



ROYAUME DU MAROC UNIVERSITE  
MOHAMMED V FACULTE DE MEDECINE  
ET DE PHARMACIE RABAT



Année 2022

Numéro d'ordre : 32/21/CSVS

## THÈSE DE DOCTORAT

Centre d'Etude Doctoral Science de la vie et de la Santé

Formation Doctorale: Biologie médicale, Pathologie Humaine et Expérimentale  
et Environnementale

*Intitulé de la thèse*

**Etude du mécanisme d'interaction et de signalisation des acides gras  
via CD36 avec et sans infection à la COVID 19 : Développement de  
Workflow d'identification de molécules bioactives *in silico***

JIHANE AKACHAR

Présentée et soutenue le 28 Novembre 2022

### JURY:

**Professeur M'hammed ANSAR**

Faculté de Médecine et de Pharmacie, Université Mohammed V, Rabat

**Président**

**Professeur Rachid EL JAOUDI**

Faculté de Médecine et de Pharmacie, Université Mohammed V, Rabat

**Directeur de thèse**

**Professeur Azeddine IBRAHIMI**

Faculté de Médecine et de Pharmacie, Université Mohammed V, Rabat

**Co-Directeur de thèse**

**Professeur Anas KETTANI**

Faculté des Sciences Ben Msik, Université Hassan II de Casablanca

**Examineur**

**Professeur Khalil HAMMANI**

Faculté Polydisciplinaire de Taza, Université Sidi Mohamed Ben Abdellah Fès

**Rapporteur**

**Professeur Jaouad EL HARTI**

Faculté de Médecine et de Pharmacie, Université Mohammed V, Rabat

**Rapporteur**

**Professeur Youssef RAMLI**

Faculté de Médecine et de Pharmacie, Université Mohammed V, Rabat

**Rapporteur**

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

فَالسُّلْبُ بِحَانِكَ لَأَسْلَمُ  
إِنِّي أَنبِئُكَ الْعِمْلَ الْحَكِيمَ

صَدَقَ اللَّهُ الْعَظِيمَ

## REMERCIEMENTS

Ce document est le produit de centre de Biotechnologie Médicale (MedBiotech), Faculté de Médecine et de Pharmacie, Rabat, un effort continu qui n'aurait pas été possible sans la contribution de nombreuses personnes. C'est votre chance de leur montrer votre appréciation :

**A Mon Co-Directeur de thèse le Pr. Azeddine IBRAHIMI**, Professeur et Chef du Laboratoire de Biotechnologie Médicale (Médbiotech) à la Faculté de Médecine et de Pharmacie de Rabat, Université Mohammed V, pour me faire partie d'une équipe de recherche, pour votre confiance dès le départ, soutien, disponibilité infatigable malgré un horaire très chargé, et persévérance. Aussi, pour tous vos précieux conseils, vos connaissances scientifiques et vos encouragements durant toutes les années que j'ai passées dans votre laboratoire menant à cette thèse de doctorat en modélisation moléculaire et analyse bioinformatique, à chaque étape je tiens à vous remercier pour la liberté d'action que vous avez m'a donné. Voici tous mes remerciements et mes plus profonds respects.

**A Monsieur le Pr Rachid El Jaoudi, Directeur de thèse** Professeur à la Faculté de Médecine et de Pharmacie de Rabat, Université Mohammed V. Merci de votre intérêt pour la recherche scientifique et les progrès biomédicaux au Maroc. Voici ma gratitude, ma profonde gratitude et ma profonde gratitude.

Mes plus sincères remerciements vont également à Madame **Pr.Catherine Etchebest**, professeure à l'université Paris Diderot et directrice de l'équipe DSIMB (Dynamiques des structures et des interactions des macromolécules biologiques) évoluant au sein du laboratoire Biologie Intégrée du Globule Rouge (Inserm/Paris Diderot), pour ces conseils et ses commentaires m'ont été très utiles en raison de l'intérêt particulier qu'elle m'a témoigné, de son expérience et de ses capacités.

Je suis également reconnaissant envers le Professeur **Naim Akhtar KHAN** pour m'avoir accueilli, au sein de son équipe NUTox, dans un excellent environnement scientifique au sein de son Laboratoire Physiologie de la Nutrition & Toxicologie de l'Université de Bourgogne France.

Je voudrais ensuite remercier Professeur **Youssef Ramli**, le tout avec humour et beaucoup d'autonomie pour leur rigueur scientifique, leur disponibilité et, bien sûr, pour toujours partager leurs connaissances et leur expérience.

Un grand merci à tout le personnel du Laboratoire Biotechnologie MedBiotech de la Faculté de Médecine et de Pharmacie de Rabat.

Je tiens à remercier **Pr. Mohammed Ansar** d'avoir accepté d'être président du jury. Je remercie également tous les membres du jury d'avoir accepté d'assister à la présentation de ce travail, particulièrement le **Pr. Youssef Ramli, Jaouad EL Harti, Pr. Hammani Khalil et Pr. Anas Kettani** pour avoir généreusement accepté de rapporter et d'examiner mon travail.

**A Mr. Le président de jury** ainsi que **les membres de jury**, malgré vos nombreuses activités, vous êtes honoré d'évaluer mon travail en acceptant de faire partie de mon jury.



# Dédicace

*Chers parents, pour tous vos sacrifices, votre amour, votre gentillesse, votre soutien et vos prières durant mes études,  
A mes chers frères et sœurs pour leurs encouragements constants et leur soutien moral*

*A toute ma famille qui m'a soutenu tout au long de mon parcours universitaire,*

*Que ce travail soit l'accomplissement de votre soi-disant désir et le résultat de votre soutien constant.*

*Merci d'être toujours à mes côtés.*

## LISTE DES PUBLICATIONS

### ARTICLES

- 1. Akachar, J**, Etchebest, C, Bouricha,E, El Jaoudi,R and Ibrahimi (2022).A. *In Silico* Workflow Analysis of Fatty Acid Signaling Mechanism via the CD36 Receptor and Design of New compounds Derived from Natural anti-obesity agents as Potential of the Fat-taste Analogues. *Journal of chemical information and modeling*. (en cours de soumission)
- 2. Akachar, J**, Bouricha,E, El Jaoudi,R and Ibrahimi,A (2022). *In silico* structural mecanisme analysis of COVID-19 Induced Fatty Acid Taste Impairments via the CD36 Receptor. *Spring Journal of 3Biotech*. (en cours de soumission)
- Guerrab W, **Akachar J**, Jemli ME, Abudunia AM, Ouaabou R, Alaoui K, Ibrahimi A, Ramli Y. (2022) May Synthesis, molecular docking, ADMET evaluation and *in vitro* cytotoxic activity evaluation on RD and L20B cell lines of 3-substituted 5,5-diphenylimidazolidine-2,4-dione derivatives. *J Biomol Struct Dyn*. 1:1-9.
- Chkirate, K., **Akachar, J**, Hni, B., Hökelek, T., Talbaoui, A., Mague, J. T., ... & Essassi, E. M. (2022). Synthesis, spectroscopic characterization, crystal structure, DFT, ESI-MS studies, molecular docking and in vitro antibacterial activity of 1, 5-benzodiazepin-2-one derivatives. *Journal of Molecular Structure*, 1247, 131188.
- Abad, N., Sebbar, N. K., Talbaoui, A., Bakri, Y., **Akachar, J**, Ibrahimi, *et al.* (2022). Synthesis and modeling of new 1, 2, 3-triazolymethyl-quinoxaline derivatives obtained via 1, 3-dipolar cycloaddition reactions. *Moroccan Journal of Heterocyclic Chemistry*, 21(01), 21-01.
- Akachar, J**, Etchebest,C, El Jaoudi,R & Ibrahimi,A. (2021). The computational analyses,molecular dynamics of fatty-acid transport mechanism to the CD36 receptor. *Nature Journal of Scientific Report*.
- Guerrab, W., El Jemli, M., **Akachar, J**, Demirtaş, G., Mague, J. T., Taoufik, J., *et al.*(2021). Design, synthesis, structural and molecular characterization, toxicity, psychotropic activity and molecular docking evaluation of a novel phenytoin derivative: 3-decyl-5, 5-diphenylimidazolidine-2, 4-dione. *Journal of Biomolecular Structure and Dynamics*, 1-18.
- Akachar, J**, Bouricha, E. M., Hakmi, M., Belyamani, L., El Jaoudi, R., & Ibrahimi, A. (2020). Identifying epitopes for cluster of differentiation and design of new peptides inhibitors against human SARS-CoV-2 spike RBD by an in-silico approach. *Heliyon*, e05739.
- Hakmi, M., Bouricha, E. M., **Akachar, J**, Lmimouni, B., El Harti, J., Belyamani, L., & Ibrahimi, A. (2020). *In silico* exploration of small-molecule  $\alpha$ -helix mimetics as inhibitors of SARS-COV-2 attachment to ACE2. *Journal of Biomolecular Structure and Dynamics*, 1-12.
- Bouricha, E. M., Hakmi, M., **Akachar, J**, Zouaidia, F., & Ibrahimi, A. (2020). In-silico identification of potential inhibitors targeting the DNA binding domain of estrogen receptor  $\alpha$  for the treatment of hormone therapy-resistant breast cancer. *Journal of Biomolecular Structure and Dynamics*, 1-8.

11. Bouricha, E. M., Hakmi, M., **Akachar, J.**, Belyamani, L., & Ibrahimi, A. (2020). *In silico* analysis of ACE2 orthologues to predict animal host range with high susceptibility to SARS-CoV-2. *3 Biotech*, 10(11), 1-8.
12. Chemlal, L., **Akachar, J.**, Makram, S., Zoubir, B., Cherrah, Y., Eljaoudi, R., ... & Faouzi, M. A. (2019). The displacement study of <sup>99m</sup>Tc-DTPA—Human serum albumin binding in presence of furosemide and metformin by using equilibrium dialysis and molecular docking. *IUBMB life*, 71(12), 2003-2009.
13. **Akachar, J.**, Abudunia, A, El Jaoudi, R, Bennani, FA, Ansar, M, Ibrahimi, A, Al-kaf, A. (2018) Synthesis, Anti-hypoxic Activity and Molecular Docking Studies of New Quinazolinone-4 Derivatives. *ISESCO Journal of Science and Technology*.
14. Allam, L., Lakhli, W., Tarhda, Z., **Akachar, J.**, Ghrifi, F., Amri, H. E., & Ibrahimi, A. (2017). Three-dimensional structure prediction of the human LMTK3 catalytic domain in DYG-in conformation. *J Biomol Res Ther*, 6(151), 2.
15. Al-kaf, A, **Akachar, J.** (2017). Synthesis, analgesic, anti-inflammatory and molecular docking studies of new quinazolin-4-one derivatives. *Moroccan Journal of Heterocyclic Chemistry*, 16(1).

#### COMMUNICATIONS ORALES

**J.Akachar, L.Lhalou, H.Lemriss, R.Eljaoudi, A.Ibrahimi.** (2015). Mathematical and computational modeling CD36-Ligands, *ASNME Agadir School on Numerical Methods in Engineering, Morocco*

#### COMMUNICATIONS AFICHEES

**J.Akachar, R.Eljaoudi, A.Ibrahimi.** (2017). Mathematical and computational modeling CD36-Ligands, *Journé de Savoir La JMFB, la Jeunesse Musulmane de France de Bourgogne Dijon, France*,

## RESUME

**Title:** Etude du mécanisme d'interaction et de signalisation des acides gras via CD36 avec et sans infection COVID 19 : Développement de Workflow d'identification de molécules *bioactives in silico*

**Auteur :** Jihane AKACHAR

**Mot-clés:** Acide gras, CD36, Dynamique moléculaire, Modélisation moléculaire, Bioinformatique, Conception de molécule, COVID19.

Une consommation excessive de lipides en particulier d'Acides Gras Saturés associé à un déséquilibre qualitatif contribue à la prévalence de l'obésité. Le récepteurs lipidiques CD36 est un candidat plausible en tant que capteur lipidique gustatif, également induisant une cascade de signalisation cellulaire. Le but de cette thèse a été la conception, le criblage et l'identification de nouvelles molécules bioactives par le développement de workflow basés sur l'approche *in silico*. Nous avons utilisé des séries de simulations de dynamique moléculaire de la protéine CD36, en interaction avec un à trois acides palmitiques (PLM), afin d'établir un premier workflow pour découvrir le mécanisme du fonctionnement du complexe CD36/Acides gras et de déterminer le rôle structural de la Lys 164. Les résultats ont montré que la Lys164 influence d'une manière importante sur la forme de stabilité du CD36 et. La conception grâce aux résultats du premier workflow de 15 molécules activant la voie de signalisation des acides gras et plus particulièrement ceux qui sont spécifiques au goût de gras. Uniquement, l'analogue CIN4 a montré, par la suite, par criblage dynamique, une dimérisation des deux domaines TM1 et TM2 qui ressemble à l'acide palmitique. En outre, les étapes de ce premier workflow ont permis d'explorer le dysfonctionnement gustatif de la chimio-réception des acides gras spécifiquement le "goût de gras" des patients testés positifs au COVID-19. Les résultats de cette étude ont montré une affinité élevée et une stabilité du domaine RBD du SARS-CoV-2 non mutant au récepteur CD36, un deuxième workflow a été établi. Le criblage virtuel a été utilisé dans un premier temps pour identifier les épitopes critiques pour trois cibles (CD147, CD209 et CD299). Les résultats ont montré une affinité élevée du domaine RBD du SARS-CoV-2 pour le récepteur CD299. Le CD36 et CD299 représente, donc, une cible puissante pour la découverte et la conception d'un nouveau groupe thérapeutiques ciblant le dysfonctionnement du goût par le Covid 19 et neutralisant le SRAS-CoV-2.

## ABSTRACT

**Title:** Study of the mechanism of interaction and signaling of fatty acids via CD36 with and without COVID 19 infection: Development of Workflow for the identification of bioactive molecules *in silico*

**Author:** Jihane Akachar

**Keywords:** Fatty acid, CD36, Molecular dynamics, Molecular modeling, Bioinformatics, Molecule design, COVID19.

Excessive consumption of lipids, in particular Saturated Fatty Acids, associated with a qualitative imbalance contributes to the prevalence of obesity. The CD36 lipid receptor is a plausible candidate as a taste lipid sensor, also inducing a cellular signaling cascade. The aim of this thesis was the design, screening and identification of new bioactive molecules through the development of workflows based on the *in silico* approach. We used series of molecular dynamics simulations of the CD36 protein, in interaction with one to three palmitic acids (PLM), in order to establish a first workflow to discover the mechanism of the functioning of the CD36/fatty acid complex and to determine the structural role of Lys 164. The results showed that Lys164 has an important influence on the form of stability of CD36 and. The design thanks to the results of the first workflow of 15 molecules activating the signaling pathway of fatty acids and more particularly those which are specific to the taste of fat. Only the CIN4 analog subsequently showed, by dynamic screening, a dimerization of the two TM1 and TM2 domains which resembles palmitic acid. In addition, the steps of this first workflow made it possible to explore the gustatory dysfunction of fatty acid chemoreception specifically the "fat taste" of patients who tested positive for COVID-19. The results of this study showed high affinity and stability of the non-mutant SARS-CoV-2 RBD domain at the CD36 receptor, a second workflow was established. Virtual screening was first used to identify critical epitopes for three targets (CD147, CD209 and CD299). The results showed high affinity of the SARS-CoV-2 RBD domain for the CD299 receptor. CD36 and CD299 therefore represent a powerful target for the discovery and design of a new therapeutic group targeting taste dysfunction by Covid 19 and neutralizing SARS-CoV-2.

# ملخص

**العنوان:** دراسة آلية تفاعل وإشارات الأحماض الدهنية عبر CD36 مع وبدون عدوى COVID 19 : تطوير سير العمل لتحديد

الجزينات النشطة بيولوجيًا في السيليكو

**تأليف:** جيهان أكشر

**الكلمات المفتاحية:** الأحماض الدهنية ، CD36 ، الديناميات الجزيئية ، النمذجة الجزيئية ، المعلوماتية الحيوية ، تصميم الجزينات ، COVID19.

