

# الجمهورية الجزائرية الديمقراطية الشعبية

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

Ministère de l'Enseignement Supérieur  
et de la Recherche Scientifique

Université Mouloud Mammeri  
FACULTE DE MEDECINE  
TIZI OUZOU



وزارة التعليم العالي والبحث العلمي  
جامعة مولود معمري  
كلية الطب  
قسم الصيدلة  
تيزي وزو

Département de Pharmacie  
N° D'ordre:

†.⊙∶∧∧.⊥ξ†∩∶∩∶∧.†∩∶∩∩∶∩

## MEMOIRE DE FIN D'ETUDES

Présenté sous forme d'article et soutenu publiquement

En vue de l'obtention du Diplôme de Docteur en Pharmacie

Le : 25/06/2024

*Sous le Thème*

Modélisation par homologie et Docking protéine-protéine d'une  
métalloprotéinase hémorragique CcMP-II

**Réalisé par:**

MOUSSOUNI Katia

SAHI Hayet

SI TAYEB Thiziri

**Encadré par:**

Dr BELAZOUGUI Ourdia

**Membres du jury:**

Pr N. HADHOUM

MCA en Chimie thérapeutique

Faculté de médecine

Présidente

UMMTO

Dr O. BELAZOUGUI

MAHU en Toxicologie

Faculté de médecine

Rapporteur

UMMTO

Dr A. MATMAR

Assistant en Toxicologie

CHU de Tizi-Ouzou

Examineur



République algérienne démocratique et populaire

الجمهورية الجزائرية الديمقراطية الشعبية

Ministère de l'enseignement Supérieur et de la recherche scientifique

Université Mouloud Mammeri Tizi-Ouzou

Faculté de Médecine

Département de pharmacie



#### Article :

Modélisation par homologie et Docking protéine-protéine d'une métalloprotéinase hémorragique CcMP-II

MOUSSOUNI Katia\*, SAHI Hayet\*, SI TAYEB Thiziri\*, BELAZOUGUI Ourdia

\*Correspondance : [katiamoussouni00@gmail.com](mailto:katiamoussouni00@gmail.com), [sahihayet309@gmail.com](mailto:sahihayet309@gmail.com),  
[thizirisita@gmail.com](mailto:thizirisita@gmail.com)

**Mots-clés :** prédiction, in silico, modélisation par Homologie, Docking moléculaire, HADDOCK.

**Keywords:** Prediction, in silico, homology modeling, molecular docking, HADDOCK.

**Résumé :** La biologie structurale étudie le lien entre la structure des molécules et leurs activités biologiques. La connaissance des structures des complexes protéine-protéine permet de mieux comprendre les mécanismes des interactions biomoléculaires. Cette étude vise à valider les prédictions du docking moléculaire et à démontrer son utilité dans l'élucidation d'un mécanisme de toxicité lié à une Snake venom metalloproteinase (SVMP), en comparant ces résultats avec ceux obtenus expérimentalement. Nous avons étudié l'interaction de *Cerastes cerastes* métalloprotéinase type II (CcMP-II), une métalloprotéinase hémorragique du venin de serpent *Cerastes cerastes*, avec le collagène de type IV et l'intégrine alpha-2 beta-1 en utilisant le logiciel de docking moléculaire « HADDOCK ». Une modélisation par homologie a été réalisée au préalable, car la structure tridimensionnelle de la protéine n'était pas disponible dans la Protein DATA BANK (PDB). L'évaluation des complexes dockés a révélé des interactions favorables entre les deux protéines et des valeurs énergétiques significatives, suggérant des liaisons potentiellement fonctionnelles. La concordance des résultats in silico avec les données expérimentales valide cette approche et encourage l'intégration synergétique des deux méthodes pour comprendre les interactions biologiques complexes.

**Abstract:** Structural biology focuses on the relationship between molecular structure and biological activities. Understanding the structures of protein-protein complexes enhances our comprehensions of biomolecular interaction mechanism. This study aims to validate predictions from molecular docking and demonstrate its utility in elucidating a mechanism of toxicity associated with a snake venom metalloproteinase (SVMP), comparing these results with experimental findings. We investigated the interaction of *Cerastes cerastes* metalloproteinase type II (CcMP-II), a hemorrhagic metalloproteinase from the venom of *Cerastes cerastes*, with collagen type IV and integrin alpha-2 beta-1 using the molecular docking software "HADDOCK." Homology modeling was conducted beforehand due to the protein's unavailability in the PDB. Evaluation of the docked complexes revealed favorable interactions between the two proteins and significant energy values, suggesting potentially functional bindings. The alignment of in silico results with experimental data validates this approach and encourages synergistic integration of both methods to understand complex biological interactions.

## 1. INTRODUCTION

Dans le domaine de la recherche en toxicologie, la nécessité de comprendre les interactions moléculaires complexes entre les agents toxiques et les organismes vivants est primordiale pour évaluer les risques sur la santé humaine et l'environnement. Auparavant, cette exploration reposait principalement sur des méthodes expérimentales souvent laborieuses et coûteuses. Cependant, l'avènement de la bio-informatique a révolutionné la manière dont nous abordons ces défis scientifiques.

La bio-informatique, ou biologie *in silico*, fusion de la biologie et de l'informatique est constituée par l'ensemble des concepts et des techniques nécessaires à l'interprétation informatique et la prédiction des informations issues des données de la biologie expérimentales. Cette révolution a ouvert la voie à de nouvelles approches, notamment le criblage virtuel par docking moléculaire. Cette méthode, combinant diverses simulations moléculaires, offre une plateforme puissante pour étudier les interactions entre des entités biologiques (1).

Dans ce contexte, notre recherche se concentre spécifiquement sur l'application du docking moléculaire dans l'étude de l'effet du venin de serpent sur la santé humaine.

L'étude du venin est particulièrement pertinente étant donné les enjeux sanitaires associés aux envenimations ophidiennes, une forme d'empoisonnement causée par des morsures de serpents venimeux. Avec environ 2.7 millions d'envenimations de serpent chaque année dans le monde, dont jusqu'à 580 000 nécessitant un traitement médical (2).

Ces envenimations constituent en Afrique un problème de santé publique majeur. Plus de 100 000 cas de décès sont enregistrés dans le monde, dont 20 000 sont sur le continent Africain. Par ailleurs, quatre cents mille des victimes envenimées conservent des séquelles fonctionnelles graves et permanentes (3).

En Algérie, deux espèces sévissent dans les zones désertiques et les hauts plateaux : *Macrovipera lebetina* et *Cerastes cerastes*. Le venin de la vipère *C. cerastes* présente une composition complexe, dominée par diverses enzymes protéolytiques telles que les métalloprotéinases (SVMP), les phospholipases A2 et les protéases sériques. Ces enzymes sont responsables des effets toxiques du venin, incluant la nécrose tissulaire, l'hémolyse et les troubles de la coagulation (4).

Les (SVMPs) sont largement répandues dans les venins des Viperidae et représentent 30 à 60 % de l'ensemble des protéines du venin. Les SVMPs sont des enzymes dépendantes du zinc, elles sont subdivisées en trois classes P-I, P-II et P-III, selon leur structure moléculaire et leur domaines (5).

Les SVMPs de classe P-II sont particulièrement intéressantes, étant impliquées dans des processus biologiques complexes tels que la modulation de l'intégrité des matrices extracellulaires et l'inhibition de l'agrégation plaquettaire. Leur activité hémorragique et leur capacité à interagir avec diverses protéines cibles en font des sujets d'étude importants pour comprendre les mécanismes de la toxicité des venins (6).

Dans cette étude, nous avons modélisé la structure tridimensionnelle d'une métalloprotéinase hémorragique CcMP-II précédemment purifiée et caractérisée à partir d'un venin de serpent de vipéridés qui est *Cerastes cerastes* (5), afin d'étudier son activité hémorragique ainsi que son effet antiagrégant plaquettaire, respectivement liée à son interaction avec le collagène type IV et aux plaquettes par docking moléculaire. A l'issue de ce docking, il sera question de comparer ces résultats avec les résultats obtenus expérimentalement au laboratoire avec cette même enzyme.

## **2. MATERIEL**

### **2.1 Enzyme métalloprotéinase hémorragique CcMP-II**

Lors de cette étude, nous avons choisi de travailler sur la protéine CcMP-II purifiée précédemment par [Bennacer et al., 2021] (5), une métalloprotéinase hémorragique de 472 acides aminés issue du venin de *Cerastes cerastes*, classée dans la catégorie P-II des métalloprotéinases du venin de serpent (SVMP) (5).

CcMP-II est structurée en deux domaines principaux : un domaine métalloprotéinase à l'extrémité N-terminale, comportant un motif de liaison au zinc (HXXGHNLGIDH) ainsi que la séquence CIM (Cys-Ile-Met) près du Met-turn, et un domaine de type désintégrine à l'extrémité C-terminale contenant les séquences RGD (Arg-Gly-Asp), essentielles pour l'inhibition de l'agrégation plaquettaire (5).

En l'absence de données sur la structure tridimensionnelle de CcMP-II, une modélisation par homologie a été nécessaire.

### **2.2 Protéines cibles**

#### **2.2.1 Collagène**

Dans notre étude, nous avons choisi de tester l'interactivité de notre enzyme avec le collagène type IV, référencé sous le code 5NAY dans la Protein Data Bank (PDB). La structure tridimensionnelle de cette protéine, déterminée par cristallographie aux rayons X avec une résolution de 1,8 Å, comprend six chaînes (A, B, C, D, E, F) et est liée aux ligands Co-cristallisés.

