation de la taille du cluster. Le cluster doit alors dépasser une certaine taille critique pour que sa croissance devienne énergétiquement favorable. Dans ce chapitre, nous avons montré que le phosphate de calcium adopte des mécanismes de nucléations s'éloignant du chemin prévu par la CNT :

- Tout d'abord, le rayon de la particule ne constitue pas une bonne coordonnée pour étudier l'évolution de la nucléation principalement due à un mécanisme par agrégation de clusters de prénucléation.
- De plus, nous avons soulevé le fait qu'un tel mécanisme génère la formation d'agrégats non sphériques.
- Enfin, nous avons également montré que la formation de clusters déficients en calcium, implique un mécanisme d'absorption des ions calcium par les clusters, au cours de la nucléation.

De plus, les simulations effectuées ont permis de démontrer que dans des conditions dites supersaturées, une solution composée de complexes de prénucléation est favorisée par rapport à une solution composée d'ions libres solvatés. Ces complexes constituent les briques élémentaires nécessaires à la formation de clusters.

Ces résultats ouvrent la voie à l'étude des paramètres clés (température, supersaturation, molécules supplémentaires, etc.) pouvant influencer les barrières d'énergies impliquées dans la

nucléation. En particulier, il sera intéressant d'étudier l'effet de certaines molécules et/ou surfaces sur le déclenchement de la nucléation. Enfin, il sera nécessaire d'étudier des conditions plus proches du pH physiologique, impliquant en particulier la présence d'ions  $\text{HPO}_4^{3-}$  dans la solution.

1. Daudon, M. Épidémiologie actuelle de la lithiase rénale en France. *Ann. Urol. (Paris)*. **39**, 209–231 (2005).
2. Sorokin, I. *et al.* Epidemiology of stone disease across the world. *World J. Urol.* **35**, 1301–1320 (2017).
3. Colas, H. *et al.* Whewellite,  $\text{CaC}_2\text{O}_4 \cdot \text{H}_2\text{O}$ : structural study by a combined NMR, crystallography and modelling approach. *Crystengcomm* **15**, 8840–8847 (2013).
4. Tazzoli, V. & Domeneghetti, C. The crystal structures of whewellite and weddellite: re-examination and comparison. *Am. Mineral.* **65**, 27–334 (1980).
5. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
6. Franchini-Angela, M. & Aquilano, D. Theoretical growth morphology of whewellite  $\text{CaC}_2\text{O}_4 \cdot \text{H}_2\text{O}$ . *Phys. Chem. Miner.* **10**, 114–120 (1984).
7. Leroy, C. Oxalates de calcium et hydroxyapatite : des matériaux synthétiques et naturels étudiés par techniques RMN et DNP. (Université P. et M. Curie, 2016).
8. Benzerara, Y., Bazin, D., Daudon, M., Haymann, J.-P. & Rode, J. Catechin for treating Renal Lithiasis. (2015).
9. Chen, Z., Wang, C., Zhou, H., Sang, L. & Li, X. Modulation of calcium oxalate crystallization by commonly consumed green tea. *CrystEngComm* **12**, 845 (2010).
10. Rhee, Y. M. & Pande, V. S. Multiplexed-Replica Exchange Molecular Dynamics Method for Protein Folding Simulation. *Biophys. J.* **84**, 775–786 (2003).
11. Barducci, A., Bonomi, M. & Parrinello, M. Metadynamics. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **1**, 826–843 (2011).
12. Millan, A. Crystal Growth Shape of Whewellite Polymorphs: Influence of Structure Distortions on Crystal Shape. *Cryst. Growth Des.* **1**, 245–254 (2001).
13. Pipolo, S. *et al.* Navigating at Will on the Water Phase Diagram. (2017).
14. Gallet, G. A. & Pietrucci, F. Structural cluster analysis of chemical reactions in solution. *J. Chem. Phys.* **139**, 074101 (2013).
15. Gibbs, J. W. A. du texte. The collected works / of J. W. Gibbs,... (1928).