يساهم الاستهلاك المفرط للدهون ، وخاصة الأحماض الدهنية المشبعة ، المرتبط باختلال التوازن النوعي في انتشار السمنة. يُعد مستقبل CD36 للدهون مرشحًا معقولاً كمستشعر طعم للدهون ، كما أنه يحفز سلسلة إشارات خلوية. كان الهدف من هذه الأطروحة هو تصميم وفحص وتحديد الجزينات النشطة بيولوجيًا الجديدة من خلال تطوير تدفقات العمل بناءً على نهج *in silico*. استخدمنا سلسلة من عمليات محاكاة الديناميكيات الجزيئية لبروتين CD36 ، بالتفاعل مع واحد إلى ثلاثة أحماض نخلية (PLM) ، من أجل إنشاء سير عمل أول لاكتشاف آلية عمل مركب CD36 / الأحماض الدهنية ولتحديد التركيب الهيكلي. دور Lys 164. أظهرت النتائج أن Lys164 له تأثير مهم على شكل استقرار CD36 و. التصميم بفضل نتائج سير العمل الأول المكون من 15 جزيءًا لتنشيط مسار إشارات الأحماض الدهنية وبشكل أكثر تحديدًا تلك الخاصة بمذاق الدهون. أظهر التناظرية CIN4 فقط لاحقًا ، من خلال الفحص الديناميكي ، إضعاف نطاقي TM1 و TM2 اللذين يشبهان حمض البالميتيك. بالإضافة إلى ذلك ، أتاحت خطوات سير العمل الأول هذا استكشاف الخلل الوظيفي الذوقي في الاستقبال الكيميائي للأحماض الدهنية على وجه التحديد "طعم الدهون" للمرضى الذين ثبتت إصابتهم بـ COVID-19. أظهرت نتائج هذه الدراسة تقاربيًا عاليًا واستقرارًا لمجال SARS-CoV-2 RBD 2 غير المتحور في مستقبل CD36 ، وتم إنشاء سير عمل ثانٍ. تم استخدام الفحص الافتراضي لأول مرة لتحديد الحلقات الحرجة لثلاثة أهداف (CD147 و CD209 و CD299). أظهرت النتائج تقاربيًا كبيرًا لمجال SARS-CoV-2 RBD لمستقبل CD299. لذلك يمثل CD36 و CD299 هدفًا قويًا لاكتشاف وتصميم مجموعة علاجية جديدة تستهدف اختلال الذوق بواسطة Covid 19 وتحييد SARS-CoV-2.

## SOMMAIRE

<b>CHAPITRE I : Revue Bibliographique</b> .....	19
I.1.Définition du récepteur CD36.....	20
I.1.1.Historique de la découverte du CD36.....	22
I.1.2.Structure et localisation du gène CD36 .....	22
I.1.3.Structure du récepteur CD36 et les ligands .....	23
I.1.4.Distribution cellulaire et régulation de l'expression du CD36.....	24
I.1.5.Modification post-traditionnelles.....	25
I.1.6. Rôles du CD36 dans la perception sensorielle lipidique et l'Obésité.....	25
I.1.6.1. Les lipides alimentaires .....	26
I.1.6.2.Rôle du CD36 dans la préférence spontanée aux lipides alimentaires et signalisation cellulaire.....	31
I.1.6.3.CD36, senseur de lipides au niveau hypothalamique : impact sur la régulation de la prise alimentaire.....	32
I.1.7.Rôles pathologiques du CD36 dans les infections par malaria et Covid 19.....	34
I.1.8.Acides gras et les sites de liaison du récepteur CD36 .....	34
I.1.8.1. Identification d'une poche de liaison des AGLC dans CD36.....	34
I.1.8.2.L'importance de la Lysine-164 dans la fixation et l'induction de la signalisation des acides gras .....	35
I.1.8.3.L'internalisation des AGLC par l'accès dans un tunnel au sein du CD36 .....	36
I.1.8.4.Lesite de liaison des AG chevauche avec ceux des oxLDL et de la phosphatidylcholine oxydée .....	37
I.2. Matériel et méthodes : Etudes <i>in silico</i> .....	39
I.2.1.L'approche <i>in silico</i> dans le domaine de la biologie (Bio-Informatique) .....	40
I.2.1.1. Modélisation moléculaire (Modélisation par homologie) .....	40
I.2.1.2.Conception et découverte de molécules assistées par ordinateur (Drug design) .....	41
I.2.1.3.Amarrage moléculaire (Docking) .....	44
I.2.1.4.La Dynamique Moléculaire (Molecular Dynamics).....	44
<b>CHAPITRE II: The computational analyses, molecular dynamics of fatty-acid transport mechanism to the CD36 receptor. Nature Journal of Scientific Report.</b> .....	48
II.1.Résumé de l'article .....	49
II.2.Article original.....	51
<b>CHAPITRE III: In silico Method for Design and Discovery a New Compound as Potential of the Fat-taste Analogues.</b> .....	65
III.1.Résumé de l'article.....	66
III.2.Article original.....	<b>Error! Bookmark not defined.</b>
<b>CHAPITRE IV: Analyse <i>in silico</i> du mécanisme structurel des troubles du goût des acides gras induits par le COVID-19 via le récepteur CD36</b> .....	70
IV.1.Résumé de l'article .....	98

IV.2.Article original .....	<b>Error! Bookmark not defined.</b>
<b>CHAPITRE V: Identifying epitopes for cluster of differentiation and design of new peptides inhibitors against human SARS-CoV-2 spike RBD by an <i>in-silico</i> approach.</b> .....	<b>111</b>
V.1.Résumé de l'article.....	120
V.2.Article original.....	122
<b>DISCUSSION GENERALE ET PERSPECTIVES</b> .....	<b>135</b>
<b>CONCLUSION GENERALE</b> .....	<b>Error! Bookmark not defined.</b>
<b>REFERENCES BIBLIOGRAPHIQUES</b> .....	<b>141</b>



## LISTE DES ABREVIATIONS

<b>AA</b>	Acide aminé
<b>AG</b>	Acides gras
<b>AGLC</b>	Acides gras à longue chaîne
<b>AGS</b>	Acides Gras Saturés
<b>ACS</b>	AcylCoA Synthetase
<b>AGL</b>	Acides Gras Libres
<b>AgrP</b>	Agouti Related Peptide
<b>AGL</b>	Acides Gras Libres
<b>AGCC</b>	Acides Gras à Chaîne Courte
<b>AGCM</b>	Acides Gras à chaîne moyenne
<b>AGMI</b>	Acides Gras Mono-Insaturés
<b>AGPI</b>	Acides Gras Polyinsaturés
<b>AMPc</b>	Adénosine Monophosphate Cyclique
<b>acylCoA</b>	Acyl-coenzyme A,
<b>ACS</b>	Noyau du Tractus Tolitaire
<b>AGS</b>	Acides Gras Saturés
<b>AGE</b>	Acides Gras de glycosylation
<b>AGPI</b>	Acides Gras Polyinsaturés
<b>ADMET</b>	Absorption, Distribution, Metabolism, Excretion, and Toxicity
<b>AND</b>	Acide Désoxyribo Nucléique (ADN)
<b>ARNm</b>	Acide Ribonucléique messenger
<b>ApoB48</b>	Apolipoprotein B
<b>AMBER</b>	Modèle assisté immeuble avec l'énergie Raffinement
<b>CD36</b>	Cluster of Cifferenciation 36
<b>CPT1</b>	Carnitine Palmitoyl Transferase 1
<b>COVID 19</b>	Corona Virus Disease 2019
<b>CHARMM</b>	Chemistry at HARvard Macromolecular Mechanics
<b>CPT1</b>	Carnitine Palmitoyl Transferase I
<b>CCK</b>	Cholécystokinine
<b>CHO</b>	Hamster Chinois
<b>DM</b>	Dynamique Moléculaire
<b>ER</b>	Réticulum endoplasmique
<b>FAT</b>	Fatty Fcid Translocase

<b>FABPpm</b>	Fatty Acid-Binding Protein plasma membrane
<b>FATP4</b>	Fatty Acid Transfer Protein 4
<b>FAT</b>	Fatty Acid Transporter
<b>GPR</b>	G-protéine récepteur couplée
<b>GCase</b>	Glucocérébrosidase
<b>GPVI</b>	Glycoprotéine IV
<b>GPIV</b>	Glycoprotein VI
<b>GPIIIb</b>	Glycoprotéine IIIb
<b>LIMP2</b>	Lysosomal Integral membran protein 2
<b>LPL</b>	Lipo Protein Lipase
<b>LBP</b>	Lipid-Binding Proteins
<b>LBDD</b>	Ligand-Based Drug Design
<b>MCV</b>	Maladies Cardiovasculaires
<b>ERK1</b>	Extracellular Signal-Regulated Kinases 1
<b>MTP</b>	Microsomal triglyceride transfer protein
<b>MG</b>	Monoglycéride
<b>NHS</b>	N-hydroxysuccinimidyle
<b>oxLDL</b>	Lipoprotéines de basse densité oxydées
<b>OEA</b>	Oléoyléthanolamide
<b>PDB</b>	Protein Data Bank
<b>POMC</b>	Melano Cortin
<b>PPAR</b>	Peroxisome Proliferator Activated Receptor
<b>PCox</b>	Phospholipide oxydé
<b>POMC</b>	Pro Opio Melano Cortin
<b>QSAR</b>	Relation Quantitative Structure-Activité
<b>RMSD</b>	Root Mean Square Deviation
<b>Rg</b>	Rayon de Giration
<b>RMSF</b>	Root Mean Square Fluctuations
<b>RMN</b>	Résonance Magnétique Nucléaire
<b>SOC</b>	Store-Operated Calcium Channels
<b>SSO</b>	Sulfo-N-Succinimidyle-Oléat
<b>SR-BII</b>	Scavenger Receptor class B membrane protein-2
<b>TG</b>	Triglyceride
<b>TM</b>	Transmembrane
<b>VLDL</b>	Lipoprotéines de très basse densité

## LISTE DES FIGURES

**Figure 1** : Rôle du CD36 dans l'absorption intestinale des lipides.

**Figure 2** : Cartographie cytogénétique montrant la position du gène codant pour la protéine CD36 (Genbank)

**Figure 3** : Structure du gène CD36 (GenBank)

**Figure 4** : Structure de CD36, une glycoprotéine transmembranaire aux multiples ligands et fonctions.

**Figure 5** : Représentation schématique des domaines de liaison et les modifications post-traditionnelles.

**Figure 6** : Représentation schématique de structure d'un phospholipide

**Figure 7** : La nomenclature des acides gras

**Figure 8** : Les lipides alimentaires : apports alimentaires et rôles cellulaires.

**Figure 9** : Principales étapes de l'absorption intestinale des acides gras à longue chaînes (AGCL).

**Figure 10** : Rôle du CD36 dans la détection orale des lipides alimentaires.

**Figure 11** : Rôle de CD36 dans la détection des lipides et la régulation de la prise alimentaire au niveau central.

**Figure 12** : Le site de fixation des acides gras sur CD36 et leur accès au tunnel

**Schéma 13** : Methods la détermination du site de liaison sur la cible thérapeutique

**Figure 14** : Liaison hydrogène : cas de deux molécules identiques de H<sub>2</sub>O

# Introduction

Une consommation excessive de lipides, en particulier d'acides gras saturés (AGS), est associée à des déséquilibres qualitatifs et contribue à l'épidémie d'obésité. Une corrélation a également été trouvée entre la quantité de graisse consommée et le poids corporel chez l'Homme (Bray *et al.*, 2004) et chez la souris (Wit *et al.*, 2011). Il est associé à des troubles métaboliques telles que l'insulino-résistance, l'intolérance au glucose, des dyslipidémies, l'hypertension, qui constituent des facteurs de risque des maladies cardiovasculaires (MCV). Bien que l'origine de cette épidémie soit clairement multifactorielle, le bouleversement récent de notre alimentation est la principale cause de ce phénomène. L'une des conséquences d'une alimentation riche est qu'elle encourage la consommation d'aliments " plaisir ", généralement riches en sucre et en matières grasses (Simmons *et al.*, 2005). Comprendre les mécanismes moléculaires à l'origine de la régulation des mécanismes d'absorption des lipides représente donc une approche pertinente pour développer de nouvelles approches thérapeutiques pour lutter contre l'obésité.

Au cours des deux dernières décennies, plusieurs progrès ont été réalisés dans le transport des acides gras et plusieurs protéines intestinales, regroupées sous le terme générique de lipid-binding protéine (LBP) présentant une haute affinité pour les Acide gras à longue chaîne (AGLC), ont été clonées. Il s'agit de la Plasma membrane fatty acid-binding Protein ou FABPpm (Stremmel *et al.*, 1985), du fatty acid transporter ou FAT/CD36 (Abumrad *et al.*, 1993) et enfin du fatty acid transfer protéine 4 ou FATP4 (Abumrad *et al.*, 1993). La FABPpm a été la première LBP membranaire à avoir été isolée (Stremmel *et al.*, 1985). En ce qui concerne leur rôle dans le transport et l'absorption des AGLC nos connaissances restent encore limitées, bien que plusieurs chercheurs attribuaient le transport à une simple diffusion. (Shiau *et al.*, 1985) Plusieurs données suggèrent l'implication du CD36 dans le transport et le captage des AGLC. Le fatty acid transporter (FAT) est une glycoprotéine de 88 kDa (Abumrad *et al.*, 1993), qui présente une structure en épingle à cheveux avec un large domaine extracellulaire comportant une poche hydrophobe pouvant accueillir jusqu'à 3 AGLC ainsi que de nombreuses autres molécules comme les LDL oxydés, phospholipides anioniques ou la thrombospondine (Cechetto *et al.*, 2002) . Cette protéine est trouvée dans les tissus caractérisés par un métabolisme lipidique intense. Son implication dans le transfert intra-entérocytaire des AGLC a été suggérée par plusieurs arguments. Tout d'abord, le FAT/CD36 lie avec une affinité de l'ordre du nanomolaire les AGLC (Cechetto *et al.*, 2002 ; Ibrahimi *et al.*, 2002). Ensuite, il a été démontré de façon convaincante qu'il est impliqué dans le captage des AGLC par le tissu adipeux et les muscles striés, squelettiques et cardiaques (Ibrahimi *et al.*, 2002). Bien qu'il a été décrit dans la physiopathologie cardiovasculaire, comme l'athérosclérose,

(Kuang *et al.*, 2004) le métabolisme lipidique, (Koonen *et al.*, 2005), la maladie de diabète, (Corpeleijn *et al.*, 2006), l'angiogenèse, (Simantovet *et al.*, 2005) la biologie du cancer, (Huang *et al.*, 2004) l'ischémie cérébrale, (Cho *et al.*, 2005) la maladie d'Alzheimer (Moore *et al.*, 2002), CD36 est impliqué dans la préférence spontanée aux lipides, puisque les souris CD36 (-/-) perdent la préférence spontanée pour des solutions contenant des lipides. Il a été montré que cette détection des lipides médiée par CD36 passe par la voie gustative. En effet, CD36 est exprimé en particulier au niveau des papilles caliciformes de souris, de rat et d'Homme ((Drover *et al.*, 2008) ; (Nauli *et al.*, 2006); (Press *et al.*, 2022)). Récemment, dans une série de 81 sujets testés positifs pour le variant SARS-CoV-2 Omicron, l'altération de l'odorat et du goût a été autodéclarée par 12 % et 23 % des patients, respectivement (Press, 2022). L'identification récente de récepteurs lipidiques (par exemple CD36) est un candidat plausible en tant que capteur lipidique gustatif, également connu sous le nom de transporteur d'acides gras.

Aujourd'hui, la multitude de ses ligands et la découverte plus récente de ses capacités à induire différentes voies de signalisation cellulaire laissent supposer de faire émerger l'existence de d'un chemin spécifique de CD36 au transport lipidique. Le récepteur cristallographique CD36 est capturé dans la structure à l'état ouvert sans extrémité N-terminale et c-terminale (TM) (Hsieh *et al.*, 2016).

La modélisation et dynamique moléculaire sont devenues un outil essentiel pour la détermination de la structure tridimensionnelle des protéines et comprendre le mécanisme moléculaire, et les interactions mis en jeu dans la fixation des acides gras sur la protéine CD36. Notre laboratoire s'intéresse bien à la protéine FAT, et plusieurs travaux ont été réalisés dans ce contexte (Sabaouni *et al.*, 2013 ; Zineb *et al.*, 2015). Les travaux de recherche de notre thèse ont rentré dans le cadre de la continuité des études déjà réalisés sur ces protéines. Dans ce contexte et en collaboration avec le laboratoire physiologie de la nutrition et toxicologie, Dijon, et le laboratoire de Dynamique des Structures et Interactions des Macromolécules Biologiques (DSIMB) Paris, France, et le laboratoire de chimie thérapeutique Faculté de médecine et pharmacie, Rabat, Maroc, l'objectif de notre travail de thèse est de déterminer des mécanismes d'interaction et de signalisation des acides gras via CD36 avec et sans infection COVID-19 dans le but de développer des workflow basés sur l'approche *in silico* par modélisation et dynamique moléculaire nécessaires pour la conception, le criblage et l'identification de nouvelles molécules bioactives.