### **2.2.2 Intégrine**

Nous avons analysé l'interaction de notre enzyme avec l'intégrine alpha-2 beta-1, référencée sous le code 1Q7D. Cette intégrine présente une structure tridimensionnelle déterminée par cristallographie aux rayons X à une résolution de 1,8 Å et est composée de trois chaînes (A, B, C).

## **2.3 Ressources numériques et logiciels**

### **2.3.1 Protein Data Bank**

La banque de données sur les protéines (PDB) créée en 1971 est la première ressource de données numériques en son genre. Recensant plus de 200 000 structures tridimensionnelles de protéines et d'acides nucléiques, Elle offre une mine d'informations sur l'architecture et les interactions de ces biomolécules (7).

### **2.3.2 SWISS-Model**

C'est un serveur dédié à la modélisation comparative automatisée des structures tridimensionnelles (3D) des protéines. Il permet de générer des modèles de haute qualité en utilisant des séquences d'acides aminés comme point de départ (8).

La bibliothèque de modèles SWISS-MODEL est une vaste base de données contenant des structures de protéines déterminées expérimentalement, issues de la Protein Data Bank. Elle constitue la principale source d'informations structurales pour le processus de modélisation, en fournissant les coordonnées atomiques des structures protéiques et en maintenant des bases de données de séquences et de profils consultables via BLAST et HHblits. Les propriétés des modèles, indépendantes de l'alignement, sont précalculées et stockées dans la base de données (9).

### **2.3.3 BIOVIA Discovery Studio (version 2021)**

BIOVIA Discovery Studio est un logiciel de simulation et de modélisation moléculaire utilisé pour la préparation des protéines avant le docking moléculaire. Il permet de nettoyer les structures protéiques, d'ajuster les atomes manquants, de minimiser l'énergie des conformations et d'ajouter les hydrogènes nécessaires. Grâce à ses outils avancés de modélisation moléculaire, Discovery Studio optimise les protéines, assurant des résultats plus précis et fiables pour l'analyse des interactions biomoléculaire (10).

### **2.3.4 PyMOL**

C'est une interface graphique intégrant des packages académiques individuels conçus pour la préparation des protéines, les applications de mécanique moléculaire ainsi que le docking et le

scoring. PyMOL est également un système de visualisation moléculaire offrant un éventail de fonctionnalités bien plus étendu, permettant l'édition moléculaire, le rendu d'images et la réalisation de simulations dynamiques (11).

### **2.3.5 Swiss-Pdb Viewer**

C'est un logiciel de visualisation et d'analyse moléculaire qui permet la minimisation d'énergie pour optimiser les structures tridimensionnelles et atteindre des configurations plus stables. Cela est réalisé en utilisant des champs de force pour calculer l'énergie potentielle des atomes et en appliquant des algorithmes de minimisation pour ajuster progressivement les positions atomiques jusqu'à ce que l'énergie soit minimisée. Cette optimisation aide à corriger les erreurs structurelles, à modéliser les effets des mutations et à préparer les structures pour des simulations dynamiques, garantissant ainsi que les modèles reflètent des conformations réalistes et stables (3, 12).

### **2.3.6 HADDOCK (High Ambiguity Driven biomolecular docking)**

C'est un logiciel de docking flexible guidé par des informations expérimentales et bio-informatiques pour modéliser les complexes biomoléculaires. Il intègre diverses données, telles que celles provenant de la résonance magnétique nucléaire, de la spectrométrie de masse et de techniques de prédiction bio-informatique, pour diriger la modélisation et traiter la flexibilité conformationnelle. HADDOCK utilise un protocole en trois étapes comprenant la minimisation d'énergie du corps rigide, le raffinement semi-flexible et le raffinement final en solvant explicite. Il est largement utilisé pour la modélisation de complexes protéine-protéine, protéine-acides nucléiques et protéine-petites molécules, et est disponible en tant que serveur web avec une interface conviviale (13).

### **2.3.7 PDBsum**

PDBsum est une ressource en ligne qui recueille des informations détaillées sur les structures de protéines provenant de la Protein Data Bank (PDB), incluant leur séquence, leur topologie et leur interaction avec d'autres molécules. Cet outil est essentiel pour explorer et analyser les interactions entre deux protéines à un niveau moléculaire précis, en utilisant les données de structures tridimensionnelles de la PDB pour offrir une vue complète et visuelle des interactions protéine-protéine (12).

### 3. METHODE

#### 3.1 Modélisation de l'enzyme CcMP-II

Pour réaliser la modélisation tridimensionnelle (3D) de la protéine CcMP-II (*Cerastes cerastes* metalloproteinase II), la séquence d'acides aminés a été obtenue à partir d'un article publié dans The Protein Journal en 2021 intitulé "Computational and Functional Characterization of a Hemorrhagic Metalloproteinase Purified from *Cerastes cerastes* Venom», la protéine a été isolée à partir du venin de serpent de *Cerastes cerastes* (**figure 1**).

En utilisant le serveur automatisé de modélisation par homologie SWISS-MODEL, la séquence a été introduite au format FASTA, et une recherche de templates a été effectuée.

```
Nt...1MIQVLLVTICLAVFPYQSSIIILESGNVNDYEVVYPRKVTALPKGAVQKYEDAMQYEFKVNGEFVVLHLEKNKGLFS  
EDYSETHYSPDGREITTYPPVEDHCYHGGRIENDADSTASISACNGLKGFHKLQGEYLIEPLKLSNSEAHAVYKYENVEKE  
DEAPKMGVTEETNWESEYEPKASQLNLPPEQQQRYIELVVVADHRMYTKYDGDKTEISSIYEIVNLTQIFRPLHIRVAL  
IGLEIWSSGELSKVTLASDDTLEAFGEWRETVLMNRKRHDNAQLLTGMIFNTIEGRTYKSMCDPKHSVGIVRDHRTRRFV  
ANRMAHXXGHNLGIDHRDSCGANSIMSATVSNPSSQFSDCSLNQYLNHIIHYTTCLYNEPSKTDIVSPVCGNYYE  
VGEDDCGPPANCQNPCDAATCKLTTGSQCAEGLCCDQCKFMKEGTVCRVARGDDDYCNGQSACPRNGFH472...Ct
```

**Figure 1** : Séquence protéique de la métalloprotéinase hémorragique CcMP-II purifiée du venin de *Cerastes cerastes*.

##### 3.1.1 Critères de sélection des templates

La sélection du template s'est basée sur les critères suivants :

- Pourcentage d'identité de séquence : un seuil minimum de 35% d'identité de séquence a été appliqué pour garantir une similarité structurale adéquate entre la protéine CcMP-II et le template potentiel (14).
- Couverture de la séquence : le modèle doit couvrir autant que possible la séquence cible. Les lacunes (gaps) ou les régions manquantes dans l'alignement peuvent entraîner des inexactitudes dans le modèle, affectant ainsi la fiabilité de la structure modélisée (15).
- Qualité de la structure expérimentale : la qualité de la structure déterminée expérimentalement est un facteur clé. La résolution et le facteur R d'une structure cristalline sont indicatifs de leur précision (15).
- Similarité fonctionnelle : la similarité fonctionnelle entre CcMP-II et la protéine modèle était prise en compte. Les protéines qui remplissent des fonctions similaires ou qui

participent aux mêmes voies biochimiques sont plus susceptibles de posséder des caractéristiques structurelles communes (16).

- Méthode expérimentale: les templates obtenus par cristallographie aux rayons X ont été privilégiés par rapport aux modèles générés par d'autres méthodes, car ils offrent une meilleure précision structurale (17).

### 3.1.2 Procédure de Modélisation

Le modèle 3D de la protéine CcMP-II a été généré. L'évaluation initiale de la qualité du modèle s'est basée sur deux scores clés fournis par SWISS-MODEL : le Global Model QualityEstimate (GMQE) et Qualitative Model Energy ANalysisDisCo (QMEANDisCo) (13, 18).

- ❖ **Global Model QualityEstimate (GMQE)** : ce score composite, qui varie entre 0 et 1, intègre diverses caractéristiques de l'alignement entre la séquence cible et le modèle proposé, ainsi que des aspects de la structure tridimensionnelle prédite. Des valeurs proches de 1 indiquent une haute fiabilité, tandis que des valeurs proches de 0 suggèrent une faible qualité attendue du modèle. Cette évaluation globale permet d'estimer la qualité attendue du modèle dans son ensemble, fournissant ainsi une indication de sa fiabilité (19).
- ❖ **QMEANDisCo (Qualitative Model Energy ANalysisDisCo)** : analyse des distances CA-CA (carbone alpha-carbone alpha) entre un modèle donné et des structures protéiques homologues. En mesurant la cohérence de ces distances, QMEANDisCo fournit une estimation précise de la similitude d'un modèle avec des structures de référence (20). Les principaux résultats sont donnés par :
  - Score Global de QMEANDisCo : ce score fournit une évaluation globale de la qualité du modèle de protéine. Un score proche de 1 indique une bonne concordance des distances CA-CA prédites avec celles des structures de référence, ce qui suggère une bonne qualité du modèle. En revanche, un score plus bas, proche de 0, indique des écarts significatifs, signalant une moindre qualité du modèle.
  - Graphe "Local Quality" : ce graphique permet une analyse détaillée de la qualité locale du modèle. Chaque résidu du modèle est représenté sur l'axe des x, tandis que sur l'axe des y est indiquée la similitude attendue avec la structure native. Les résidus avec un score inférieur à 0,6 sont généralement considérés comme de faible qualité, mettant en évidence les zones critiques nécessitant potentiellement des ajustements ou des révisions dans le modèle.