Le manuscrit est organisé de la façon suivante :

- Un premier chapitre présente une revue bibliographique sur la protéine CD36 et l'analyse bioinformatique *in silico*.
- Un second chapitre concerne le développement de workflow d'analyse du mécanisme de transport des acides gras via le récepteur CD36 en temps réel par dynamique moléculaire.
- Un troisième chapitre décrit le développement d'une méthode d'analyse *in silico* du mécanisme de signalisation des acides gras via le récepteur CD36 et la conception de nouveaux analogues photochimiques naturels du goût gras
- Le chapitre 4 présente l'analyse *in silico* du mécanisme structurel des troubles du goût des acides gras induits par le COVID-19 via le récepteur CD36
- Le chapitre 5 illustre la méthode de modélisation moléculaire utilisée pour identification des épitopes pour un cluster de différenciation et conception de nouveaux inhibiteurs de peptides contre la RBD du pic de SARS-CoV-2 humain
- Discussion des résultats et aux perspectives ouvertes de ce travail.
- Conclusion générale

# **CHAPITRE I : Revue Bibliographique**

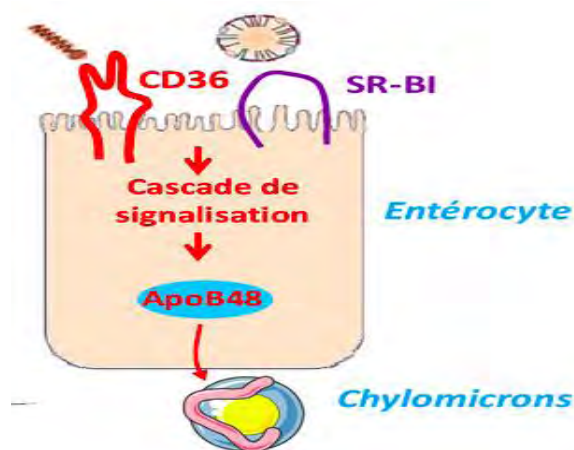
## CHAPITRE I : Revue Bibliographique

### I.1.Définition du récepteur CD36

La Protéine CD36 est une glycoprotéine multifonctionnelle capable de lier les acide gras a longue chaine (AGLC) saturés ou non saturés, ionisés, avec une affinité de l'ordre du nanomolaire.Elle est localisée principalement au niveau des membranes des bordures en brosse des deux tiers supérieurs des villosités dans la partie proximale de l'intestin grêle (Silverstein *et al.*, 2009). De nombreux chercheurs ont montré que l'expression intestinale de la CD36 est étroitement corrélée avec la teneur en lipides du régime ; où un régime pauvre en lipides diminue son expression et un régime hyper lipidique chronique l'induit fortement (Lobo *et al.*, 2001 ; Poirier *et al.*, 1996). Cependant, à la différence de l'adipocyte et du myocyte, la contribution significative de la CD36 dans le captage entérocytaire des AGLC a été mise en doute (Sukhotnik *et al.*, 2001). En dépit de la contribution de cette dernière, une recherche a montré que la délétion du gène du CD36 peut diminuer le captage des AGLC par les entérocytes isolés dépourvus de couche d'eau non agitée ou n'a pas d'effet (Nassir *et al.*, 2007). Une classe des travaux a démonté la pertinence physiologique de ces données ; puisque des anses jéjunales isolées in situ de souris sauvages et déficientes en CD36 qui conservent la couche d'eau non agitée et la dynamique de l'absorption, ont la même efficacité de captage (Drover *et al.*, 2005). De nombreux recherches ont étudié pourquoi l'invalidation du gène de CD36 chez la souris ne se traduit pas *in vivo* par une modification de l'absorption nette des AGLC excepté pour des acides gras à très longues chaînes rares dans l'alimentation (Figure 1) ((Nassir *et al.*, 2007) ; (Drover *et al.*, 2008) ; (Nauli *et al.*, 2006); (Tran *et al.*, 2011)).

Cependant, le CD36 participe au mécanisme d'absorption des lipides alimentaires puisque sa déficience réduit considérablement la sécrétion lymphatique des chylomicrons qui sont moins riches en TG et de plus petites tailles. Ces effets sont associés à une rétention de TG au niveau entérocytaire en régime hyperlipidique chronique (Masuda *et al.*, 2009). Ces données démontrent que la CD36 joue un rôle très important dans la formation des chylomicrons, mais pas comme transporteur clé et efficace des lipides (Drover *et al.*, 2005).





**Figure 1 :** Rôle du CD36 dans l'absorption intestinale des lipides (Tran *et al.*, 2012).

La Figure 1 montre :

- 1 : liaison des acides gras à longue chaînes (AGLC) issus de la digestion des triglycérides alimentaires avec le CD36;
- 2: Cascade de signalisation par activation des ERK1/2 ;
- 3 et 4 : le CD36 via l'activation des ERK1/2, déclenche l'induction de l'ApoB48, de la MTP et de la L-FABP, trois protéines qui optimisent la synthèse des chylomicrons ;
- 5 : les AGLC déclenchent une dégradation du CD36 par la voie ubiquitine-protéasome : désensibilisation.

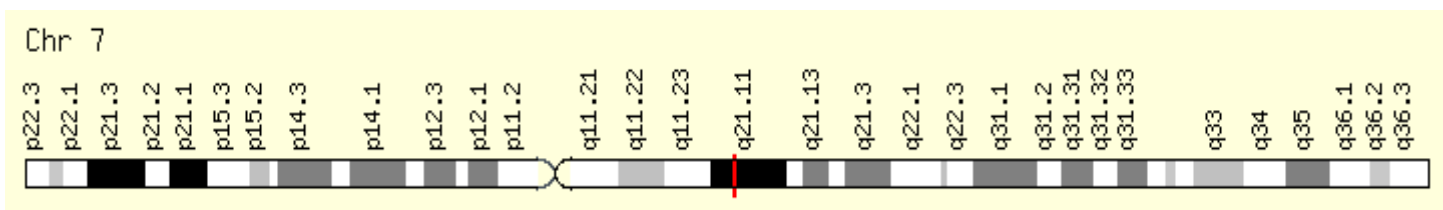
L'existence de la couche d'eau non agitée, en plus des propriétés structurales et fonctionnelles des LBP membranaires, indique que la contribution d'un transport facilité dans le captage intestinal des acides gras alimentaires n'est probablement pas prépondérante en période postprandiale. En revanche, il est important de noter que l'efficacité de la diffusion passive est conditionnée par le maintien d'un gradient de concentration en AGLC entre les milieux extra- et intracellulaires dépendant de l'efficacité du métabolisme entérocytaire des AGLC. Cependant, pour des concentrations plus faibles, en situation interprandiale par exemple, la diffusion facilitée pourrait s'avérer plus efficace, du fait d'une plus forte affinité pour les AGLC. Le transport à forte capacité et à faible affinité est particulièrement bien adapté et éviterait ainsi que le captage des AGLC ne devienne un facteur limitant dans l'absorption intestinale des lipides après un repas (Chow *et al.*, 1979).

### I.1.1. Historique de la découverte du CD36

Le récepteur CD36 fut initialement identifié à partir de membranes de plaquettes sanguines à la fin des années 1970. Un premier groupe décrivit le CD36, alors nommé glycoprotéine IV (GPIV), comme une protéine de 87 kDa résistante à la protéolyse lorsque des plaquettes humaines étaient digérées par la trypsine et la chymotrypsine. Le récepteur CD36 a par la suite été identifié par un autre groupe comme étant l'antigène réagissant avec l'anticorps OKM5 et fut alors nommé antigène OKM5 (Okumura *et al.*, 1976). L'immuno-précipitation de l'antigène OKM5 des monocytes a mené à la caractérisation d'une protéine d'environ 88 kDa, soit un poids moléculaire similaire à la GPIV (Talle *et al.*, 1983). Les termes GPIV et antigène OKM5 furent remplacés par CD36 (Okumura *et al.*, 1976). Cependant le récepteur porta d'autres noms, notamment FAT (fattyacid translocase), GPIIb (glycoprotéine IIIb) et PAS IV (Poirier *et al.*, 1996).

### I.1.2. Structure et localisation du gène CD36

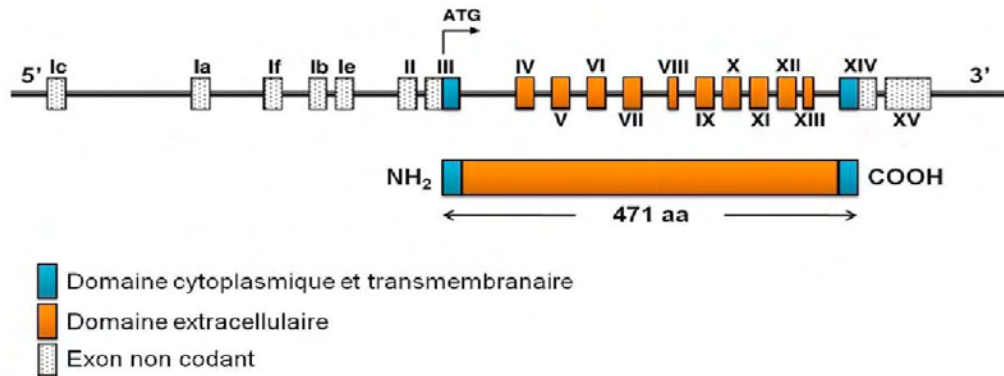
CD36 est un gène du 77089 paire de bases, sa localisation cytogénétique est : 7q21.11 précisément dans le bras long du chromosome 7 position 11.2. bien que sa localisation moléculaire sur l'ADN : du 80,602,188 au 80,679,277 paires de bases (Figure 2) (NCBI) (Armesilla *et al.*, 1994)..



**Figure 2** : cartographie cytogénétique montrant la position du gène codant pour la protéine CD36 (Genbank) (<https://www.genecards.org/cgi-bin/carddisp.pl?gene=CD36>) (Armesilla *et al.*, 1994)

CD36 est encodée par 15 exons dont (Figure 3):

- Exons 3 et 14 codent respectivement pour NH-Terminale et COOH-Terminale.
- La région 5'-non traduite de l'ARNm est encodée par 3 exons.
- Exon 3, contient 89 nucléotides de la région 5'-non traduite et encode aussi pour l'extrémité NH-Terminale et domaine transmembranaire (Armesilla *et al.*, 1996).
- La région 3'-non traduite est contenue dans l'exon 14 seul ou 14 et 15 (Armesilla *et al.*, 1994).



**Figure 3:** structure du gène CD36 (GenBank) (Armesilla *et al.*, 1994).

Ce gène est aussi appelé :

- BDPLT10
- CHDS7
- FAT
- GP3B
- GP4s
- GPIV
- PASIV
- SCARB3

### I.1.3. Structure du récepteur CD36 et les ligands

CD36 est constitué de 472 acides aminés, il appartient à une classe de protéines qui partagent des caractéristiques structurales communes, qui comprennent une épingle à cheveux, une topologie membranaire avec deux domaines transmembranaires et avec deux extrémités dans le cytoplasme (Figure 4) (Canton *et al.*, 2013). Le domaine extracellulaire du CD36 fortement glycosylé comprend trois ponts disulfures dans la moitié carboxy-terminale qui ont été montrées importantes pour le recrutement membranaire du CD36. La partie amino-terminale contient des domaines de liaison pour hexaréline, AG, le LDL oxydé ou des phospholipides, la thrombospondine, et les érythrocytes infectés par *P.falciparum* (Canton *et al.*, 2013). Un segment de résidus d'acides aminés hydrophobes (186-204) pourrait faire une boucle dans la partie externe de la membrane, (Coburn *et al.*, 2003) et pourrait faire partie d'une poche de liaison pour les ligands lipidiques de CD36, tel que discuté dans la section suivante. Les queues de la protéine sont cytoplasmiques et la queue courte, mais carboxylique est actif dans la transduction du signal à travers sa capacité à interagir avec un certain nombre de tyrosine – kinases (Canton *et al.*, 2013).

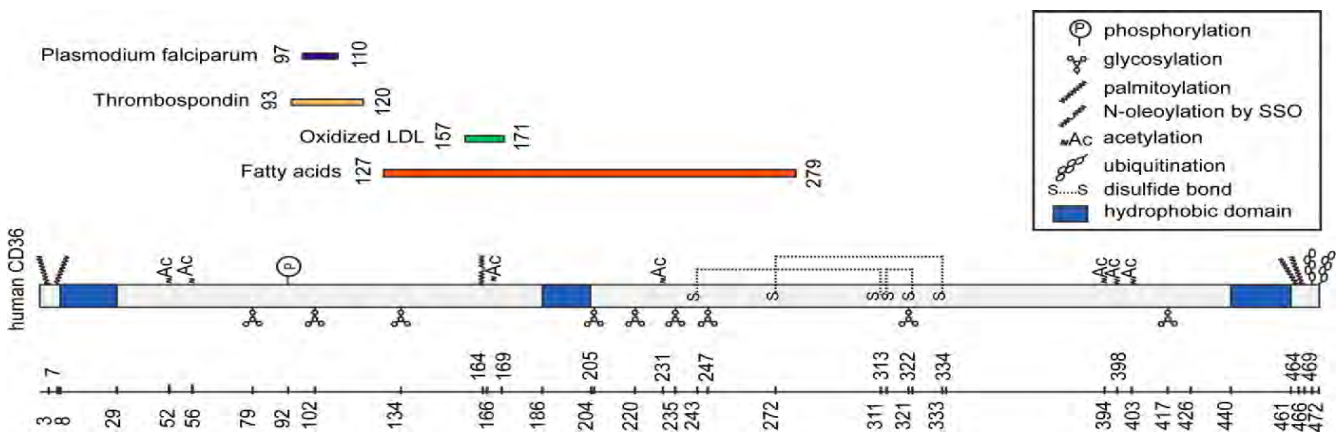


Les contributions relatives des régulations transcriptionnelle et traductionnelle dans l'expression du CD36 ne sont actuellement pas clairement définies (Yesner *et al.*, 1995).

Des études récentes de la région promotrice du gène CD36 ont mis en évidence la présence de séquences caractéristiques d'éléments régulateurs de la transcription en 5' de l'exon 1 (PEBP2/CBF) (Amersilla *et al.*, 1996), mais aussi internes à l'intron 1 (Asch *et al.*, 1996). Cependant, seules ces dernières possèdent des caractéristiques érythroïdes et mégacaryocytiques et notamment des séquences consensus GATA. L'utilisation alternative de ces deux promoteurs pourrait expliquer certains aspects de la régulation de ce gène. Enfin, le CD36 est exprimé par les adipocytes, les cellules épithéliales mammaires, les kératinocytes et certaines lignées de cellules tumorales (Asch *et al.*, 1996).

### I.1.5. Modification post-traditionnelles

CD36 est une protéine subissant à des plusieurs modifications post-traditionnelles, y compris la phosphorylation, la glycosylation, palmitoylation, ou ubiquitination modulent les niveaux ou le trafic de CD36 et par conséquent, ses fonctions (**Figure 5**)



**Figure 5** : Représentation schématique des domaines de liaison et les modifications post-traditionnelles (Pepino *et al.*, 2014).

### I.1.6. Rôles du CD36 dans la perception sensorielle lipidique et l'Obésité

L'épidémie d'obésité pose un grave problème à l'échelle mondiale menace pour la santé humaine et le monde entier qui ne peut pas plus être ignoré. Les facteurs de la génétique, de l'environnement et les habitudes alimentaires sont impliqués dans le développement de l'obésité. Les graisses alimentaires sont consommées en grande quantité par les sujets obèses en raison de ses indices

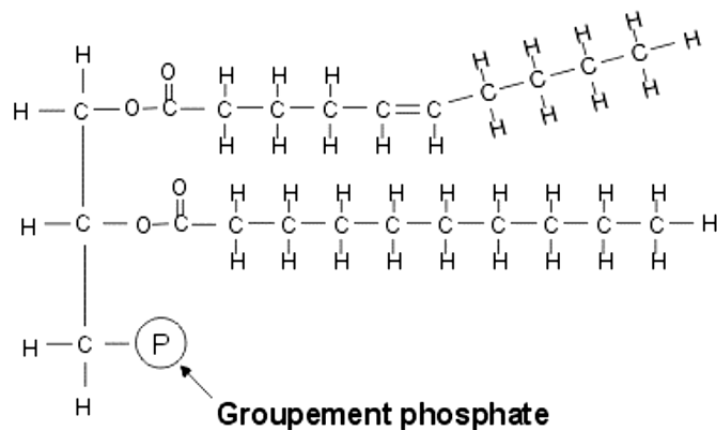
olfactifs, visuels et texturaux. Croissance les preuves suggèrent que la détection de la graisse buccale joue un rôle important rôle dans le développement de l'obésité. Sur la base des données de corrélation entre CD36 polymorphismes et l'obésité, on a émis l'hypothèse que CD36 peut jouer un rôle important dans la perception de la graisse buccale et préférences individuelles pour les aliments gras. (Djeziri *et al.*, 2018)

### **I.1.6.1. Les lipides alimentaires**

Les graisses alimentaires sont consommées en grande quantité par les sujets obèses, cependant leur consommation, en accord avec les Apports Nutritionnels Conseillés, est essentielle au maintien d'un bon état de santé. Ce chapitre sera consacré aux lipides alimentaires et à leurs différents rôles au sein de l'organisme. Les lipides jouent un rôle essentiel dans la composition des membranes des cellules la fabrication des hormones, le métabolisme cellulaire, le transport et le stockage de certaines vitamines, ils représentent une source importante d'énergie. L'alimentation apporte de 30 à 40 % de l'énergie sous forme de lipides, principalement environ 95% des triglycérides (TG) (Guesnet *et al.*, 2005). Du fait de leur caractère hydrophobe, l'absorption des TG au niveau intestinal est un processus complexe comportant plusieurs étapes (Niot *et al.*, 2009). 4.5% des phospholipides (PL) et les 0.5% restant sont les stérols, essentiellement du cholestérol retrouvé sous forme d'ester. On peut aussi retrouver dans les aliments quelque acides gras libres (AGL).