### **3.1.3 Optimisation du Modèle**

Pour améliorer la qualité structurale du modèle, une minimisation d'énergie a été réalisée à l'aide de Swiss PDB viewer. Cette étape a permis de corriger les contraintes stériques et d'optimiser la conformation globale du modèle construit. Ensuite, la structure affinée a été visualisée et analysée à l'aide de PyMOL Molecular Viewer, permettant une inspection détaillée des ajustements effectués (21).

### **3.1.4 Validation du modèle**

La structure de la protéine a été évaluée pour sa qualité à l'aide de divers outils de validation en ligne. Le serveur SAVES (SAVESv6.0 - Structure Validation Server) a été utilisé pour cette validation, incluant :

- L'analyse du tracé de Ramachandran qui est un outil graphique permettant d'identifier les pourcentages d'acides aminés et d'analyser la conformation du squelette polypeptidique des protéines en représentant les valeurs des angles dièdres  $\phi$  et  $\psi$  sur un graphique, délimitant ainsi les régions énergétiquement favorables et défavorables (22).
- Le programme ERRAT attribue un score global compris entre 0 et 100, basé sur une analyse statistique de la distribution des facteurs de qualité pour chaque résidu. Un score ERRAT supérieur à 90 indique une bonne qualité du modèle (23).

En outre, ProSA-Web a été utilisé pour évaluer la qualité des structures de protéines en comparant les scores Z aux structures expérimentales disponibles dans la base de données PDB (24).

## **3.2 Docking protéine-protéine**

### **3.2.1 Préparations des protéines**

Nous avons préparé nos protéines, CcMP-II obtenue par modélisation, ainsi que le collagène type IV et l'intégrine alpha-2 beta-1, référencées sous les codes PDB 5NAY et 1Q7DL respectivement dans la protéine Data Bank. Les fichiers PDB des trois protéines ont été nettoyés pour éliminer toutes les molécules d'eau.

Les résidus actifs, définis comme étant exposés au solvant et directement impliqués dans l'interaction, ont été identifiés pour chaque protéine (25). Pour le collagène et l'intégrine, ces résidus actifs ont été déterminés par visualisation des interactions entre les protéines et les ligands Co-cristallisés avant que ces derniers ne soient retirés des fichiers PDB.

Pour CcMP-II, l'identification des résidus actifs a été réalisée en utilisant CASTp et FTSite. D'abord, CASTp (Computed Atlas of Surface Topography of proteins) a analysé la topologie et la géométrie des cavités de la protéine pour déterminer les sites de liaison potentiels en

fonction de leur accessibilité au solvant et de leur profondeur (26) . Ensuite, FTSite (Binding site prediction), un serveur web de cartographie des sites de fixation basé sur l'énergie, a été utilisé pour confirmer ces sites actifs en identifiant les points chauds de liaisons à travers l'interaction de sondes organiques avec la protéine (27, 28). Les résidus passifs ont été automatiquement définis par HADDOCK.

### **3.2.2 Docking et sélection des clusters**

Après cette préparation, les protéines au format PDB ont été soumises au processus de docking. Deux dockings ont été réalisés : l'un entre la protéine CcMP-II et le collagène (5 NAY), et l'autre entre CcMP-II et l'intégrine (1Q7DL). Pour chaque docking, le meilleur cluster a été sélectionné en se basant sur le score HADDOCK, qui intègre plusieurs termes énergétiques, tels que l'énergie de van der Waals résultant des interactions entre les dipôles instantanés et induits des atomes et des molécules, l'énergie électrostatique provenant des interactions entre les charges électriques des atomes et des molécules, ainsi que l'énergie de désolvation, qui reflète l'énergie nécessaire pour éliminer les molécules d'eau entourant les surfaces des protéines lors de la formation du complexe.

Les contraintes de distance sont également prises en compte, se référant aux limitations imposées sur la distance entre des paires spécifiques d'atomes ou de résidus dans les protéines pour garantir une conformation adéquate du complexe.

### **3.2.3 Visualisation moléculaire et analyse des complexes générés**

Les acides aminés formant des liaisons hydrogène, des liaisons non covalentes et des ponts salins ont été identifiés à l'aide de PDBSum, qui offre une visualisation détaillée de ces interactions au sein du complexe. Cette analyse permet ainsi de mieux appréhender la stabilité et la spécificité des complexes formés (12). Enfin, les conformations dockées ont été visualisées à l'aide de PyMol Molecular Viewer.

## **4. RESULTATS**

### **4.1 Modélisation par homologie**

#### **4.1.1. Sélection du modèle**

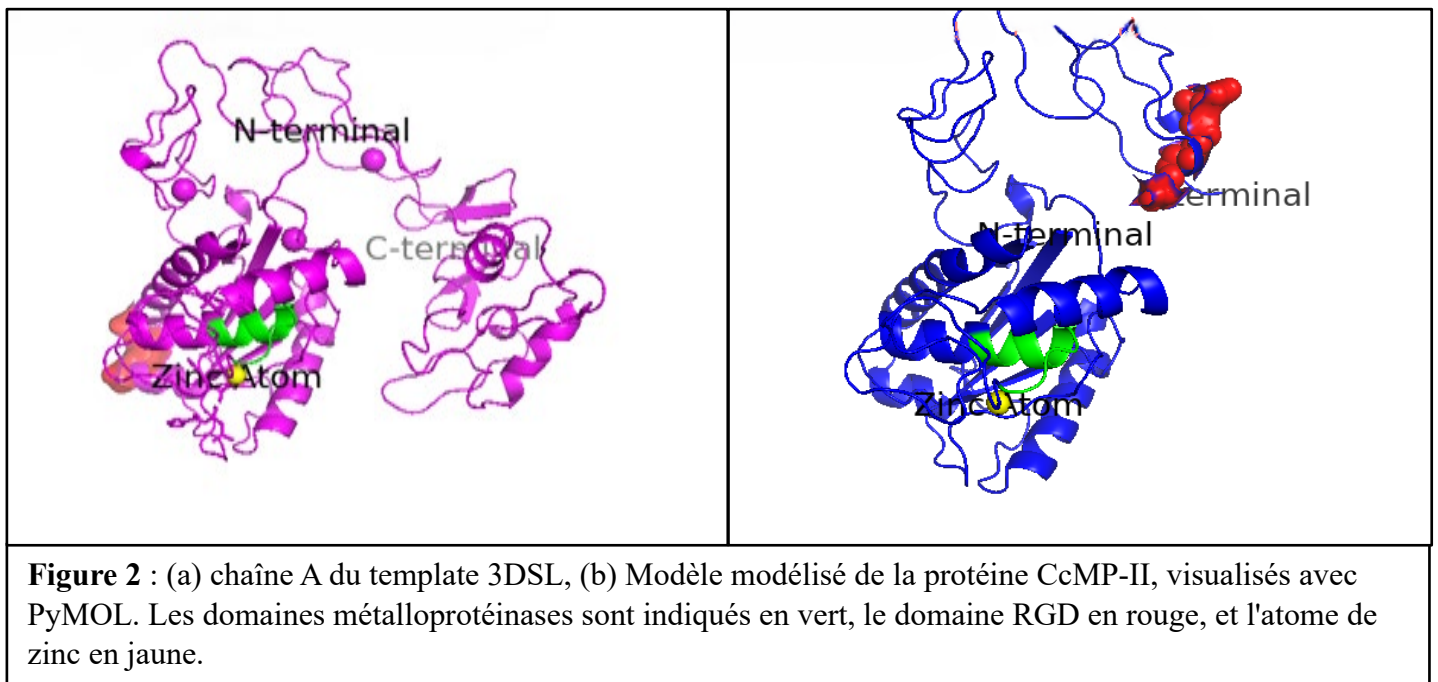
Au total, 50 templates ont été initialement identifiés par SWISS-MODEL. Après application des critères de sélection, la structure 3D de la protéine CcMP-II a été construite sur la base de

la structure d'une metalloproteinase-disintegrine Bathropasin (PDB-code : 3DSL) dérivant du venin de *Bothrops jararaca* et dont les caractéristiques sont résumées dans le **tableau I**.

**Tableau I** : Informations structurales données par SWISS-MODEL pour la protéine template 3DSL-A

PDB ID	Protéine	Similarité	Identité	Couverture	Etendue	Résolution	Méthode
3DSL	Zinc metalloproteinase - disintegrinbothropasin	55.71%	0.47	0.59	193-472	2.7 Å	Cristallographie aux rayons X

Les structures tridimensionnelles du template 3DSL-A et le modèle modélisé de la protéine CcMP-II sont illustrées dans la **Figure 2**.



#### 4.1.2 Evaluation initiale par Swiss-Model

Le score GMQE obtenu pour le modèle de la protéine CcMP-II est de 0.45.

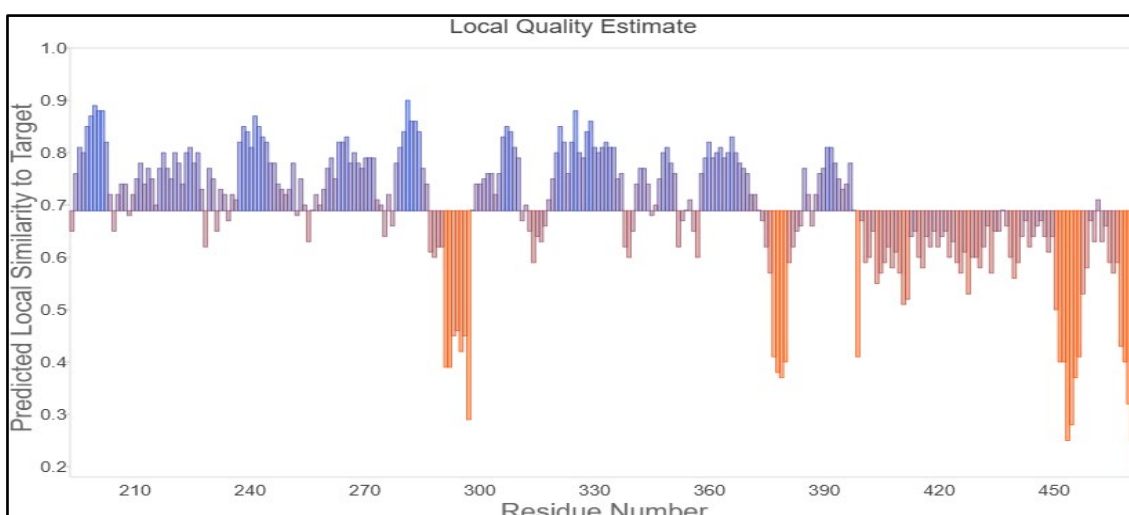
Le score QMEANDisCo global est de 0.69.

Le graphique "Local Quality Estimate" (**figure 3**) montre la similitude locale prédite avec la structure cible pour chaque résidu du modèle sur une échelle de 0 à 1.

- Résidus avec une bonne qualité : Les barres situées au-dessus de la ligne de similitude de 0,6 (en bleu) indiquent des résidus avec une bonne qualité locale. Ces régions du

modèle sont en forte cohérence avec la structure de référence, avec des valeurs comprises entre 0,7 et 0,9.

- Résidus de faible qualité : Les barres situées en dessous de la ligne de similitude de 0,6 (en orange) représentent des résidus de faible qualité. Ces zones présentent des discordances avec la structure 3DSL-A, certaines atteignant même des valeurs autour de 0,3, indiquant des zones problématiques dans le modèle.

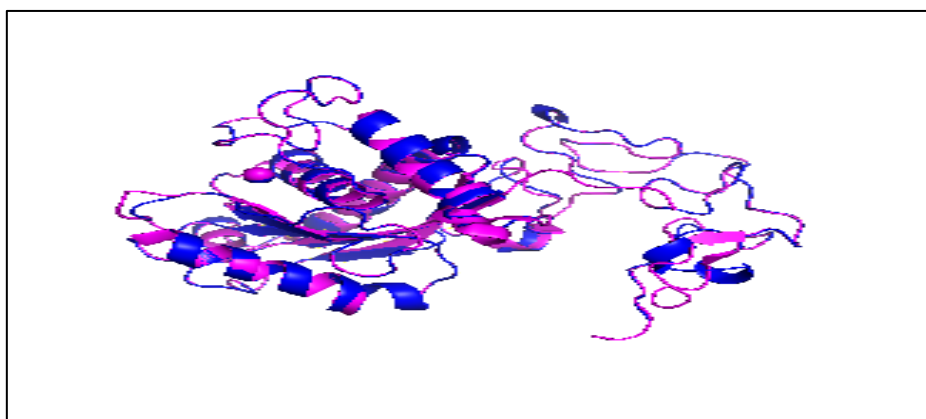


**Figure 3** : Graphique de l'estimation de la qualité locale donné par SWISS-MODEL pour CcMP-II

Ces résultats constituent une première étape dans l'évaluation de la fiabilité du modèle, nécessitant des validations complémentaires par des méthodes supplémentaires.

#### 4.1.3 Optimisation du modèle

L'alignement de la structure modélisée par homologie avec la structure tridimensionnelle (3DSL-A) montre une superposition (**figure 4**), confirmant ainsi la précision de la prédiction tridimensionnelle.



**Figure 4** : Alignement de la structure 3D des protéines : 3DSL-A en violet, CcMP-II en bleu

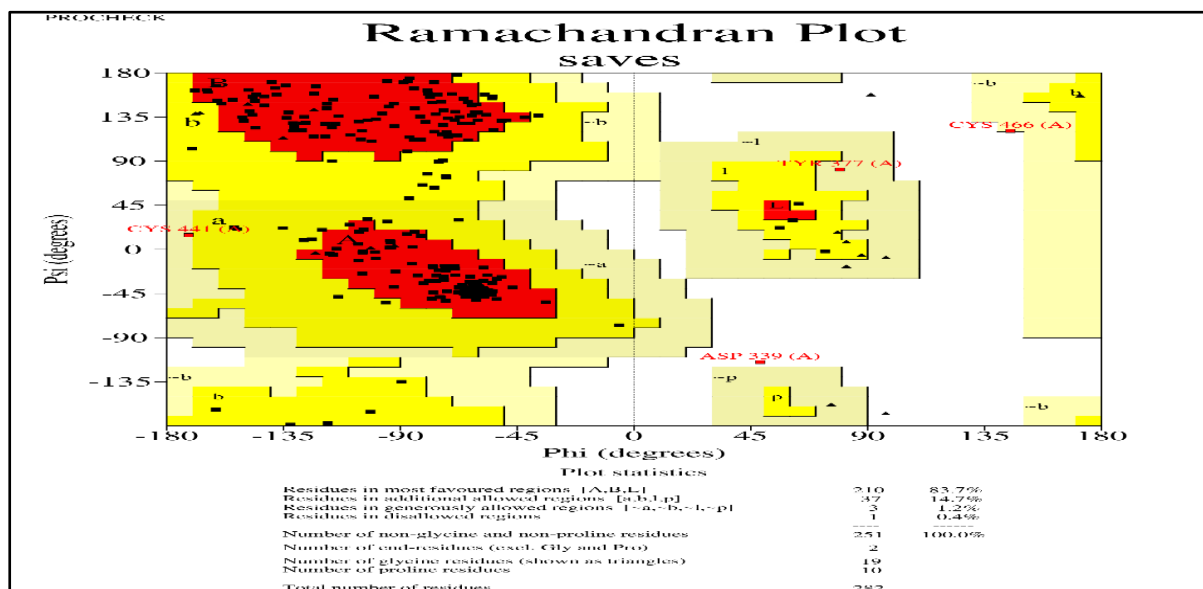
#### 4.1.4 Validation du modèle

Le diagramme de Ramachandran, élaboré à l'aide de l'outil PROCHECK version 3.5.4, a servi à valider le modèle de CcMP-II (**figure 5**). L'analyse du diagramme de Ramachandran pour CcMP-II a montré que 83,7% des résidus se trouvent dans les régions les plus favorisées, 14,7% des résidus sont dans les régions additionnellement autorisées, 1,2% des résidus sont dans les régions généreusement autorisées, et 0,4%, soit un seul, est dans une région désapprouvée.

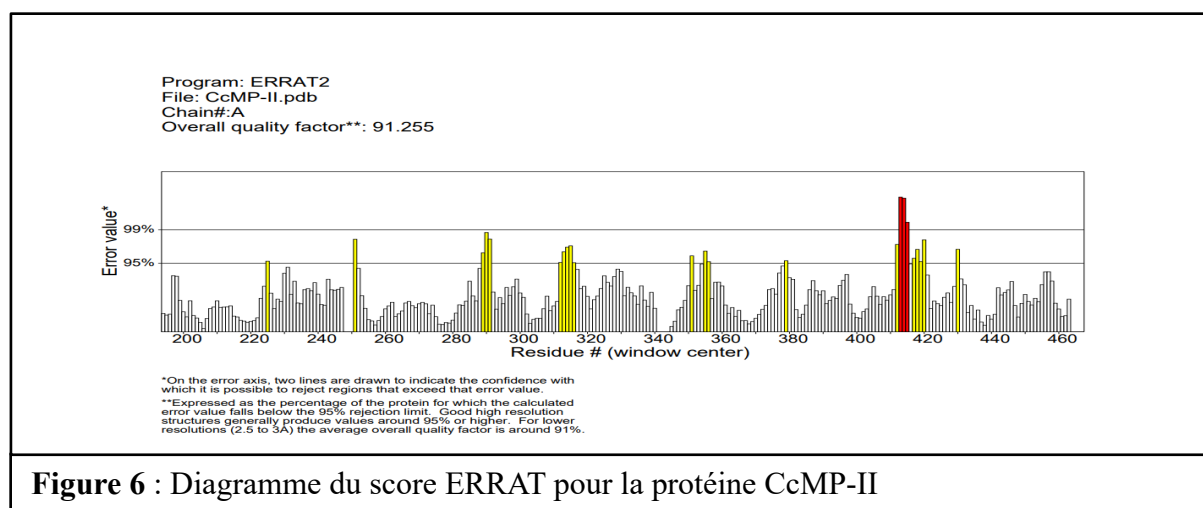
Le facteur de qualité global (ERRAT score) a été calculé à 91,255 (**figure 6**).

Le score Z global pour le modèle de la protéine CcMP-II obtenu avec ProSa-Web est de -7.09.