#### **I.1.6.1.1 .Structure des lipides alimentaires**

Les lipides alimentaires sont très majoritairement sous forme de triglycérides sont les plus abondants, sont des esters constitués de l'union de trois acides gras (AG) et d'une molécule de glycérol. Ils sont les principaux constituants des huiles et graisses végétales qu'animales. Une caractéristique de ces TG sont totalement apolaires et stable forme la classe des graisses neutres la plus importante. On retrouve principalement stockés dans les cellules adipeuses (la graisse corporelle), de phospholipides, sont la plupart des phosphoglycérides amphipathiques dont les principaux constituants sont des molécules avec une tête polaire et de deux queues apolaires d'hydrocarbures hydrophobes. Ils sont formés le plus souvent d'un résidu glycérol lié à un groupement phosphate et à deux acides gras (Figure 6). On retrouve plus précisément des phosphatidylcholines ou des phosphatidyléthanolamines.



**Figure 6** : Représentation schématique de structure d'un phospholipide

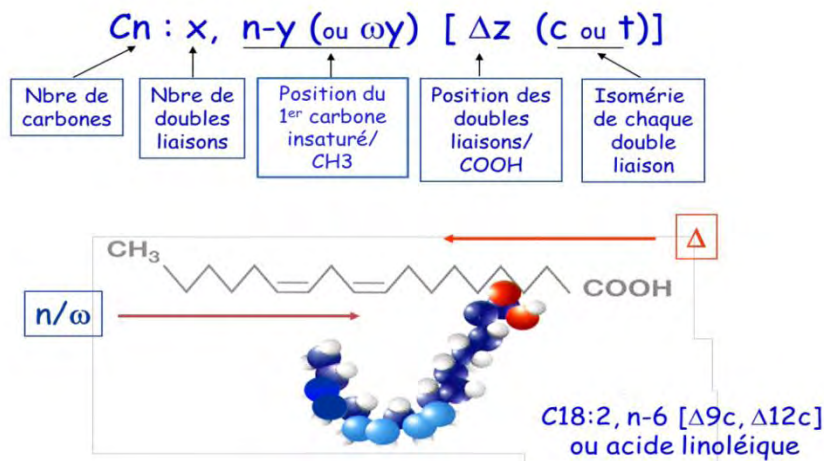
(Source : <https://www.medical-actu.com/lipides/>)

En milieu aqueux, les PL vont s'organiser de telle sorte que seule la partie hydrophile sera en contact avec les molécules d'eau ce qui aboutit à des structures en micelles, en liposomes ou en bicouche lipidique. Ainsi, les PL représentent la base constituants des membranes cellulaires. Les stérols sont des lipides contenant un noyau cyclopentano-phénanthrénique dans le groupe duquel se trouvent le cholestérol, d'origine animale, et les phytostérols, d'origine végétale, sous forme d'esters d'acides gras. On les trouve en petite quantité dans certaines céréales, légumes et fruits. Le représentant principal de cette classe est le cholestérol qui joue un rôle essentiel dans l'organisme et le fonctionnement cellulaire et est un précurseur de multiples molécules telles que des vitamines et d'hormones.

Ainsi, les acides gras qui sont apportés par l'alimentation sont les constituants élémentaires les plus essentiels au niveau de l'organisme. Ils constituent une chaîne hydrocarbonée comportant un groupement carboxyle COOH à une extrémité et un groupement méthyle CH<sub>3</sub> à l'autre extrémité. Ainsi cet ensemble sont classés selon différents critères. Ils se définissent par leur nombre de carbones sur la chaîne carbonée leur nombre et position d'insaturation (Figure 7). Les acides gras à chaîne courte (AGCC) contiennent d'un maximum de six atomes de carbone, les acides gras à chaîne moyenne (AGCM) composées de 8 à 14 d'atomes de carbones et enfin pour des chaînes carbonées supérieures à 14 atomes de carbone on parle d'acides gras à longue chaîne (AGLC). Les acides gras insaturés n'en contiennent aucune double liaison (insaturation), soit il ne comprend qu'une seule, on parle d'acides gras mono-insaturés (AGMI), ou soit il comporte au moins deux doubles liaisons et dans ce cas on parle d'acides gras polyinsaturés (AGPI). On distingue deux principales séries d'AGPI en fonction de leur degré d'insaturation avec le groupement méthyl de la



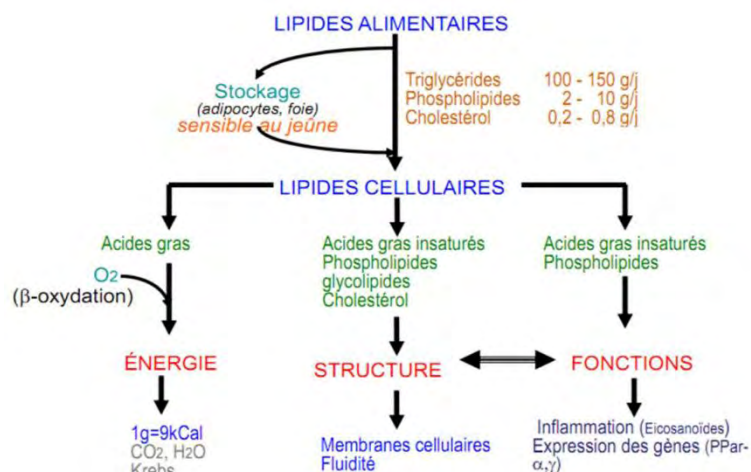
chaîne carbonée et position de leurs doubles liaisons. Les principaux AGPI alimentaire est l'acide  $\alpha$ -linoléique pour oméga3 (ALA, 18:2 n-3) et l'acide linoléique pour les oméga 6 (AL, 18:2, n-6). On parle d'acides gras essentiels ou indispensables car notre organisme étant incapable de les synthétiser tout seul. C'est pourquoila consommation d'aliments riches en oméga 3 et en oméga 6 est primordiale (Chevrot *et al.*, 2013).



**Figure 7.** La nomenclature des acides gras (Chevrot *et al.* 2013)

### I.1.6.1.2. Rôle des lipides alimentaires

Les lipides alimentaires jouent différents rôles cruciaux dans notre organisme, notamment des rôles structuraux mais également de régulation et énergétique (Figure 8).



**Figure 8:** Les lipides alimentaires : apports alimentaires et rôles cellulaires (Isabelle *et al.*, 2010)



### **I.1.6.1.3.Rôle structural**

Les lipides sont des molécules qui entrent dans la composition des membranes des cellules (fluidité, activité enzymatique...). En effet, les membranes cellulaires sont constituées d'une bicouche phospholipidique. Cette bicouche est composée des acides gras fondamentaux les acides linoléique, arachidonique et docosa-hexaénoïque. On trouve également une proportion non négligeable enrichis en acides gras saturés, glycosphingolipides et en cholestérol, auxquels sont associés des complexes protéiques actifs (Davletov *et al.*, 2010).

La fluidité membranaire est liée au degré d'insaturation des acides gras présents. De même, l'augmentation de la proportion de cholestérol peut réduire sa fluidité. Une membrane cellulaire fluide est beaucoup plus efficace pour l'activité de différents médiateurs cellulaires (hormones, eicosanoïdes, protéines etc.)(Davletov *et al.*, 2010).

### **I.1.6.1.4.Rôle régulateur**

Les lipides AGPI à 20 carbones sont les précurseurs de molécules à fonction biologique : les eicosanoïde. Les eicosanoïdes interviennent au niveau de la reproduction, la fonction rénale, l'activité neurale, la physiologie cardiaque, la fonction circulatoire locale, la coagulation sanguine, l'inflammation et la pression sanguine (Guesnet *et al*, 2005).

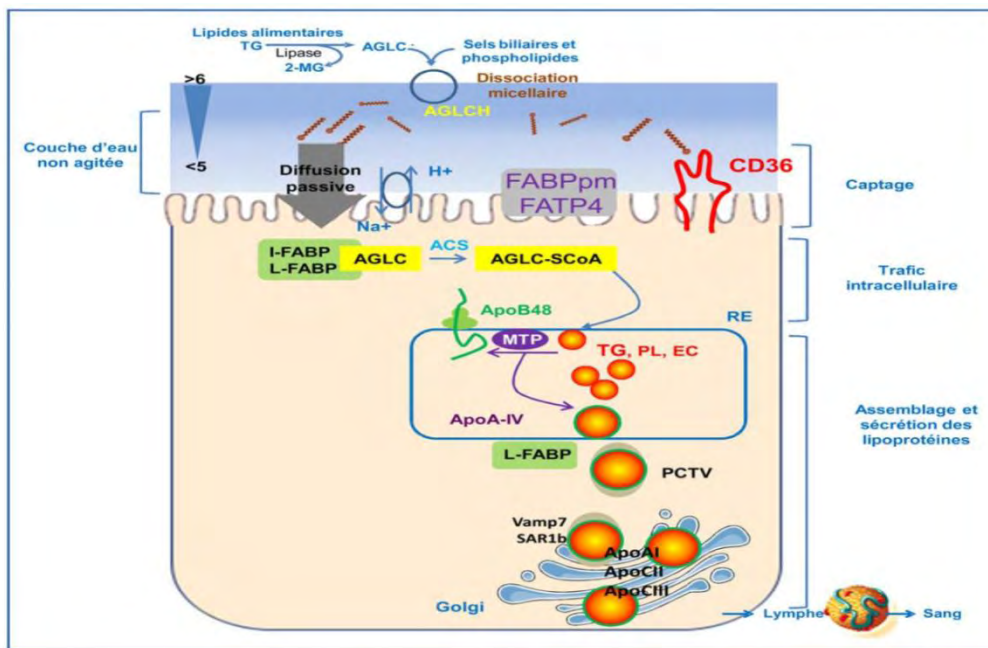
Leur production issue par deux voies, l'une issue de l'action de la cyclooxygénase donnant naissance aux prostaglandines, thromboxanes et prostacyclines et l'autre issue de l'action de la lipoxygénase donnant naissance aux leucotriènes. Ces eicosanoïdes sont impliqués dans la production de second messenger (AMPc ou  $Ca^{++}$ ), l'activation d'une protéine kinase, ou une modification du potentiel membranaire (Ben-Shlomo *et al*, 1991; Sumida *et al*, 1993). C'est également un autre élément important de lipide alimentaire. Il s'agit du cholestérol qui est un précurseur de nombreuses molécules de l'organisme. Il est impliqué dans la synthèse de nombreuses molécules stéroïdiennes comme la progestérone, la testostérone et les œstrogènes; les glucocorticoïdes (cortisone, cortisol) ou des sels biliaires ainsi la fabrication de substances comme la vitamine D. De plus, les AGLCalimentaire et plus particulièrement les AGPI jouent un rôle direct dans l'expression de divers gènes par l'activation ou la répression des facteurs de transcription (Kliwer *et al*, 1997). Cette induction transcriptionnelle s'effectue, notamment, par le biais des récepteurs nucléaires PPAR (peroxysome proliferatoractivatedreceptor) delta et gamma.

### I.1.6.1.5. Rôle énergétique

Les acides gras alimentaires par leur catabolisme par la  $\beta$ -oxydation mitochondriale fournissent une source très importante d'énergie pour l'organisme 1 g d'acides gras équivaut à 9 kilocalories. (Davletov *et al.*, 2010).

### I.1.6.1.6. Absorption et transport

À l'arrivée dans la lumière intestinale, les lipides alimentaires, principalement des TG, sont hydrolysés en monoglycéride (MG) et AGLC grâce aux lipases gastrique et pancréatique. Ces produits sont ensuite associés au cholestérol, aux phospholipides, aux sels biliaires et forment des micelles mixtes. Il est admis que les AGLC sont absorbés préférentiellement au niveau jéjunal (Guesnet, *et al.*, 2005) par les entérocytes hautement différenciés se trouvant dans les deux tiers supérieurs des villosités intestinales (Figure 9) (Niotet *al.*, 2009).



**Figure 9** : Principales étapes de l'absorption intestinale des acides gras à longue chaîne (AGCL)(Mariadason *et al.* 2005)

MG : monoglycérides ; TG : triglycérides ; PL : phospholipides ; EC : ester de cholestérol ; LBPs : lipid binding protein ; CD36 : fattyacid transporter ; FATP4 : fattyacidtransporterprotein 4 ; I-FABP : intestinal fattyacid-binding protein ; L-FABP : liverfattyacid-binding protein ; ACS : acylCoA synthétase ; AGLC-SCoA : acyl-CoA ; ACBP : acyl-CoA-binding protein ; MTP : microsomal triglyceridetransferprotein ; ApoB48 : apolipoprotéine B48 ; ApoA-IV : apolipoprotéine A-IV ; RE : réticulum endoplasmique ; ApoAI : apolipoprotéine AI ; ApoCII : apolipoprotéine CII ; ApoCIII : apolipoprotéine CIII ; SAR1b : smadanchor for receptor activation 2 ; Vamp7 : vesicle-associated membrane protein 7.

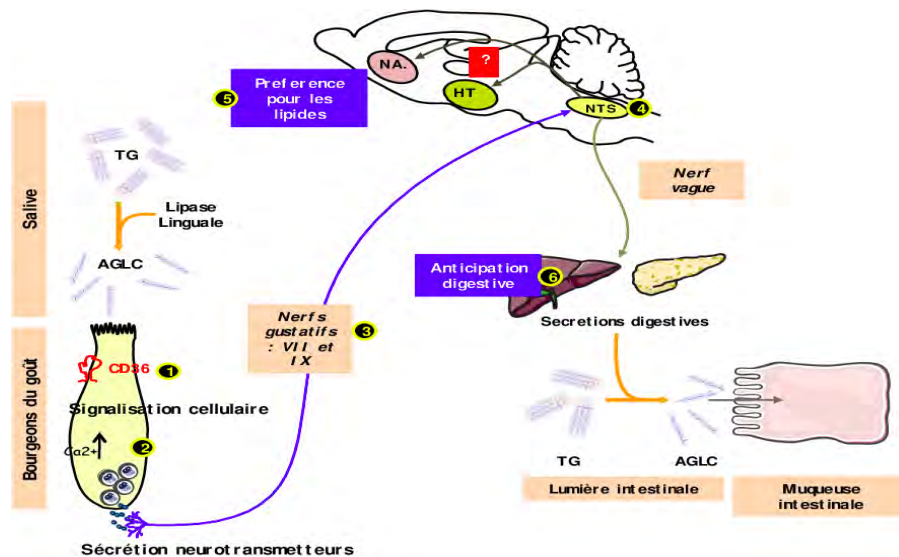
En ce qui concerne l'absorption intestinale des lipides, les mécanismes cellulaires et moléculaires ne sont pas encore totalement élucidés. Cela s'explique par le fait que l'intestin a longtemps été considéré comme une simple barrière passive capable d'absorber de grandes quantités de lipides alimentaires. Étant donné les rôles fondamentaux des acides gras à longue chaîne (AGLC) dans la cellule (fonctionnel, énergétique, métabolique et de signalisation cellulaire), on conçoit que leur biodisponibilité ne soit pas limitante. Le fait que l'excrétion fécale des lipides soit de l'ordre de 5 % chez l'Homme malgré de grandes variations de l'apport lipidique et que la stéatorrhée soit diagnostiquée dès que la perte excède 7 %, souligne l'efficacité de l'absorption intestinale des AGLC (Ross *et al.*, 1993). Ces dernières années, plusieurs lipid-binding proteins (LBP) ont été identifiées au niveau de l'entérocyte, il s'agit de la Plasma membrane fattyacid-binding Protein ou FABPpm, du fattyacidtransporter ou FAT/CD36 et enfin du fattyacidtransferprotein 4 ou FATP4 montrant la complexité du mécanisme d'absorption et suggérant que des modifications de leur niveau d'expression intestinale retentissent sur l'efficacité de l'absorption et la lipémie postprandiale, et plus généralement sur l'homéostasie des lipides de l'organisme (Swerlick *et al.*, 1992). C'est pourquoi des revues aborderaient le rôle des LBP, en particulier membranaires, dans le mécanisme d'absorption des lipides et sa régulation. Nous nous intéressons plus de la protéine CD36.

#### **I.1.6.2. Rôle du CD36 dans la préférence spontanée aux lipides alimentaires et signalisation cellulaire**

CD36 est impliqué dans la préférence spontanée aux lipides puisque les souris CD36 (-/-) perdent la préférence spontanée pour des solutions contenant des lipides (Armesilla *et al.*, 1996). Il a été montré que cette détection des lipides médiée par CD36 passe par la voie gustative. En effet, CD36 est exprimé en particulier au niveau des papilles caliciformes de souris, de rat et d'Homme.

La liaison des AGLC avec CD36 induit une cascade de signalisation cellulaire qui passe par l'activation des Src kinases, Fyn et Yes qui activent la phospholipase C ((Drover *et al.*, 2008); (Nauli *et al.*, 2006)). L'inositol triphosphate ainsi libéré provoque la libération de calcium à partir du RE, générant l'ouverture de canaux store-operated calcium channels (SOC) conduisant à un influx de calcium. L'ensemble de ce mécanisme conduit à l'exocytose de neurotransmetteurs activant les nerfs gustatifs (Simon *et al.*, 2011) (Figure 10). Ils transmettent ce signal aux fibres gustatives afférentes qui à leur tour relaient l'information au niveau du tractus solitaire (NTS) (Laugerette *et al.*, 2005) (Figure I.10). Cette activation cérébrale participe en retour à la régulation de la phase céphalique de la digestion caractérisée par une augmentation de flux pancréato-biliaire et du contenu en protéines du suc pancréatique (Laugerette *et al.*, 2005).