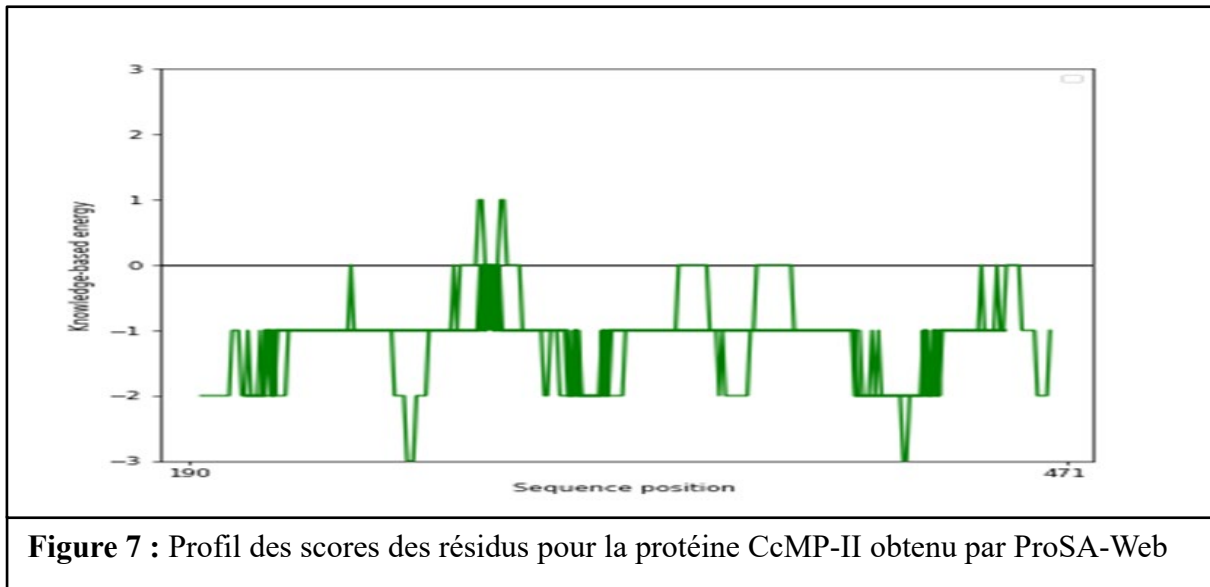
Le profil des scores des résidus obtenu par ProSA-Web montre la distribution des scores à travers la protéine. La plupart des résidus présentent des scores conformes aux attentes pour des structures validées, avec quelques régions ayant des scores légèrement déviants (**figure 7**).



**Figure 5 :** Diagramme de Ramachandran montrant les angles  $\psi$  et  $\phi$  des résidus de la protéine



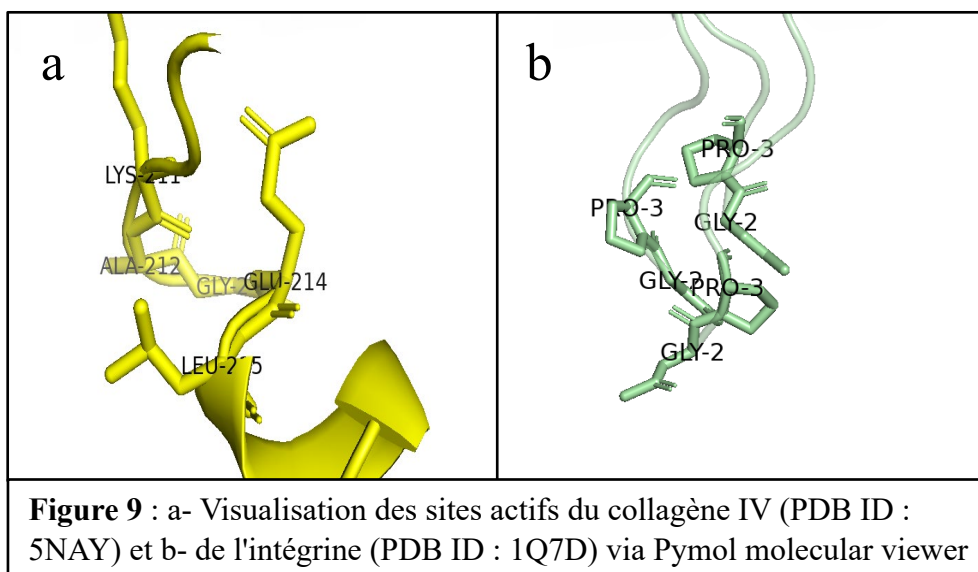
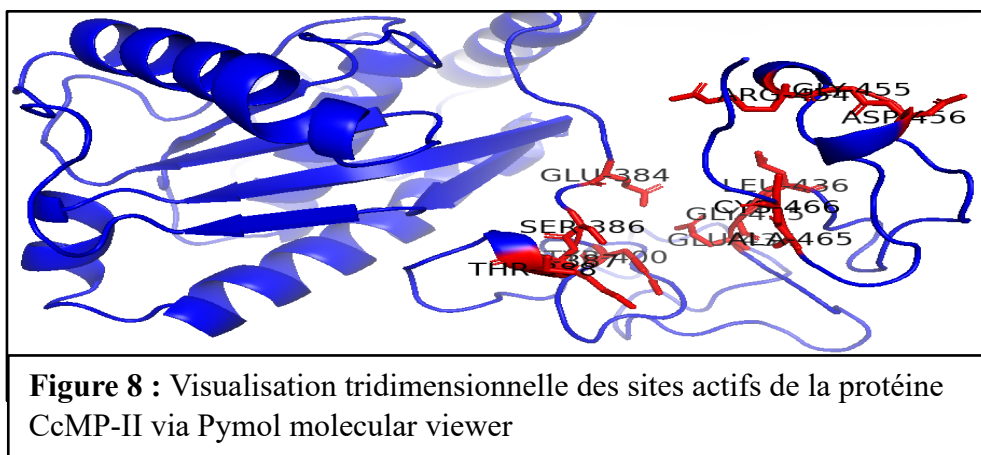
**Figure 6 :** Diagramme du score ERRAT pour la protéine CcMP-II



## 4.2 Docking protéine-protéine

### 4.2.1 Préparation des protéines

Les sites actifs identifiés pour les trois protéines ont été visualisé via Pymol (**figure 8 et 9**)



#### 4.2.2 Docking moléculaire

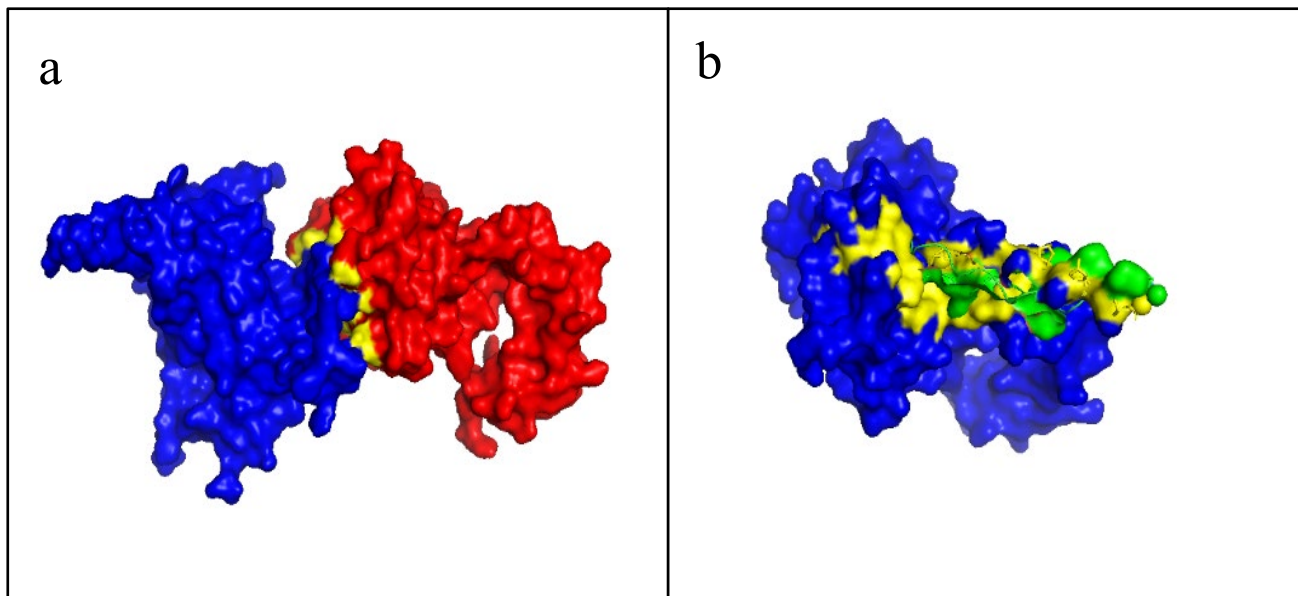
La structure PDB modélisée de la CcMP-II et celles du collagène (5NAY) et de l'intégrine obtenues à partir de la PDB ont été soumises séparément au serveur HADDOCK, générant dans chaque cas 10 clusters. Pour le docking CcMP-II-collagène, le meilleur cluster (cluster 1) a été sélectionné et ses caractéristiques sont résumées dans le **tableau II**. De même, pour le docking CcMP-II-intégrine, le cluster 2 a été sélectionné et ses caractéristiques sont résumées dans le **tableau III**. Les clusters sélectionnés ont été visualisés via PyMOL (**figure 10**).