Cette nouvelle découverte suggère qu'il existe une 6ème saveur : le goût du gras (Kuda *et al.*, 2013; El-Yassimi *et al.*, 2008).



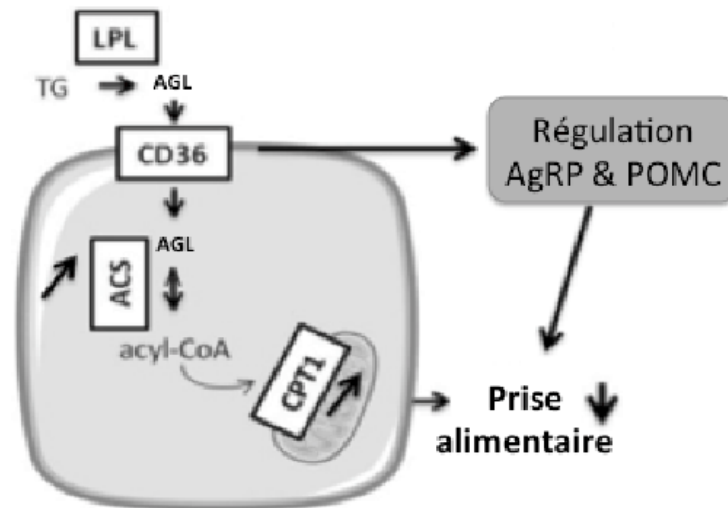
**Figure 10** : Rôle du CD36 dans la détection orale des lipides alimentaires (Chevrot *et al.*, 2013)  
 (CD36 intestinal : un récepteur aux acides gras à longue chaîne qui contrôle l'hypertriglycéridémie post prandiale, l'endotoxémie et l'intégrité de l'épithélium intestinal)

Cette fonction a une pertinence physiologique chez l'Homme puisque chez les individus et les souris obèses, il existe une relation entre l'IMC et le seuil de détection de lipides (Chevrot *et al.*, 2014). Par ailleurs, le génotype de CD36 affecte la détection orale des lipides. En effet, les sujets mutés pour CD36 (homozygotes pour l'allèle rs1761667-G) ont une diminution du seuil de détection des lipides associée à une diminution de l'expression de CD36. CD36 joue donc un rôle dans la détection des lipides au niveau lingual qui conduit à la préparation du système digestif à l'arrivée de lipides (Pepino *et al.*, 2012).

### I.1.6.3. CD36, senseur de lipides au niveau hypothalamique : impact sur la régulation de la prise alimentaire.

Au niveau hypothalamique, certains neurones sont spécialisés dans la détection des nutriments énergétiques tels que le glucose et les AG, qui permettent de détecter le statut énergétique au niveau central. Il a été montré que l'injection d'intralipide hépariné au niveau cérébral diminue la prise alimentaire (Moullé *et al.*, 2014). En utilisant *in vivo*, un shRNA chez le rat, il a été démontré que sur l'ensemble des neurones répondant à l'OA, 50 % répondent via CD36, qui déclenche un flux calcique. Cet effet est confirmé par Moullé et al, qui ont démontré que l'effet inhibiteur des lipides sur la prise alimentaire dépend de CD36 et de l'activité d'une enzyme ACS, qui active les AG en acylCoA avant leur acheminement vers la mitochondrie où ils sont oxydés (Figure 11) (Moullé *et al.*, 2013).

Enfin, la réduction de 35 % à 70 % de CD36 conduit à une diminution de l'expression de l'ACS, de CPT1 et diminue l'expression des peptides agouti related peptide (AgRP) et ProOpio Melano Cortin (POMC) au niveau du Noyau Arqué, régulant la prise alimentaire (Le Foll *et al.*, 2013) (Figure 11). Ces résultats suggèrent fortement que même au niveau central, CD36 joue un rôle de récepteur et participe à la régulation de la prise alimentaire.



**Figure 11** : Rôle de CD36 dans la détection des lipides et la régulation de la prise alimentaire au niveau central (Le Foll *et al.*, 2013)

(CD36 intestinal : un récepteur aux acides gras à longue chaîne qui contrôle l'hypertriglycéridémie post prandiale, l'endotoxémie et l'intégrité de l'épithélium intestinal).

Ainsi l'ensemble de ces données démontre le rôle du senseur des AGLC dépendant de CD36 au niveau oro-intestinal et central sur la digestion, l'absorption des lipides et la régulation de la prise alimentaire. Un dysfonctionnement de ce senseur pourrait être à l'origine d'un défaut de la biodisponibilité des AGLC et de l'homéostasie énergétique. CD36 joue également un rôle fondamental dans la défense de l'organisme en liant certains pathogènes (*Plasmodium falciparum*), le peptide beta amyloïde, les cellules apoptotiques ou encore des lipides modifiés (LDL oxydées...) (Su *et al.*, 2009)

### **I.1.7.Rôles pathologiques du CD36 dans les infections par malaria et Covid 19**

L'infection par *Plasmodium falciparum*, responsable de la forme la plus grave du paludisme, est l'adhérence des globules rouges infectés (iRBCs) à l'endothélium capillaire. La protéine humaine CD36 est un récepteur majeur de *Plasmodium falciparum*-iRBC et peut contribuer au développement du paludisme en séquestrant iRBCs et inhibant la réponse immunitaire au parasite. En revanche, cette glycoprotéine est également associée au développement de lésions athérosclérotiques via sa participation au métabolisme des LDL oxydées, des phospholipides et principalement des acides gras (Glatz *et al*, 2017; Glatz *et al*, 2018; Li *et al*, 2019). Le CD36 est un récepteur piègeur majeur responsable de la reconnaissance et de l'internalisation des oxLDL (Erol *et al*, 2020). L'absorption médiée par les récepteurs d'oxLDL par les macrophages dérivés de monocytes active la reprogrammation des réponses immunitaires innées, appelées « immunité entraînée » (Erol *et al*, 2020). Il est intéressant de noter que le réseau représentatif de modules de transcription sanguine enrichis dans les monocytes des patients COVID-19 par rapport aux témoins, ainsi que l'activation des cellules T et NK, la cytopénie et la régulation positive des gènes du cycle cellulaire et des immunoglobulines ont détecté une augmentation des signaux de l'activation des monocytes, y compris la régulation à la hausse de CD36 (Gardinassi *et al*, 2020). Cela peut être dû au fait que le COVID-19 sévère est caractérisé par une fréquence élevée de sous-ensembles de monocytes classiques exprimant CD36 (Chevrier *et al*, 2021). Nous avons émis l'hypothèse qu'une régulation à la hausse de l'expression de CD36 pourrait être un facteur prédictif de COVID-19 sévère avec des résultats létaux possibles. Ainsi, une modification de l'expression du gène CD36 peut être associée à l'évolution défavorable de la maladie chez les patients atteints d'une forme critique de COVID-19 par au moins deux mécanismes différents : l'activation de l'immunité innée contre les cellules contenant le virus et l'activation du processus d'athérogenèse via une meilleure absorption des acides gras et des lipoprotéines oxydées.

Les troubles du goût sont reconnus comme des symptômes courants chez les patients COVID 19, les relations entre le récepteur CD36 et le trouble du goût sont mal connues même en l'absence d'études pertinentes sur l'effet des altérations du goût et de l'odorat sur la consommation alimentaire pendant la COVID-19.

### **I.1.8.Acides gras et les sites de liaison du récepteur CD36**

#### **I.1.8.1.Identification d'une poche de liaison des AGLC dans CD36**

La fonction du CD36 dans l'absorption des AGLC et de la signalisation induite par ces derniers peut être irréversiblement inhibée par la N-hydroxysuccinimidyle (NHS), des esters à longue chaîne d'AG. Ces réactifs de marquage des protéines réagissent rapidement pour former des liaisons



stables principalement avec les groupes amines dans les chaînes latérales de la lysine. Les esters sulfo-NHS contiennent un groupe sulfonate chargé négativement qui limite la perméabilité de la membrane. Les esters de palmitate, myristate, et de l'oléate se sont révélés inhiber efficacement l'absorption des AG par des adipocytes isolés, propriété utilisée dans les premières études pour identifier le CD36 comme récepteur membranaire aux AG. (Cartoni *et al*, 2010)

Le mécanisme moléculaire sous-jacent de l'inhibition par SSO a été récemment exploré dans une étude où des cellules d'un ovaire du hamster chinois (CHO) exprimant de manière stable CD36 humain ont été traitées avec du SSO avant d'être soumises à la protéomique. La cible du SSO à CD36 humaine a été identifiée comme la lysine 164, un résidu situé dans une poche déjà suggérée comme un site potentiel de la fixation des AG contraignant, fondé sur des simulations avec la poche de liaison de la protéine (Baillie *et al*, 1996).

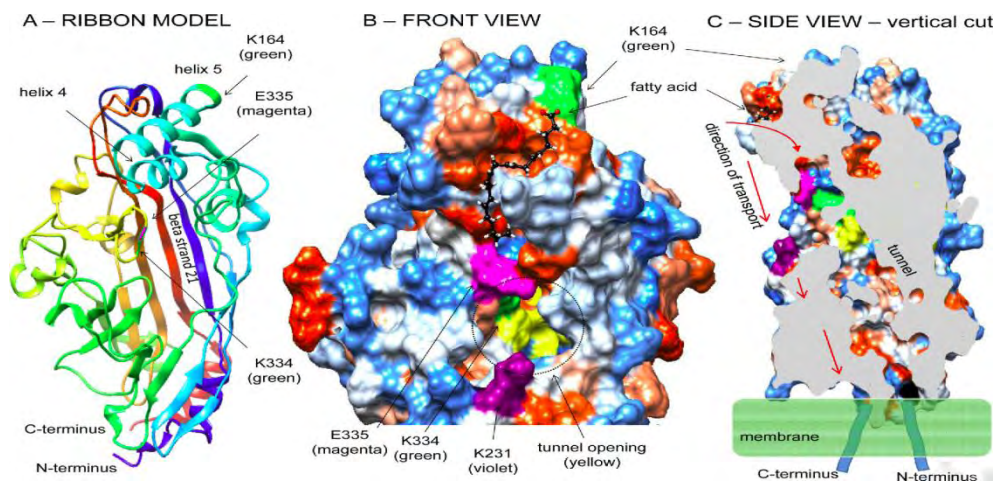
#### **I.1.8.2.L'importance de la Lysine-164 dans la fixation et l'induction de la signalisation des acides gras**

L'importance de la Lys-164 pour l'absorption des AG par CD36 et la signalisation induite a été démontrée en utilisant la mutagenèse et l'expression dans des cellules CHO. La poche de liaison CD36 lipidique a une prépondérance de résidus hydrophobes et le modèle de travail est que la chaîne alkyle de l'acide gras servirait à diriger et positionner l'AG dans la poche (Storch *et al*, 2009). Cette interprétation est compatible avec les données des plusieurs travaux montrant qu'une chaîne plus courte NHS dérivé tel que sulfo-N-succinimidyle myristate liée à plusieurs protéines dans les membranes des adipocytes, tandis que SSO a été récupéré principalement sur CD36. Une fois dans la poche, le carboxyle de l'AG peut former des interactions électrostatiques avec Lys-164 et ceux-ci induirait un changement de conformation dans CD36 pour promouvoir l'absorption AG et / ou la signalisation induite par l'acide gras. Cette interprétation est basée sur l'observation que l'expression d'un CD36 muté où la lysine 164 est substituée par des résidus d'alanine dans la signalisation qui apparaît limitée celle induite par l'AG (Samovski et Abumrad, observations non publiées). Dans ce contexte, l'interaction de l'AG avec la lysine dans la poche de liaison de la protéine intestinale a été documentée et se sont révélés influencer le transfert des AG pour modéliser les membranes. Fait intéressant, la poche de liaison de CD36 contient une autre lysine, Lys-166, qui ne se lie pas au SSO. Cette lysine a été trouvée être acétylée (ainsi que Lys 52, 231 et 403), qui neutralise sa charge et accroît le caractère hydrophobe de la poche de liaison (Poirier *et al*, 1996). Potentiellement l'acétylation de la Lys-166 peut réglementer l'accès à la poche de liaison, ou aider le positionnement de l'AG dans le site, ce qui pourrait réguler le changement de conformation induit dans CD36 et la signalisation ultérieure. Par exemple, l'acétylation / désacétylation qui régule

la fonction d'un certain nombre de protéines impliquées dans le métabolisme du glucose et des AG (Lundby *et al.*, 2012).

### I.1.8.3.L'internalisation des AGLC par l'accès dans un tunnel au sein du CD36

Un aperçu du mécanisme potentiel du transport des AGLC par CD36 peut être dérivé de la structure cristalline du LIMP2 un membre de la famille du CD36. La structure cristalline identifie la présence d'une grande cavité qui traverse toute la longueur de la molécule qui a été proposée pour servir comme un tunnel pour le transfert des lipides à la membrane (Neculai *et al.*, 2013). Des chercheurs ont modélisé la structure de CD36 et ont réalisé l'amarrage des AGLC sur CD36 (Neculai *et al.*, 2013). Le modèle obtenu prévoit que la poche de liaison des AG forme une rainure qui coulisse vers le tunnel (Figure 12). L'interaction du groupe carboxyle de l'AG avec la Lys-164 au sommet de la poche positionnerait l'acide gras afin qu'il puisse coulisser dans le tunnel avec sa chaîne acyle en tête (Silverstein *et al.*, 2014). L'interaction est également susceptible d'induire un changement de conformation, une interprétation conforme à la conclusion que la substitution de Lys-164 par des résidus d'alanine dans les changements qui limitent la liaison de l'AG et donc limiter la signalisation (Marta *et al.*, 2014).



**Figure 12 :** Le site de fixation des acides gras sur CD36 et leur accès au tunnel. (Marta *et al.*, 2014).

Le changement de conformation pourrait déplacer le glutamate 335 situé au sommet du tunnel pour faciliter l'accès des AG à ce dernier. L'acide gras empruntera le tunnel pour atteindre la membrane plasmique où la sortie du tunnel est susceptible d'être enterrée. Le tunnel a une ouverture sur la surface de CD36, mais la modélisation proposée dans ces anciens travaux et la liaison de données d'authentification unique suggèrent que AG serait rarement accéder au tunnel ((Marta *et al.*, 2014) ; (Baillie *et al.*, 1996)). La Lysine 231 présente à l'ouverture de la surface du tunnel est acétylée qui pourrait limiter l'accès des AG au tunnel (Figure 12). Que deux sites d'entrée du tunnel (via la



poche de liaison des AG ou l'ouverture de la surface du tunnel) sont possibles pour l'AG, si cela implique le type de l'AG donc différentes fonctions de signalisation auraient besoin d'être testées expérimentalement ((Baillie *et al.*, 1996) ; (Marta *et al.*, 2014)).

#### **I.1.8.4. Lesite de liaison des AG chevauche avec ceux des oxLDL et de la phosphatidylch-oline oxydée**

Une caractéristique importante du site de liaison des AG identifiée dans CD36 est qu'il chevauche avec ceux de la LDL oxydée (oxLDL) et de la phosphatidylcholine oxydée étant donné que ces deux ligands ont été proposés pour interagir avec les lysines 164 et 166 (Karet *al.*, 2008). En accord avec la possibilité d'un site commun de liaison, l'incubation avec SSO (25 uM, 30 min) inhibe efficacement l'absorption des oxLDL par les macrophages. Les implications physiologiques de la communauté des résidus du site de liaison et des interactions médiatrices d'AG et oxLDL avec CD36 sont inconnues à l'heure actuelle. Toutefois, des résultats pourraient fournir un cadre pour la façon dont oxLDL pourrait induire un dysfonctionnement du métabolisme des AG CD36 réglementés. (Kar *et al.* , 2008).

## Objectif

L'objectif général de cette thèse est de déterminer des mécanismes d'interaction et de signalisation des acides gras via CD36 avec et sans infection COVID-19 dans le but de développer des workflow basés sur l'approche *in silico* par modélisation et dynamique moléculaire nécessaires pour la conception, le criblage et l'identification de nouvelles molécules bioactives.

Spécifiquement à:

1. Etudier le mécanisme de transport des acides gras via le récepteur CD36 en temps réel.
2. Déterminer le mécanisme de signalisation des acides gras via le récepteur CD36 et la conception de nouveaux analogues phytochimiques naturels du goût gras
3. Analyser le mécanisme structurel des troubles du goût des acides gras induits par le COVID-19 via le récepteur CD36
4. Identifier des épitopes pour un cluster de différenciation et conception de nouveaux inhibiteurs de peptides contre la RBD du pic de SARS-CoV-2 humain

## **I.2. Matériel et méthodes : Etudes *in silico***

## **I.2. Matériel et méthodes : Etudes *insilico***

### **I.2.1.L'approche *in silico* dans le domaine de la biologie (Bio-Informatique)**

Au cours de ces trente dernières années, la récolte de données en biologie a connu un boom quantitatif grâce notamment au développement de nouveaux moyens techniques servant à comprendre l'ADN et d'autres composants d'organismes vivants. Pour analyser ces données, plus nombreuses et plus complexes aussi, les scientifiques se sont tournés vers les nouvelles technologies de l'information. L'immense capacité de stockage et d'analyse des données qu'offre l'informatique leur a permis de gagner en puissance pour leurs recherches. Et la rencontre entre la biologie et l'informatique, c'est ce qu'on appelle la bioinformatique. Celle-ci couvre des disciplines des sciences de la vie telles que la génomique, la protéomique et la biologie des systèmes (institut suisse de la bioinformatique).