**Tableau II** : Score HADDOCK et valeurs énergétiques du cluster sélectionné pour CcMP-II-collagène

Paramètre	Valeur
HADDOCK score	-70 +/- 3.8
Cluster size	43
RMSD from the overall lowest-energy structure	1.7 +/-1.0
Van der Waals energy	-31.1+/-3.7 kcal/mol
Electrostatic energy	-234.6+/-44.3 kcal/mol
Desolvation energy	4.6+/-5.6 kcal/mol
Restraints violation energy	33.5+/-5.4 kcal/mol
Buried Surface Area	1466.4+/-194.3 Å <sup>2</sup>
Z-Score	-1.7

**Tableau III** : Score HADDOCK et valeurs énergétiques du cluster sélectionné pour CcMP-II-intégrine

Paramètre	Valeur
HADDOCK score	-24.8 +/- 5.0
Cluster size	33
RMSD from the overall lowest-energy structure	6.3 +/- 3.8
Van der Waals energy	-12.5+/- 3 kcal/mol
Electrostatic energy	-47+/- 10.3 kcal/mol
Desolvation energy	-3.1+/- 0.4 kcal/mol
Restraints violation energy	1.4+/- 1.0 kcal/mol
Buried Surface Area	518.5+/- 85.0 Å <sup>2</sup>
Z-Score	-1.5

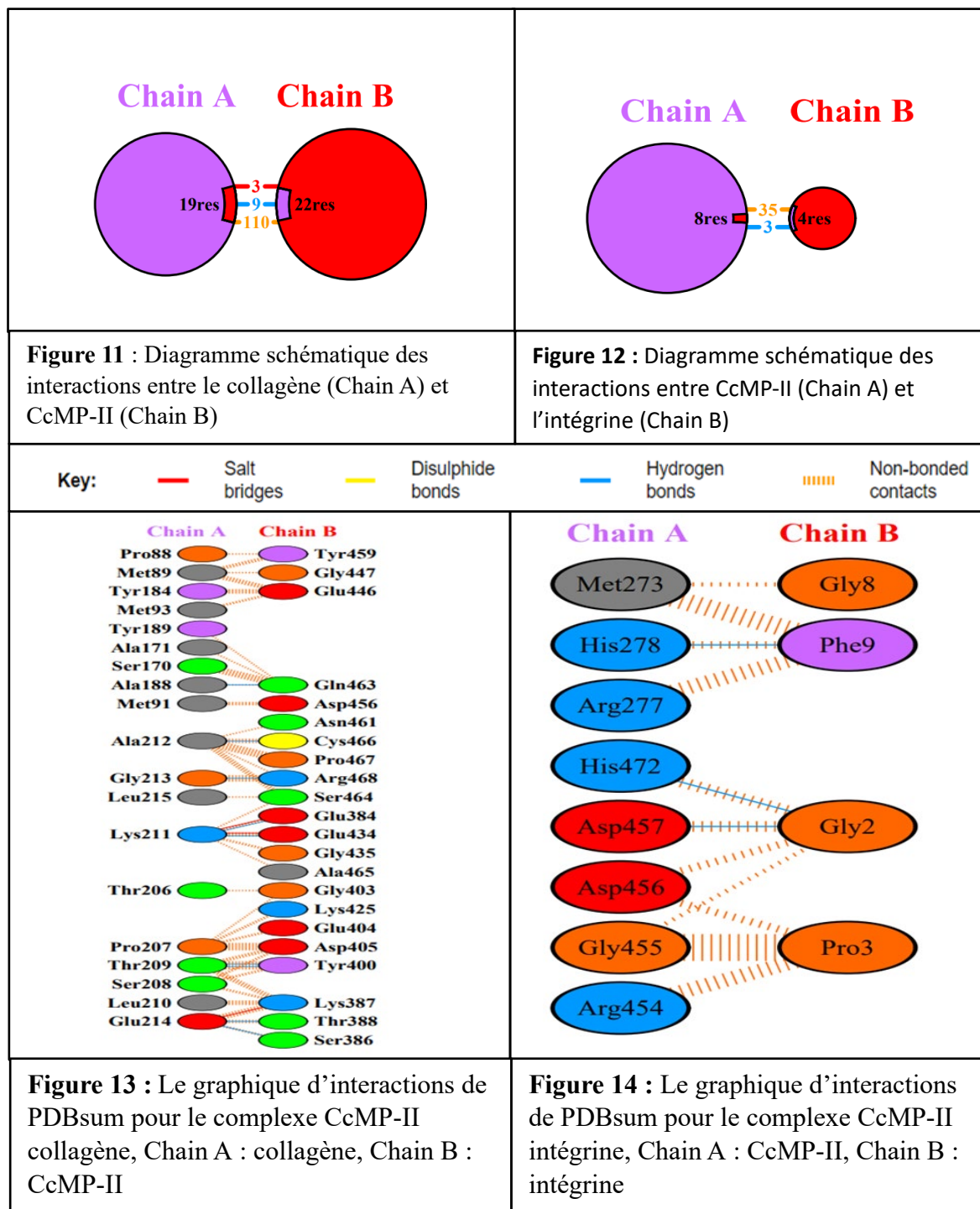


**Figure 10** : Interactions des Complexes Protéiques CCMP-II/Collagène IV et CCMP-II/Intégrine  $\alpha 2\beta 1$ . (a) Interaction CCMP-II/Collagène IV dont CCMP-II (bleu) et Collagène IV (rouge), avec la partie en interaction en jaune. (b) Interaction CCMP-II/Intégrine  $\alpha 2\beta 1$  dont CCMP-II (bleu) et Intégrine  $\alpha 2\beta 1$  (vert), avec la partie en interaction en jaune.

Les analyses des complexes CcMP-II-Collagène et CcMP-II-Intégrine ont été approfondies à l'aide de PDBSum pour identifier les interactions spécifiques impliquant des liaisons hydrogène, des interactions non liées et des ponts salins.

Pour le complexe CcMP-II-Collagène, l'analyse à l'aide de PDBSum a révélé plusieurs interactions moléculaires clés. 3 ponts salins ont été identifiés, ainsi que 9 liaisons hydrogène et 110 interactions non liées. En outre, les résidus du site actif de la protéine CcMP-II, tels que Glu384, Ser386, Lys387, Thr388 et Tyr400, jouent un rôle essentiel dans la liaison avec le site actif du collagène. Selon l'analyse par PDBSum, ces résidus établissent plusieurs liaisons hydrogènes : Thr388 de CcMP-II avec Glu214 du collagène, Ser386 avec Glu214, Tyr400 avec Thr209 et Glu434 et Lys211.

De manière similaire, l'analyse du complexe CcMP-II-Intégrine a révélé trois liaisons hydrogène spécifiques : His472 de CcMP-II avec Gly2 de l'intégrine, His278 avec Phe9 et Asp457 avec Gly2 de l'intégrine. De plus, le complexe présente 35 contacts non liés renforçant la stabilité et la spécificité de la liaison. Une interaction notable a été observée entre le domaine RGD de CcMP-II (Arg454, Gly455, Asp456) et le site actif de l'intégrine, impliquant spécifiquement Gly2 et Pro3. Ces interactions qui incluent également des liaisons hydrophobes, jouent un rôle dans la reconnaissance moléculaire et la liaison spécifique entre CcMP-II et l'intégrine.



## 5. DISCUSSION

Le docking moléculaire a été utilisé pour comprendre l'interaction entre une protéine isolée d'un venin de serpent (CCMP-II) avec le collagène type IV ainsi qu'avec l'intégrine alpha-2 beta-1 par le serveur HADDOCK.

Le collagène type IV, composant majeur de la matrice extracellulaire, assure la stabilité des parois des vaisseaux, particulièrement des artères qui doivent résister à des pressions élevées

tout en maintenant leur forme. Sa dégradation affaiblit ces parois, les rendant plus vulnérables aux ruptures sous pression, ce qui peut entraîner des saignements.

Les plaquettes sanguines jouent un rôle majeur dans l'hémostase, qui représente l'ensemble des processus physiologiques permettant de prévenir les hémorragies spontanées et d'arrêter le saignement résultant d'une brèche vasculaire. Lorsque les vaisseaux sanguins sont endommagés, le collagène exposé dans la matrice extracellulaire devient un signal d'adhésion pour les plaquettes. Les plaquettes se lient initialement au collagène via l'intégrine alpha-2 beta-1, facilitant ainsi leur activation et agrégation pour former un clou plaquettaire efficace. La perturbation de ce processus pourrait conduire à une réponse hémostatique affaiblie, augmentant potentiellement le risque de saignement prolongé et affectant la capacité du système vasculaire à répondre efficacement aux lésions et aux blessures.

### **5.1 Modélisation par Homologie de la Protéine CcMP-II**

Les scores obtenus, avec un GMQE de 0.45 et un QMEANDisCo global de 0.69, offrent une évaluation initiale de la qualité du modèle de la protéine CcMP-II.

Le GMQE de 0.45, bien qu'inférieur à 0.5, indique une structure globalement correcte. Ce score modéré est typique pour des protéines présentant une faible identité de séquence avec leurs templates.