Trois activités principales :

- Acquisition et organisation des données biologiques
- Conception de logiciels pour l'analyse, la modélisation des données
- Analyse des résultats produits par les logiciels

#### **I.2.1.1. Modélisation moléculaire (Modélisation par homologie)**

La détermination des structures tridimensionnelles des protéines par des méthodes expérimentales continue d'alimenter et de maintenir la base de données sur les protéines. En février 2012, plus de 79 356 structures protéiques expérimentales sont disponibles dans la base de données PDB (Protein Data Bank), <http://www.rcsb.org/pdb> (Westbrooke *et al.*, 2003). Les techniques expérimentales telle que la cristallographie aux rayons X et la spectroscopie RMN sont des techniques longues et plus chère. De plus, elles échouent pour toutes les protéines, en particulier les protéines membranaires (Teichmann *et al.*, 1999 ; Capener *et al.*, 2000; Floudas *et al.*, 2006). Pour surmonter ce problème, diverses approches ont été utilisées pour prédire la structure 3D des protéines à partir de leurs séquences d'acides aminés primaires. La modélisation « comparative » ou « d'homologie » est l'approche la plus précise qui exploite le principe selon lequel les protéines partagent des structures similaires (Cavasotto *et al.*, 2009 ; Vitkup *et al.* 2001). La modélisation moléculaire est un terme qui comporte diverses techniques de chimie computationnelle et de graphisme moléculaire. Ces techniques permettent d'analyser, d'afficher, de dessiner, de calculer, de simuler et de stocker les propriétés des composés pour imiter le comportement ou modéliser des molécules.

La modélisation par homologie vise à construire un modèle structurel 3D d'une protéine en utilisant les structures déterminées expérimentalement des membres de la famille apparentés comme modèles. Autrement dit, si deux protéines ont une homologie de séquence d'acides aminés élevée,

elles ont également une bonne homologie de structure. A partir des coordonnées atomiques des protéines résolues par cristallographie dite « modèle », la structure tridimensionnelle de la protéine « cible » étudiée est reconstruite. Pour une protéine cible donnée, recherchez des bibliothèques de structures protéiques expérimentales pour identifier des modèles appropriés. Un modèle tridimensionnel de la protéine cible est généré sur la base de l'alignement de séquence entre la protéine cible et la structure du modèle. La modélisation par homologie est actuellement la méthode de calcul la plus précise pour générer des modèles structuraux fiables et est largement utilisée dans de nombreuses applications biologiques. Les projets de modélisation nécessitent généralement moins de deux heures d'effort de calcul (Lorenza et al, 2009).

### **I.2.1.2. Conception et découverte des molécules assistées par ordinateur (Drug design)**

Conception de molécules bioactives ou de médicaments computationnels La « conception de médicaments » peut être définie par un ensemble de processus ingénieux permettant de trouver de nouveaux médicaments en fonction de la connaissance de leurs cibles biologiques. La conception de molécules bioactives ou de médicaments peut être étudiée par la modélisation moléculaire et la conception informatique de nouvelles molécules. “*in silico*” ((Gualtieri *et al*, 1995); (HyperChem *et al*, 2008), (MarvinSketch *et al*, 2017); (Ghose *et al*, 2011); (Wager *et al*, 2010).

Des outils informatiques sont désormais utilisés à différentes étapes de la conception de molécules bioactives (Wager *et al*, 2016; Wager *et al*, 2010) :

Approche basée sur le complexe récepteur-ligand

- Approche basée sur les propriétés du complexe récepteur-ligand “Criblage virtuel”
- Phase de sélection des composés actifs qui présente des affinités supérieures à celle de notre ligand « lead ».
- Optimisation de lead possédant des propriétés pharmacologiques recherchées essentielle pour un candidat bioactif (Lipinski *et al*, 1997).

#### **I.2.1.2.1. Criblage virtuel «*In Silico*»**

Le terme criblage virtuel est un processus qui mène à la découverte de nouvelles molécules d'intérêt thérapeutique à l'aide d'un ensemble de techniques computationnelle (Schultes *et al*, 2010). Le rôle des méthodes de criblage est de réduire le nombre considérable de molécules à tester expérimentalement et de diminuer le coût et le temps considérable nécessaire au processus de R&D et pour guider la sélection des composés les plus prometteurs, que ce soit lors des phases d'identification des hits ou d'optimisation des leads (Contreras *et al.*, 1999).

Les méthodes de criblage consistant à prédire l'affinité d'un très grand nombre de molécules (collectées en chimiothèques) pour le site actif ciblé afin d'aboutir à une liste très réduite de molécules présentant l'activité souhaitée. La diversité moléculaire des chimiothèques criblées expérimentalement ou virtuellement est un critère important pour mieux couvrir l'espace chimique et donc pour augmenter les chances d'identifier de nouvelles touches « hits ». Les molécules doivent également avoir les propriétés physico-chimiques comme la biodisponibilité et les propriétés ADME (Absorption, Distribution, Métabolisme et Excrétion) caractérisant un médicament. Pour cela, des règles élémentaires permettent de filtrer rapidement des chimiothèques et éliminer toutes les molécules qui ne répondant pas à ces règles ; les plus connues sont celles de Lipinski (Lipinski *et al*, 1997), de Veber (Veber *et al*, 2002) et de triangle d'or (Johnson *et al*, 2009).

Le criblage virtuel nécessite des informations sur les propriétés physico-chimiques et/ou géométriques des composés actifs et sur la structure 3D de la cible. Selon la disponibilité ou non de ces informations, différentes méthodes de criblage virtuel peuvent être utilisées. Les méthodes basées sur les ligands (Ligand-Based Drug Design LBDD) utilisent les ligands connus comme guides pour définir les propriétés responsables de leur affinité (Aparoy *et al*, 2012). Les méthodes basées sur la structure (Structure-Based Drug Design SBDD) sont utilisées lorsque les données sur la structure 3D de la cible biologique sont disponibles. Elles permettent d'évaluer le potentiel d'interaction entre les composés criblés et le site d'interaction de la cible (Wang *et al*, 2018). Au sein de ces deux grandes familles, différentes méthodes ont été développées.

#### **I.2.1.2.1.1.Approche basée sur le ligand «Ligand-based»**

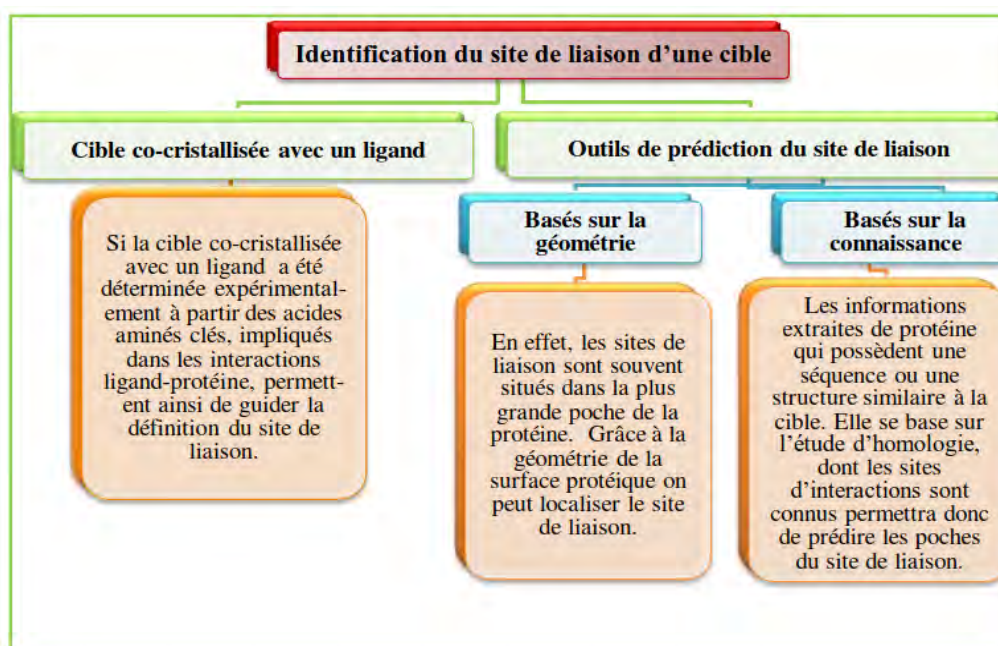
Cette démarche repose sur la connaissance préalable de ligands ayant une activité sur la cible thérapeutique connue. Il sera ainsi possible d'utiliser ces ligands comme une première base de « hits » afin d'identifier d'autres composés similaires, présentant des caractéristiques d'activité commune (Waring *et al*, 2009 ; Martin *et al*, 2005). L'hypothèse que les molécules similaires vont avoir tendance à présenter des profils d'activités très proches à similaires est le principe de base de toutes les méthodes « ligand-based ». Différents types de descripteurs moléculaires pourront être calculés pour quantifier la similitude entre les composés. Ces méthodes présentent l'avantage d'être un protocole de découverte de médicament viable, même en absence de structure expérimentale tridimensionnelle (3D) du récepteur ciblé (Pajouhesh *et al*, 2017, Johnson *et al*, 2009). Suivant le nombre de ligands connus de la cible thérapeutique, plusieurs méthodes peuvent être employées : la recherche de similarité, le criblage de pharmacophores et les approches QSAR. Ces deux dernières méthodes sont les outils les plus importants et largement utilisés dans la conception de médicaments

à base de ligands (Pajouhesh *et al*, 2005). Les modèles réussis à base de ligand permettent l'accès à de nouveaux ligands « scaffoldhopping » (Kerns *et al*, 2008; Hansch *et al*, 1979; Bazooyaret *al*, 2013) le « scaffoldhopping » fait la référence à la recherche assistée par ordinateur des composés ayant une activité biologique similaire, mais contenant des structures de noyau différent (Fichert *et al*, 1997). L'objectif fondamental de l'approche « scaffoldhopping » est de découvrir des composés structurellement nouveaux à partir des composés actifs connus en modifiant la structure centrale de la molécule (Fichert *et al*, 1997, Palm *et al*, 1997). Les nouveaux composés offrent l'occasion de modifier les affinités et les sélectivités des récepteurs (profils de toxicité), ainsi que d'améliorer les profils pharmacocinétiques (propriétés ADME) et d'obtenir des analogues brevetables (Palmet *al*, 1997).

#### **I.2.1.2.1.2.Approche basée sur la structure « Structure-based»**

Cette démarche repose sur la structure utilise les interactions perçues au sein des complexes ligands-récepteurs pour générer des modèles de liaison et les utiliser dans la découverte et l'optimisation de nouveaux ligands bioactifs. Le criblage virtuel « structure-based » est considéré comme un équivalent *in silico* d'un test expérimental étudiant la liaison ligand-cible biomoléculaire (Kryger *et al*, 1999). Cependant, il se repose essentiellement sur la disponibilité de la structure 3D de la cible thérapeutique qui est obtenue par des méthodes expérimentales telles que la cristallographie par rayons X et la spectroscopie par Résonance Magnétique Nucléaire (RMN) dans des bases de données (tel que PDB:Protein Data Bank) 78. Par contre, pour toutes les structures non déterminées expérimentalement, des méthodes de prédiction de la structure 3D par homologie de séquence peuvent être mises en œuvre (Zerroug *et al*, 2019, Murphy *et al*, 2010 ; Seltzer *et al*, 2005), basée sur la structure expérimentale d'une protéine proche (même classe voire même famille).

Comme leur nom l'indique, ces méthodes utilisent la structure de la cible pour découvrir de nouveaux composés actifs. Pour cela, plusieurs approches peuvent être employées : le docking moléculaire, les approches pharmacophoriques, la méthode QSAR (RD-QSAR) et le design de novo (Zerroug *et al*, 2019; Dunn *et al*, 2000; Ferreira *et al*, 2006). Chacune de ces méthodes requiert la détermination du site de liaison sur la cible thérapeutique. Pour cela, deux solutions sont possibles : obtenir une structure 3D de la protéine co-cristallisée avec un ligand, ou utiliser des outils de prédiction (Patrick *et al*, 1999) (Schéma 13). Cela nous donne l'information nécessaire pour identifier les groupements d'interactions du ligand dans la poche de liaison du récepteur.



**Schéma 13:** Méthodes de la détermination du site de liaison sur la cible thérapeutique (Lagarde *et al.*, (2014).

### I.2.1.3. Amarrage moléculaire (Docking)

Dans le domaine de la modélisation moléculaire, l'amarrage (en anglais, docking) est une méthode qui calcule l'orientation préférée d'une molécule vers une seconde lorsqu'elles sont liées pour former un complexe stable<sup>2</sup>. Le docking moléculaire *in silico* vise à prédire la structure d'un complexe moléculaire à partir des molécules isolées, ce qui est considérablement plus facile à mettre en œuvre, moins cher et plus rapide que l'utilisation des méthodes expérimentales *in vitro*. Les logiciels de docking sont donc des outils très utiles en biologie, pharmacie et médecine, car la plupart des principes actifs sont de petites molécules (ligand) qui interagissent avec une cible biologique d'intérêt thérapeutique, généralement protéique (récepteur), afin d'influencer le mécanisme dans lequel cette protéine est impliquée. (Kitchen *et al*, 2004)

### I.2.1.4. La Dynamique Moléculaire (Molecular Dynamics)

La dynamique moléculaire (DM) est un type de simulation fondamental et très courant en modélisation moléculaire. Le but de la dynamique moléculaire est de créer un cours représentatif de l'évolution temporelle d'un système au niveau moléculaire ((Karplus *et al*, 2002); (Hansson *et al*, 2002)). La dynamique moléculaire (DM) est un type de simulation fondamental et très courant en modélisation moléculaire. Le but de la dynamique moléculaire est de créer un cours représentatif de l'évolution temporelle d'un système au niveau moléculaire ou d'expliquer des mécanismes à l'échelle atomique comme la liaison d'un ligand à sa cible (Durrant *et al*, 2011), le repliement d'un peptide (Brooks *et al*, 1993) ou l'activation d'un récepteur (Grossfield *et al*, 2011). De nombreuses suites



logicielles comme AMBER ((Pearlman *et al*, 1995); (Case *et al*, 2012)) et CHARMM (Brooks *et al*, 1983) intègrent la possibilité de réaliser des simulations de DM. Ces dernières sont généralement constituées de trois phases successives :

- 1) le chauffage du système jusqu'à la température désirée de manière à donner une impulsion à ses atomes (notions de vitesse et d'énergie cinétique  $E_{cin}$ ).
- 2) l'équilibration du système à cette température pendant un certain temps
- 3) la production qui correspond à la trajectoire d'intérêt pour une durée donnée (nanosecondes jusqu'à millisecondes en fonction des phénomènes étudiés).

#### **I.2.1.4.1.Ecart quadratique moyen (RMSD)**

L'écart quadratique moyen (ou RMSD) est utilisé pour quantifier le changement moyen subi par une protéine lors d'une simulation (Michaud-Agrawal *et al.*, 2011). Calculé à l'aide de la formule suivante :

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^{i=N} \delta_i^2}$$

Avec  $\delta_i$  La distance entre N paires d'atomes équivalents que vous voulez comparer. RMSD est utilisé pour comparer les structures des protéines.. Plus il est proche de 0, plus la structure d'intérêt est semblable à la structure de référence. La sélection de structure peut être la fin de la structure de l'étape précédente ou la minimisation pour suivre l'évolution de la protéine à chaque pas de temps. Les mesures sont effectuées uniquement sur la chaîne peptidique principale de la protéine ou simplement sur le carbone alpha ( $C\alpha$ ) Obtenez un aperçu du mouvement des protéines. La valeur RMSD augmente progressivement pendant la production et atteint un plateau. Après cela, le système est considéré comme stable. A partir de ce moment, il devient possible pour RMSD d'effectuer d'autres analyses sur la partie de la trajectoire stable.

#### **I.2.1.4.2.Rayon de giration**

Un autre critère permettant de vérifier la stabilité globale en orbite est le rayon de braquage. ( $R_g$ ) (Michaud-Agrawal, N *et al* 2011). Il donne une idée de l'extension spatiale d'une structure à partir de son centre de masse. Une élévation substantielle ( $> 5 \text{ \AA}$  par exemple) la simulation pourrait signifier le dépliement des protéines, le désassemblage de la structure secondaire. Peut être utilisé pour suivre la dénaturation pendant les processus de chauffage à haute température. La formule du  $R_g$  est la suivante :

$$R_g = \sqrt{\frac{1}{N} \sum_{i=1}^N \|\vec{r}_i - \vec{r}_{cm}\|^2}$$

Avec N le nombre total d'atomes,  $r_i$  la position de l'atome  $i$  et  $r_{cm}$  la position du centre de masse de la protéine.