Le score global de QMEANDisCo de 0.69 indique une bonne qualité générale du modèle, suggérant que les distances CA-CA sont largement cohérentes avec celles de la structure de référence. Cependant, le graphique "Local QualityEstimate" révèle des disparités : certaines régions du modèle affichent une qualité élevée, avec des similitudes comprises entre 0.7 et 0.9, tandis que d'autres résidus présentent des scores inférieurs à 0.6, voire aussi bas que 0.3, signalant des discordances. Ces variations montrent que, bien que le modèle soit globalement fiable, certaines zones nécessitent des améliorations. Ces résidus de faible qualité pourraient être dus à la flexibilité inhérente de ces régions ou à des erreurs de modélisation. Un affinement par minimisation d'énergie a été effectué afin d'améliorer la qualité du modèle en ajustant les conformations des résidus problématiques pour mieux correspondre à la structure de référence.

### **5.2 Validation du modèle**

L'analyse du diagramme de Ramachandran a révélé que 83,7% des résidus de la protéine CcMP-II se situent dans les régions les plus favorisées, indiquant une conformation stable du squelette polypeptidique. 14,7% des résidus se trouvent dans les régions additionnellement autorisées, suggérant une flexibilité locale de la structure. La présence de 0,4% de résidus dans

une région désapprouvée est minime et peut être attribuée à des facteurs locaux spécifiques (22, 29).

Le score ERRAT global de 91,255 confirme la bonne qualité du repliement local des résidus de la protéine. Ce score indique que la structure de CcMP-II est bien définie et peu susceptible de présenter des régions mal repliées (29).

Le score Z global de -7.09 obtenu par ProSA-Web indique que le modèle de la protéine CcMP-II est de qualité comparable aux structures expérimentales similaires (24). Ce score négatif suggère que le modèle est fiable et qu'il peut être utilisé en confiance pour des études ultérieures.

L'analyse de la structure de CcMP-II confirme la validité et la qualité du modèle généré par Swiss-Model, permettant son utilisation ultérieure pour notre étude de docking moléculaire.

### **5.3 Docking moléculaire**

Les résultats de notre étude ont révélé plusieurs aspects concernant les types d'énergies de liaisons impliquées dans les interactions protéine-protéine.

#### **- Complexe CcMP-II-collagène**

L'énergie intermoléculaire de van der Waals, qui résulte des interactions entre les dipôles instantanés et induits des atomes, a été mesurée à  $-31,1 \pm 3,7$  kcal/mol pour le cluster 1. Cette valeur indique des interactions favorables et stables entre les surfaces de CCMP-II et le collagène de type IV (30). L'énergie électrostatique a été significative avec une valeur de  $-234,6 \pm 44,3$  kcal/mol, cette énergie négative substantielle suggère des interactions électrostatiques fortes, renforçant la stabilité du complexe formé. L'énergie de désolvation a été mesurée à  $4,6 \pm 5,6$  kcal/mol. Une énergie de désolvation proche de zéro suggère que l'effet de la désolvation n'est ni particulièrement favorable ni défavorable pour la stabilité du complexe (31). Les contraintes de distance de 33.5 ont été satisfaites (32). De plus, la surface enterrée (BSA), qui indique la surface protéique inaccessible au solvant une fois le complexe formé, était de  $1466,4 \pm 194,3$  Å<sup>2</sup> pour le cluster 1. Cette valeur de BSA indique que le complexe est relativement compact, ce qui suggère une interaction modérée entre CCMP-II et le collagène de type IV (33). Le RMSD (Root Mean Square Deviation) était de  $1,7 \pm 1,0$  Å, ce qui montre une cohérence structurale raisonnable parmi les différentes conformations du cluster 1 (34). Le score HADDOCK global pour ce cluster était de  $-70,0 \pm 3,8$ , et le Z-score était de -1,7, indiquant que ce cluster est l'un des meilleurs en termes de stabilité et de qualité des interactions entre les protéines étudiées.

L'analyse de l'interaction entre la CcMP-II et le collagène de type IV par PDBsum a révélé un réseau d'interactions de 13 liaisons hydrogènes, renforçant l'intégrité structurale du complexe et augmentant sa résistance aux perturbations thermiques et mécaniques. Parmi ces liaisons, certaines sont particulièrement fortes avec des distances inférieures à 3 Å, comme celles impliquant Ala188 et Gln463 (2,99 Å) ou LYS211 et Gu284 (2,55 Å) indiquant une forte contribution à la stabilité. Bien que certaines distances soient légèrement supérieures à 3 Å, elles contribuent néanmoins à la cohésion globale du complexe. L'emplacement de ces liaisons aide à stabiliser les conformations locales des résidus à l'interface, assurant un alignement optimal des chaînes polypeptidiques pour une interaction efficace et durable. La présence de 3 ponts salins indique une forte contribution électrostatique à la stabilité globale du complexe. En plus de ces liaisons, les interactions hydrophobes identifiées résultant de l'exclusion de l'eau de l'interface de liaison favorisent la formation d'un noyau hydrophobe stable à l'intérieur de la structure.

#### - **Complexe CcMP-II-intégrine**

Le score HADDOCK pour le cluster 2 était de  $-24,8 \pm 5,0$  Kcal/mol, indiquant une stabilité satisfaisante et des interactions significatives entre CcMP-II et l'intégrine alpha-2 beta-1.

L'énergie intermoléculaire de van der Waals était mesurée à  $-12,5 \pm 3,0$  Kcal/mol, indiquant des interactions stables mais moins favorables que dans d'autres clusters (30).

L'énergie électrostatique à  $-47,0 \pm 10,3$  Kcal/mol, souligne des interactions fortes qui contribuent à la stabilité du complexe.

L'énergie de désolvation, évaluée à  $-3,1 \pm 0,4$  Kcal/mol, montre un effet modéré sur la stabilité du complexe en raison de la solvation (31). La surface enterrée (BSA) de  $518,5 \pm 85,0$  Å<sup>2</sup> pour ce cluster indique une interface relativement étendue entre CcMP-II et l'intégrine alpha-2 beta-1, favorisant une interaction substantielle (33). Le RMSD par rapport à la structure d'énergie minimale de  $6,3 \pm 3,8$  Å, suggérant une diversité dans les conformations possibles mais avec une convergence raisonnable vers une structure moyenne (34).

L'analyse par PDBsum de l'interaction entre la CcMP-II et l'intégrine alpha-2 beta-1 a mis en évidence 3 liaisons hydrogènes avec des distances inférieures à 2,9 Å comme celle impliquant Asp457 et Gly2 (2,72 Å). La présence de 35 ponts salins renforce la stabilité du complexe.

Les résultats du docking protéine-protéine concorde avec les résultats obtenus au laboratoire car l'interaction entre CcMP-II et le collagène d'une part, et entre CcMP-II et l'intégrine d'autre part a été étudiée *in vivo* et *in vitro*. Expérimentalement, il a été démontré que CCMP-II compromet l'intégrité des vaisseaux capillaires par 2 mécanismes. La CcMP-II hydrolyse les substrats de la membrane basale, en particulier le collagène IV, qui est un composant essentiel

entourant les cellules endothéliales des capillaires. Les coupures induites par CcMP-II sur le collagène IV affaiblissent la stabilité mécanique de la membrane basale. Ce processus entraîne la distension des parois des micro vaisseaux sous l'influence des forces hémodynamiques circulatoires, conduisant finalement à la rupture de la paroi capillaire et à l'extravasation. Par ailleurs, cette enzyme inhibe l'agrégation plaquettaire stimulée par le collagène en interagissant avec les intégrines présentes sur les plaquettes. Cette liaison empêche le collagène de se fixer sur ces récepteurs, bloquant ainsi les voies de signalisation nécessaires à l'agrégation plaquettaire (35).

L'analyse histopathologique a révélé des dommages capillaires significatifs et des hémorragies chez les souris après injection de CcMP-II, caractérisés par des projections cytoplasmiques, un amincissement des cellules endothéliales et une perturbation de la membrane basale. CcMP-II inhibe également l'agrégation des plaquettes stimulée par le collagène même en présence des inhibiteurs de l'effet du domaine métalloprotéinase (EDTA et 1,10-phénantroline). Le pourcentage d'inhibition de l'agrégation plaquettaire par CcMP-II est estimé à 70 % en présence d'EDTA. Ce résultat démontre que l'effet anti-agrégant de la CcMP-II hémorragique est dû à son domaine de type désintégrine.

Les résultats obtenus par ces deux approches se sont révélés concordants, ce qui vient confirmer la fiabilité de l'approche *in silico*. Elle présente plusieurs atouts majeurs par rapport à l'expérimentation traditionnelle. Elle est plus rapide, moins coûteuse et permet d'explorer un grand nombre de combinaisons moléculaires. De plus, elle peut identifier des interactions qui pourraient passer inaperçues lors d'expérimentations classiques. Enfin, elle offre un niveau de détail inégal dans l'analyse des interactions moléculaires, permettant de visualiser et d'analyser les structures tridimensionnelles des molécules à l'échelle atomique.

La convergence des résultats expérimentaux et *in silico* dans cette étude démontre la puissance et la complémentarité de ces deux approches. L'expérimentation reste essentielle pour confirmer les prédictions *in silico* et fournir des informations sur les effets physiologiques des interactions moléculaires. Cependant, l'étude *in silico* s'impose comme un outil indispensable pour l'investigation des interactions protéine-protéine et la compréhension des processus biologiques complexes.