#### I.2.1.4.3. Fluctuations atomiques (RMSF)

RMSD concerne la variation dans le temps des écarts de distance atomique par rapport à une structure de référence, tandis que RMSF (ou fluctuation atomique) est l'écart moyen de chaque résidu (C  $\alpha$  pour chaque résidu) par rapport au même atome à chaque résidu représenté. Structure de référence ((Welford, et al., 1962); (Michaud et al. 2011)). Le RMSF est calculé à l'aide de la formule suivante :

$$RMSF_i = \sqrt{\frac{1}{T} \sum_{t=0}^{t=T} (x_i^t - \tilde{x}_i)^2}$$

En utilisant  $RMSF_i$ , les fluctuations atomiques de l'atome  $i$  sont calculées pour la durée de la trajectoire de T pas, en utilisant la position de l'atome au temps  $t$  et la position moyenne sur le même intervalle  $x$ . La relation entre le facteur B (ou facteur d'agitation thermique) et RMSF permet de comparer les données obtenues en MD avec les données expérimentales.

$$B_{factor} = \frac{8\pi^2}{3} RMSF^2$$

Une valeur de zéro pour la fluctuation atomique signifie que la position moyenne des résidus ne change pas de celle de la structure de référence. Cette quantité permet de localiser à la fois la région charnière de la protéine et les résidus responsables des mouvements de grande amplitude.

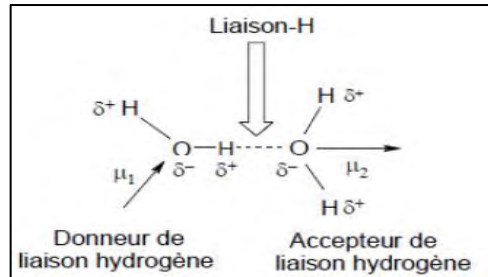
Pour affiner l'étude du mouvement des protéines, il existe des outils plus complexes qui peuvent distinguer le mouvement collectif ou local au sein de la structure secondaire entre différents domaines d'une même protéine. , entropie conformationnelle (Andricioaei *et al*, 2001).

#### I.2.1.4.4. Les liaisons hydrogènes

Les liaisons hydrogène sont d'une importance fondamentale en biologie en raison de leur rôle dans les phénomènes de solvatation et la catalyse enzymatique, pour ne citer que deux exemples importants. Ils jouent également un rôle structural par la liaison qu'ils assurent au niveau de la structure secondaire ((McRee et al, 1999; (Michaud-Agrawal et al 2011))). Elle se situe entre les liaisons fortes (covalentes, ioniques) et faibles (van der Waals), et l'énergie nécessaire pour rompre la liaison est de 5 à 30 kJ/mol. Ils sont "formés" entre un atome donneur (azote, oxygène, fluor,

soufre...) et un accepteur d'électrons. Ils sont de nature anisotrope. Autrement dit, il repose sur des contraintes géométriques entre le donneur (D) et l'accepteur (A). Cela peut être défini (C. Audouze et al., 2003) (Figure 14)::

- l'angle entre les vecteurs DH et DA doit être inférieur à  $30^\circ$
- la distance entre les atomes D et A doit être inférieure à 3,5



**Figure 14** : Liaison hydrogène : cas de deux molécules identiques de H<sub>2</sub>O

(KHERFIA et al. 2012)

**CHAPITRE II: The computational analyses,  
molecular dynamics of fatty-acid transport  
mechanism to the CD36 receptor. Nature  
Journal of Scientific Report.**

## II.1.Résumé de l'article

### Développement d'une méthode d'analyse du mécanisme de transport des acides gras via le récepteur CD36 en temps réel par la dynamique moléculaire.

Un apport excessif de lipides en particulier d'Acides Gras Saturés (AGS) associé à un déséquilibre qualitatif (excès de graisses saturées et de cholestérol, rapport n-6/n-3 trop élevé) contribue à la prévalence de l'obésité. Une corrélation a été établie entre la quantité de lipides consommés et la masse corporelle chez l'Homme (Bray, Paeratakul, et Popkin 2004) et chez la souris (de Wit *et al.*, 2011). L'obésité est associée à des perturbations métaboliques telles que l'insulino-résistance, l'intolérance au glucose, des dyslipidémies, l'hypertension, qui constituent des facteurs de risque des maladies cardiovasculaires (MCV).

La contribution active de l'intestin dans l'utilisation post prandiale des lipides a été renforcée par les dernières données de la littérature illustrent que le métabolisme de synthèse des lipoprotéines intestinales est très régulable et en particulier par les lipides alimentaires eux-mêmes. Ces données montrent, que contrairement à ce qui était admis jusque-là, l'intestin n'est pas une simple barrière passive, capable d'absorber de grandes quantités de lipides, mais qu'il est capable d'adapter ses capacités d'absorption et la qualité des lipoprotéines sécrétées à la teneur en lipides du régime. Ces modifications sont susceptibles de retentir sur la triglycéridémie posprandiale, via l'activité de la LPL, qui est conditionnée en particulier par la taille et le nombre des particules sécrétées ((Xiang *et al.*, 1999) ; (Martins *et al.*, 1996)), mais également sur le devenir des lipides alimentaires dans l'organisme. L'ensemble de ces découvertes montrent que l'intestin peut avoir un rôle clé dans la mise en place de l'obésité induite par une forte consommation de lipides par sa capacité à réguler la prise alimentaire. De ce fait, la compréhension du mécanisme moléculaire à l'origine de la régulation des mécanismes d'absorption des lipides constitue une approche pertinente pour développer de nouvelles pistes thérapeutiques permettant de lutter contre l'obésité.

Selon notre étude bibliographique, un nombre très limité de publication a décrit le choix de deux protéines (CD36 et SR-B1) en tant que candidates pour la détection des lipides au niveau intestinal ((Béaslas *et al.*, 2009) ; (Tran *et al.*, 2011)).

La CD36 est glycoprotéine membranaire située au pôle apical des entérocytes et qui lie avec une très haute affinité les AG à longues Chaînes (AGLC) et limite dans la formation des CM. En effet, la multitude de ses ligands et la découverte plus récente de ses capacités à induire différentes voies de signalisation cellulaire laissent supposer de faire émerger l'existence de nouvelles voies de CD36 au transport lipidique.

L'analyse de la séquence des acides aminés prédit une structure de récepteur. Les acides aminés de 127 à 279 forment une poche hydrophobe, permettant la liaison des AGLC saturés et insaturés avec une grande affinité, de l'ordre du nanomolaire (Baillie, Coburn, et Abumrad 1996). CD36 possède 2 domaines trans-membranaires, 2 extrémités courtes en N et C terminal intracellulaires et une large poche extracellulaire, qui contient au moins 2 sites de liaisons K164 et E335, le premier site de liaison, localisé entre les acides aminés 93 et 120 est le site de liaison pour la thrombospondine 1 alors que le 2ème site constitue une large poche hydrophobe pour la liaison avec des erythrocytes infectés par *Plasmodium falciparum* (Baruch *et al.*, 1999) sur les résidus 146-164, des LDL oxydées (Navazo *et al.*, 1996), des produits finaux de glycosylation (AGE) (Ohgami *et al.*, 2001) et des cellules apoptotiques (Navazo *et al.*, 1996) sur les résidus 155-183. Par ailleurs, il a été identifié récemment que la lysine 164 permet la liaison des AGLC avec CD36 (Kuda *et al.*, 2013). CD36 ne lie pas les AG à chaîne courte et moyenne (Su et Abumrad 2009). La poche hydrophobe et la lysine 164 est également impliquée dans la liaison de CD36 avec le SSO qui est considéré comme un inhibiteur de la liaison des AGLC avec CD36 (Kuda *et al.*, 2013).

Dans ce contexte, notre travail s'est focalisé sur l'étude du mécanisme d'absorption du CD36 par une approche *in silico*. En effet, nous avons utilisé des séries de simulations de dynamique moléculaire de la protéine de type sauvage et de type mutant K164A CD36, en interaction avec un et plusieurs acides palmitiques (PLM), afin de découvrir le mécanisme du fonctionnement du complexe CD36/Acides gras et de déterminer le rôle structural de cette mutation.

Les résultats de cette étude *in silico* ont permis de montrer que Lys164 du type sauvage, principalement exposé aux protéines de surface, a joué un rôle important dans l'absorption des acides gras et que la mutagenèse de Ala164 de type mutant K164A CD36 influence d'une manière importante sur le pliage, l'utilité, la solubilité et la forme de stabilité du CD36. En outre, ces simulations ont montré que le résidu Lys164 du type sauvage a fourni la base structurelle de la formation d'une seule entrée au canal principal du CD36 pour la dissociation d'un seul acide palmitique et qu'il pourrait assurer de préférence le transport de deux acides palmitiques au maximum.

Le contenu de ce chapitre est présenté sous forme d'un article publié dans le journal "*Nature Journal of Scientific Report*" (Akachar *et al.*, 2022). Dans cette publication, on a réalisé tout le travail d'analyse bioinformatique et de la rédaction du manuscrit.

within a nonpolar hole shared by several CD36 molecules, e.g. fatty acid and oxLDL<sup>20,21</sup>. CD36 also mediated FA uptake and signaling in CHO cells expressing mutant CD36 (K164), this can be related to that short-chain-alkyl of FA mediated the FA in relative orientation and position<sup>22,23</sup>. The second that it is in the surface of the pocket, the carboxyl group of the FA can compose electrostatic interactions with Lys-164, and this pathway followed by CD36 conformational deviation to activate FA uptake and/or signaling. This analysis is based on the observation of the mutated CD36 expression, where the substitution of even one residue Lys to Ala, triggers a signaling cascade where its stimulation by FA binding. Thus, further, Lys-164 residues mediating as well as the interaction of FA mechanism with CD36 are still unknown to the present time. As fatty acid uptake is linked to fatty acid metabolism and atherosclerosis, it would be interesting to see how we can prevent fatty acid uptake as a proof of principle of targeting this region of CD36. Although that the prevalence of studies that CD36 is an essential receptor is involved in the fat metabolism, detection, signaling pathways in specific cells. It also, could imagine that a potential drug target for the treatment of diabetes, atherosclerosis and obesity and their features of cardiovascular diseases. Therefore, it is of a great significance to identify the transport mechanism of CD36.

In this paper, we attempted to study the complex structures of wild-type and mutated CD36/Fatty acid (FA) by molecular dynamic approaches, with the aim of providing detailed mechanism for the effect of Mutagenesis Lys164 in fatty acid transport and rate. Thus, the determination of mutated structure is important as no structural information is available in the protein structure databases. Firstly, we used a computational approach to construct a three-dimensional structural model then, we utilized molecular dynamic to identify fatty acid binding site, transport trajectory and rate. Based on these computational studies, the structural analysis of Lys164 mutagenesis with FA interaction in the binding protein pocket has been documented and exposed to influence FA transfer. Yet, the findings could provide a framework for how lys164 might induce dysfunction of CD36-regulated FA which can be considered one of the parameters in protein stability. This could be a target for discovering a new drug by finding new mechanisms of FA/CD36 regulation associated with obesity and/or type2 diabetes.

### Results and discussion

MD simulations can be used as a reply to many types of issues. Here we used MD simulations to determine the mechanism of transport Fatty acid (pathway and rate) CD36 receptor. Figure 1b illustrates an overview result of the applications of Molecular Dynamics Simulations, the timescale and the molecular motions accessible with the current simulation methods.

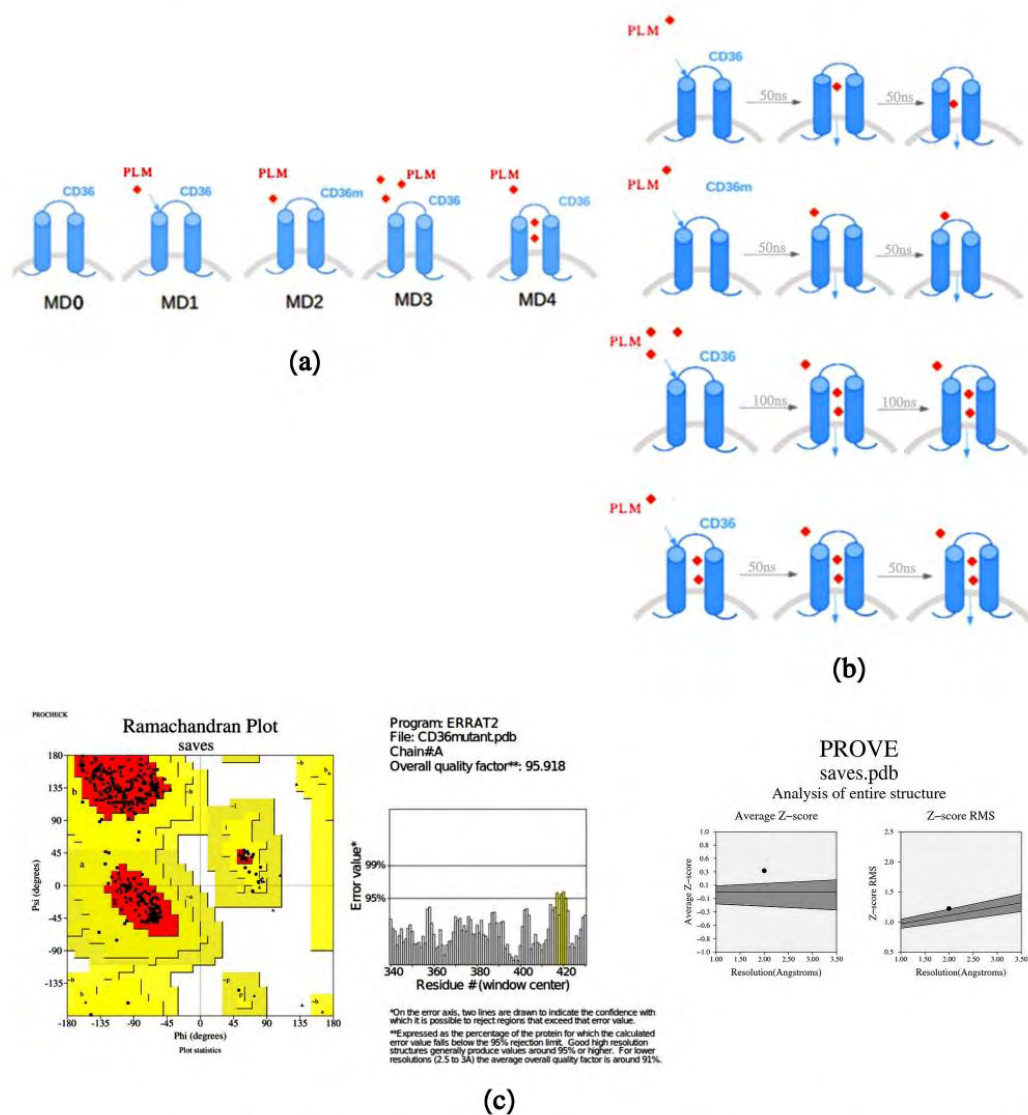
**Structural studies of mutational effects on fatty acid transport CD36 (K164A).** *The three-dimensional (3D) structure of CD36 mutant (K164A).* The tridimensional structures of virtual mutant K164A ecto-domain of CD36 predicted by using Swiss Model server to construct the homology model based on the crystal structure CD36-CIDRa (PDB ID: 5LGD). The quality of the homology model was evaluated by Ramachandran plot, the overall quality factor ERRAT2 plot and PROVE produced by the SAVS server v6.0 (<https://saves.mbi.ucla.edu/>). The stereo chemical properties of the modeled demonstrates that a majority of its residues are in favored region (92.2%) and allowed region (7.8%) of the Ramachandran plot, that indicates reasonable quality of this homology model. Moreover, the ERRAT2 score was 95.918 and the total quality index (Z-score) was 0.352. We observed that the currently mutant modeled structure have structural qualities that are acceptable<sup>24</sup> (Fig. 1c).

*Global molecular dynamics of the simulations analysis.* Because protein flexibility represents a significant function in CD36 transport and signaling, we conducted MD simulations of 100 ns on wild type in absence of PLM, in addition to the Wild type and mutant type in the presence of PLM for further analysis. Also to study the link of Lys164 residues with PLM transport and binding, we first examined the interaction of CD36 wild type crystallized ecto-domain structure with (PDB Id: 5LGD) and models that were generated for the mutant CD36 in the presence of PLM as well as alone CD36 wild type using MD simulation. After 100 ps the structure of CD36w with PLM, revealed that PLM is located above Lys164 and near the upper helix. The room in the midst of group of carboxylic group of PLM and the Lys164 is 1.2 Å whereas the mutant residue Ala164 in helix I is 5.5 Å. Clearly, this shows that the mutant residue Ala164 is far from the binding site of CD36 (Fig. 1b) and it won't instantly link with the PLM. Indeed, the mutation instantly impact the binding of PLM and hence, the K164A mutation has control of the overall conformation particularly, helix I, where the mutant amino acid K164A is existing. This is a pattern of the transportation channel of CD36 receptor<sup>25</sup>. CHO cells expressing mutated CD36 Lys-164 showed alterations FA uptake and FA-induced calcium liberation indicating the importance of Lys-164 for both FA effects<sup>20</sup>.