## **6. CONCLUSION**

Dans cette étude, 2 étapes expérimentales ont été utilisées afin d'évaluer l'effet d'une protéine vipérine (CCMP-II) sur le collagène type IV et l'intégrine alpha-2 beta-1 : la modélisation par homologie et le docking moléculaire protéine-protéine.

Il existe à l'heure actuelle deux méthodes expérimentales permettant de déterminer la structure 3D d'une protéine ou d'un complexe. Ce sont la cristallographie aux rayons X et la Résonance Magnétique Nucléaire (RMN). Cependant ces techniques demeurent coûteuses et assez longues à réaliser aux vues des nombreuses protéines présentes dans le monde du vivant (estimée à environ 20 000 pour l'Homme). C'est pour ces raisons que de nombreux chercheurs ont cherché à prédire la structure de ces complexes grâce à l'informatique, via les méthodes et algorithmes de docking. Ces méthodes sont fondées sur deux critères issus des connaissances acquises via les complexes existants dans la PDB : critère géométrique/topologique et critère énergétique (36).

Cette étude confirme que le docking moléculaire est un outil puissant et complémentaire aux approches expérimentales classiques. Il offre des avantages en termes de rapidité, de coût, de précision prédictive et de capacité à explorer efficacement l'espace chimique. Les simulations *in silico* peuvent fournir des prédictions initiales rapides permettant d'orienter les expérimentations et d'optimiser les ressources de recherche. Ces caractéristiques font du docking moléculaire un pilier essentiel dans la recherche en biologie structurale et dans l'étude des interactions biomoléculaires complexes. En intégrant les approches *in silico* et expérimentales, les chercheurs peuvent obtenir une compréhension plus complète et précise des mécanismes biomoléculaires.

## BIBLIOGRAPHIE

1. Adjerci A. Docking moléculaire et criblage ADME/Toxicité pour des inhibiteurs de l'alpha-amylase et de l'alpha-glucosidase. DSpace ANNABA. 2023.
2. Gutiérrez JM, Calvete JJ, Habib AG, Harrison RA, Williams DJ, Warrell DA. Snakebite envenoming. Nature reviews Disease primers. 2017;3:17063.
3. Guex N, Peitsch MC, Schwede T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. Electrophoresis. 2009;30 Suppl 1:S162-73.
4. Chippaux J-P. Venins de serpent et envenimations. IRD Editions. 2002.
5. Bennacer A, Boukhalfa-Abib H, Laraba-Djebari F. "Computational and Functional Characterization of a Hemorrhagic Metalloproteinase Purified from Cerastes cerastes Venom". Protein J. 2021;40(4):589-99.
6. Guyadec GL. Venins de serpent et hémostase : revue bibliographique et mesure de l'activité de protéases ophidiennes. HAL open science 2022.
7. Banque de données sur les protéines RCSB (RCSB PDB) [Internet]. 2023 [cited 2024,avr]. Available from: <https://www.rcsb.org/>.
8. Swiss M. Internet 2024 [updated 2024,avr. Available from: <https://swissmodel.expasy.org/>.
9. Kopp J, Schwede T. The SWISS-MODEL Repository of annotated three-dimensional protein structure homology models. Nucleic Acids Res. 2004;32(Database issue):D230-4.
10. Systemes D. Biovia Discovery Studio [Available from: [https://www.3ds.com/products/biovia?fbclid=IwZXh0bgNhZW0CMTAAR1MJ5q-zG6mY-DHPGT9mwCbWrj\\_4j1tRKOmFgfOq2Qh6M70frBspJcsnV8\\_aem\\_ZmFrZWR1bW15MTZieXRlcw](https://www.3ds.com/products/biovia?fbclid=IwZXh0bgNhZW0CMTAAR1MJ5q-zG6mY-DHPGT9mwCbWrj_4j1tRKOmFgfOq2Qh6M70frBspJcsnV8_aem_ZmFrZWR1bW15MTZieXRlcw).
11. Lill MA, Danielson ML. Computer-aided drug design platform using PyMOL. J Comput Aided Mol Des. 2011;25(1):13-9.
12. Laskowski RA, Jablonska J, Pravda L, Varekova RS, Thornton JM. PDBsum: Structural summaries of PDB entries. Protein Sci. 2018;27(1):129-34.
13. Dominguez C, Boelens R, Bonvin AMJJ. HADDOCK: a protein-protein docking approach based on biochemical or biophysical data. J Am Chem Soc. 2003;125(7):1731-7.

14. Pellequer J-L. Modélisation moléculaire par homologie des protéines : ses applications en Biologie et en Bioinformatique. HAL open science. 2016.
15. Fiser A. Template-based protein structure modeling. *Methods in molecular biology* (Clifton, NJ). 2010;673:73-94.
16. Cavasotto CN, Phatak SS. Homology modeling in drug discovery: current trends and applications. *Drug Discov Today*. 2009;14(13-14):676-83.
17. Coumar MS. *Molecular Docking for Computer-Aided Drug Design Fundamentals, Techniques, Resources and Applications*. ScienceDirect. 2021.
18. Benkert P, Biasini M, Schwede T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics*. 2011;27(3):343-50.
19. SWISS-MODEL. Global Model Quality Estimation (GMQE) [Available from: [https://swissmodel.expasy.org/docs/help/accuracy\\_help#gmqe](https://swissmodel.expasy.org/docs/help/accuracy_help#gmqe)].
20. Bioinformatics ESlo. QMEAN - SWISS-MODEL [Internet]. Zurich: Swiss Institute of Bioinformatics; [consulté le 15 mai 2024]. Disponible à : <https://swissmodel.expasy.org/qmean/>.
21. Suganthi M, Sowmya H, Manjunathan J, Ramasamy P, Thiruvengadam M, Varadharajan V, et al. Homology modeling and protein-protein interaction studies of GAPDH from *Helopeltis theivora* and chitinase from *Pseudomonas fluorescens* to control infection in tea [*Camellia sinensis* (L.) O. Kuntze] plants. *Plant Stress*. 2024;11.
22. Sobolev OV, Afonine PV, Moriarty NW, Hekkelman ML, Joosten RP, Perrakis A, et al. A Global Ramachandran Score Identifies Protein Structures with Unlikely Stereochemistry. *Structure*. 2020;28(11):1249-58 e2.
23. YEATES CCAT. Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Science*. 1993:1511-9.
24. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res*. 2007;35(Web Server issue):W407-10.
25. Laudet B, Prudent R, Filhol O, Cochet C. [Therapeutic agents targetting protein-protein interactions: myth or reality?]. *Med Sci (Paris)*. 2007;23(3):273-8.
26. Tian W, Chen C, Lei X, Zhao J, Liang J. CASTp 3.0: computed atlas of surface topography of proteins. *Nucleic Acids Res*. 2018;46(W1):W363-W7.

27. Ngan CH, Hall DR, Zerbe B, Grove LE, Kozakov D, Vajda S. FTSite: high accuracy detection of ligand binding sites on unbound protein structures. *Bioinformatics*. 2012;28(2):286-7.
28. Brenke R, Kozakov D, Chuang GY, Beglov D, Hall D, Landon MR, et al. Fragment-based identification of druggable 'hot spots' of proteins using Fourier domain correlation techniques. *Bioinformatics*. 2009;25(5):621-7.
29. Yamkela M, Sitobo Z, Makhoba XH. In Silico Analysis of SARS-CoV-2 Non-Structural Proteins Reveals an Interaction with the Host's Heat Shock Proteins That May Contribute to Viral Replications and Development. *Curr Issues Mol Biol*. 2023;45(12):10225-47.
30. Nathan Desdouits. Concepts et méthodes d'analyse numérique de la dynamique des cavités au sein des protéines et applications à l'élaboration de stratégies novatrices d'inhibition. *Bio-Informatique, Biologie Systémique [q-bio.QM]*. Université Pierre et Marie Curie - Paris VI, 2015. Français. NNT : 2015PA066250 . tel-01316546.
31. Kastritis PL, Rodrigues JP, Bonvin AM. HADDOCK(2P2I): a biophysical model for predicting the binding affinity of protein-protein interaction inhibitors. *J Chem Inf Model*. 2014;54(3):826-36.
32. Chandramohan UM, Katta P, Prabakaran A, Prasath M. QM, molecular docking and molecular dynamics investigation on acidic phospholipase A2 2 protein and acidic phospholipase A2 3 protein with silane dimethyl. *Chemical Physics Impact*. 2024;8.
33. Guillaume Letellier. Modélisation du complexe récepteur muscarinique/ toxique MT7 à partir de données thermodynamiques. *Sciences du Vivant [q-bio]*. Université Paris-Diderot - Paris VII, 2008. Français. NNT : . tel-00447060.
34. Damm KL, Carlson HA. Gaussian-weighted RMSD superposition of proteins: a structural comparison for flexible proteins and predicted protein structures. *Biophys J*. 2006;90(12):4558-73.
35. Gutierrez JM, Escalante T, Rucavado A, Herrera C. Hemorrhage Caused by Snake Venom Metalloproteinases: A Journey of Discovery and Understanding. *Toxins (Basel)*. 2016;8(4):93.
36. Nicolas Férey, Guillaume Bouyer, Christine Martin, Abdelhamid Drif, Patrick Bourdot, et al.. Docking de protéines en réalité virtuelle. Une approche hybride et multimodale. *Revue des Sciences et Technologies de l'Information - Série TSI : Technique et Science Informatiques*, 2009, 28 (8), pp.983-1015. 10.3166/tsi.28.983-1015 . hal-01753902.