We calculated and evaluated the three simulation systems: the structure stability of the proteins in the time of the MD and RMSD/backbone, which were considered respective to the first conformation, as we had observed in the RMSD representation Fig. 2a. The plots of RMSD reflected equilibrium of the simulations that were signal of the global protein constancy. The peak RMSDs were at 0.25 nm to 0.55 nm for the three simulation systems. The plots were depicted in three sets. The first graph corresponded to CD36 wild-type only in absence of PLM, other showing the same for CD36 WT and MT in presence of PLM were illustrated separately with black, blue and red bars correspondingly. It is evident that the time taken for both the CD36-WT and CD36-MT in presence of PLM to reach convergence was more than their corresponding CD36r WT in absence PLM. The fluctuation size had a small difference in the medium RMSD value, which had a huge constancy from 38 ns. It was a dynamic constant compartment of the CD36r WT. The wild-type in presence of PLM attained early stability due to convergence from 43 ns whereas the mutated pair reached equilibrium much later.

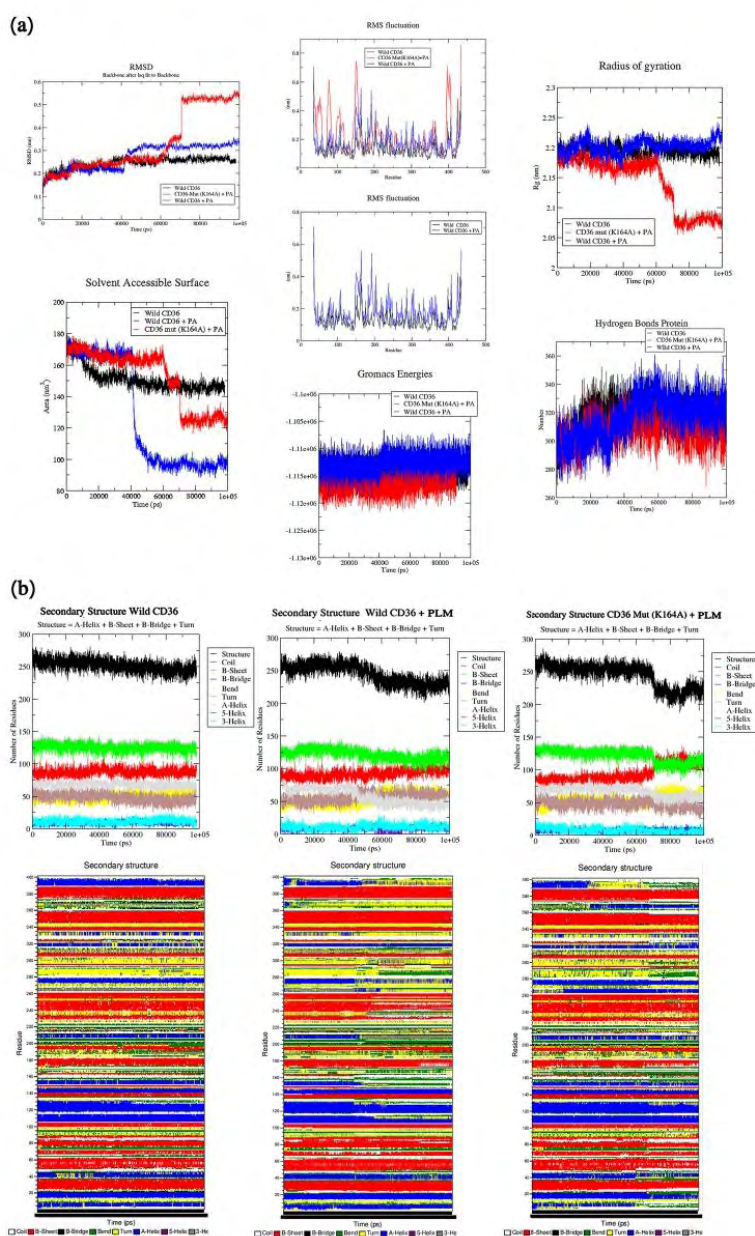
The RMSF of C-alpha atoms was used to infer with the residue specific flexibility in order to study the fluctuation of every single residue. In other words, during MD simulation three RMSF plots for CD36 wild-type in absence of PLM, the CD36 wild-type and CD36 mutant in presence of PLM were generated Fig. 2a. For





**Figure 1.** (a) Scheme of the molecular dynamics simulation procedure that the CD36 transporter (blue) and ligand palmitic acid PLM (red). *MD0* The wild type CD36 protein was simulated in the absence of palmitic acid (PLM), *MD1* the wild type CD36 protein was simulated in the presence of a single PLM, *MD2* the mutant type CD36 protein was simulated in the presence of a single PLM, *MD3* the wild type protein was simulated in the presence of a three PLM put outside, *MD4* the wild type protein was simulated in the presence of a three PLM (two PLM put inside and one PLM outside). (b) Scheme of molecular dynamic results that the CD36 transporter (blue) and ligand palmitic acid PLM (red) including the efflux of the substrate PLM along with the uptake across the CD36 receptor. (c) Evaluation of the mutant K164A CD36 (3D) structure by ERRAT2 and Prove and Ramachandran plot.





**Figure 2.** (a) Dynamics analyses of plamic acid (PLM) uptake across the wild and mutant type CD36 receptor. RMSD plot, Backbone residual fluctuations (RMSF) plot, time evolution of Radius of Gyration (Rg), Solvent Accessible Surface (SASA), Gromacs Energies and Hydrogen Bonds for wild and mutant type of CD36 in presence and absence of PLM a function of time during 100,000 ps (100 ns). Black, blue, and red, represent values obtained for Wild type CD36 in absence of PLM, wild type CD36 in presence of a single PLM, mutant K164A type CD36 in e of in presence of a single PLM respectively. (b) DSSP analysis for the secondary structure fluctuations as a function of time from 0 to 100 ns of plamic acid (PLM) uptake across the wild and mutant type CD36 receptor.

the blue curve, a trimodal distribution at the regions close to the individual portal was observed for the complex CD36-PLM wild type compared to their wild-type only black curve. The elevation of RMSF at the upper helix domain and coil was supposed to be produced by the amino acid lys 164 impacts of first assays to enter from these portals, which surround the bottom of the cavity. Interestingly, CD36 mutant red curve, exhibited elevated flexibility at N-terminal region (from 50 to 160) and C-terminal at 400 end did not display a lot of displacement at the binding region Lys164. The difference in fluctuation patterns CD36 could be attributed to their subtle importance of Lys164 residues and structural features. The flexibility of the upper part of the helix/loop is important for the organization of the unlocking and locking of the path. The representation of the RMSF showed that the conformational order occurred in the mutant. The utmost modification has taken place in the loop and upper helix region. Clearly, the access of PLM from this portal requires protein conformational change in upper helix from Lys164, which might explain why it was the most preferred route in our multiple independent runs. To conclude the RMSF plot confirmed clearly that in the helix regions of the upper loop the fluctuations of the residues were more articulated. Since the radius of gyration (Rg) is a very powerful system in analyzing the configurational compactness, the form, and the coil of the global WT, WT-PLM and MT-PLM conformation at several periods; we used it to execute Rg examination in order to look for the structural variations and dynamic stability of the wild type and mutant (K164A) structures in the presence of PLM. The average Rg values of WT, WT-PLM and MT-PLM from the trajectory were calculated as 2.2, 2.21 and 2.18 nm respectively see Fig. 2a. Throughout the simulation, the Wild type CD36 and complex wild type CD36-PLM are marked with black and blue lines respectively, exhibiting almost a similar pattern in terms of the Rg values on the contrary the complex Mutant Type CD36m-PLM which was marked with red line; This suggested that there was a change in the global conformation and coil of the preceptor after the mutation. The K164A mutation shows a significant decrease in Rg, implying a strong compactness that will influence the closure of the fatty acid transport cavity.

Moreover, in order to search for the hydrophobic center behavior of the wild type CD36 (WT) and mutant type (MT) structures in presence and absence of PLM, we had investigated the solvent surface accessibility area (SASA). The analysis of SASA displayed the residues of the complex MT-PLM structures CD36 (K634A) which had the same SASA estimation to the complex WT-PLM CD36 protein, and they both conserved the openness at 44 ns Fig. 1a and upper to WT in the absence of PLM. PLM made this amino acid more accessible to the solvent. An important decrease of the SASA mean was verified in the complex WT CD36-PLM. The CD36 K164A mutation which has brought about larger changes in the SASA value because it had a great surface area exposed to the solvent, this could cause an exposure of hydrophobic amino acids and subsequently the deployment of the receptor.

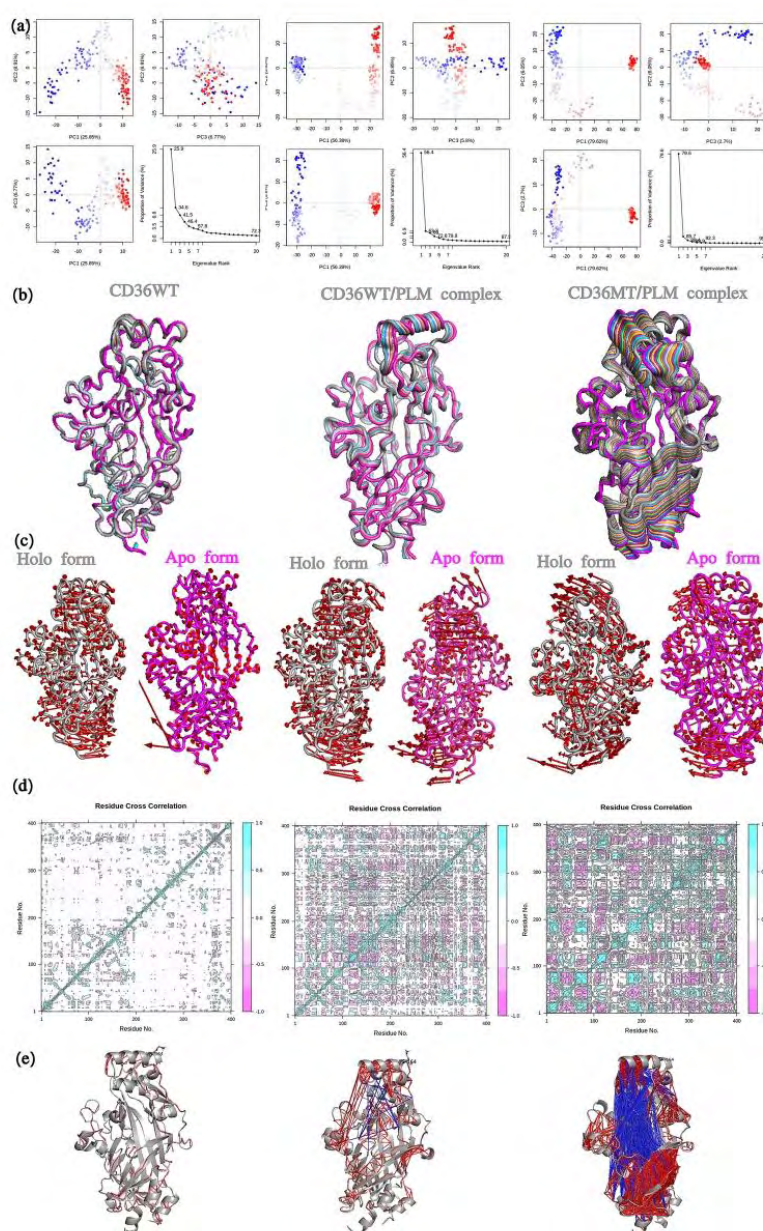
A calculation of hydrogen bond (HB) in CD36 WT, WT-PLM and MT-PLM mutants was given to interpret the stability and flexibility of the receptor. Figure 2a shows the quantity of intra molecular HB in the WT and mutant CD36 receptor in presence and absence of PLM throughout the simulation. The total of HB decreased in complex CD36 WT-PLM and MT-PLM mutants as compared to the WT at 40 ns. WT CD36 has 340 HB, whereas complex WT-PLM and MT-PLM have 320 HBs, respectively. This diminution in HB was considered a loss of conformation and compactness of the CD36 receptor in presence of PLM. Overall, a significant diminution MT-PLM after a simulation time of 40 ns disrupts the structure, dynamics, and the stability, which leads to a deregulated and dysfunctional CD36-PLM complex. Throughout the period of simulation, the energetic value of the complex WT-PLM is an indication of strong basis of the fact that the molecule has a stable structure as well as relatively higher amount of total interaction compared to complex MT-PLM. Destabilization in folding and thermodynamic stability may affect the total energy of the biomolecules. The results demonstrated a very slight rise in the total energy value in mutant structure this ruled out the notion of stability losses induced by mutation.

The secondary structural propensities become significant in exploring the structural properties of the protein. However, there were significant structural variations measured at each simulation time. These outcomes were consistent with the secondary structure evaluations with the DSSP utilizing the Fig. 2b. Therefore, this gave us a clear evidence that the secondary and tertiary constructing ( $\alpha$ -helix and  $\beta$ -sheets) of the CD36 WT-PLM and MT-PLM structures were not stable during the simulation. The mutation K164A induced change in the  $\alpha$ -helix content but it delayed at 70 ns on the contrary in WT-PLM at 40 ns. Our MD analysis has clearly showed that K164 change conformational, dynamics, and flexibility might be a possible cause of transport of fatty acid.

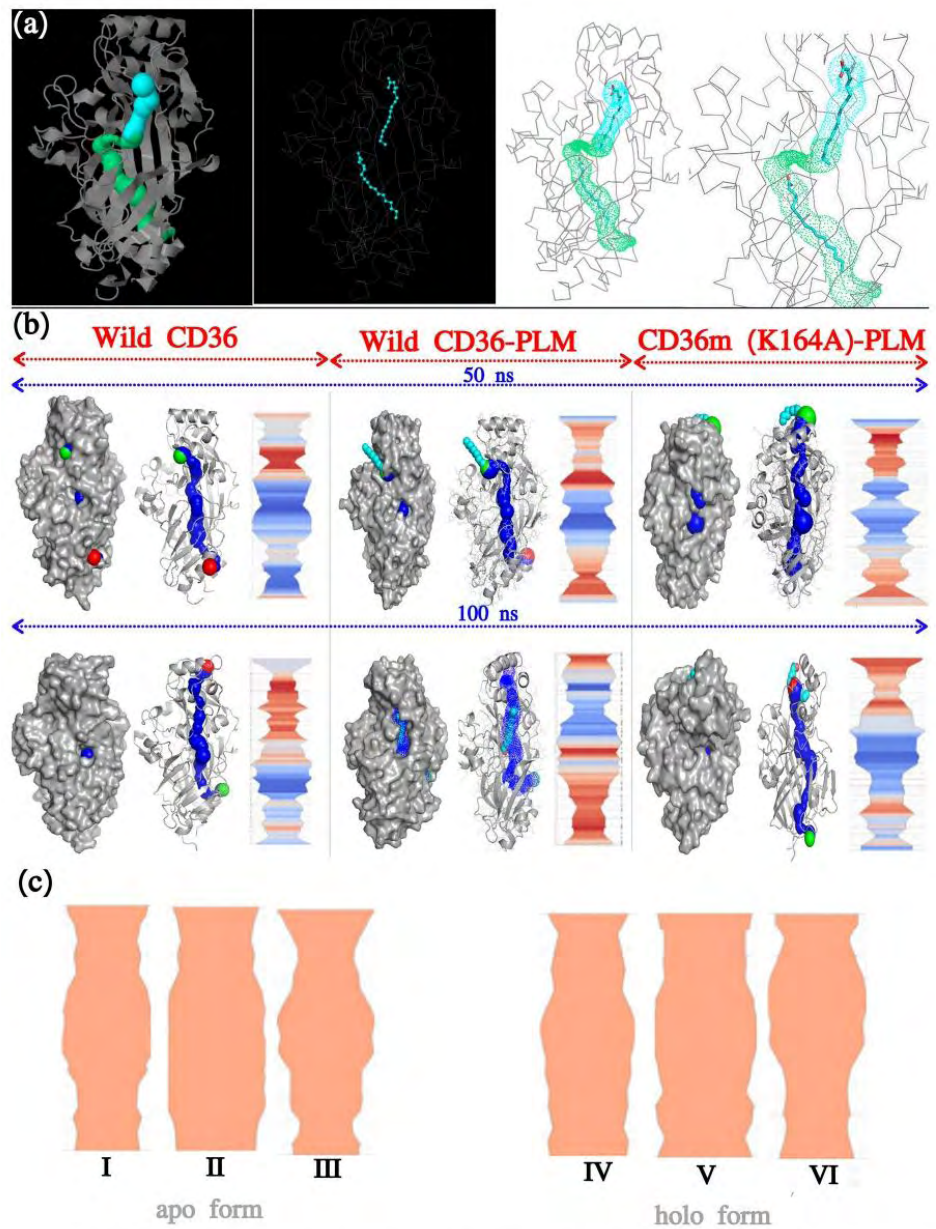
**Principal component analysis PCA.** We use the Principal component analysis (PCA) in order to identify the motions and the direction of the flexibilities in the residues. The principal components analysis was applied to the backbone atoms for CD36 WT, WT-PLM and MT-PLM such analysis showed that the first account 25%, 56% and 79% of variances of the total variance observed in the simulation data respectively Fig. 3a.

The first outstanding featured motions during the simulation were analyzed through secondary structures which includes an opening-closing motion of transport tunnel. Our observation focused on the movements of the C-coil,  $\alpha$ -helix. We found, that the flexible motions of these secondary structures in the distinct active states are quite dissimilar from CD36 wild type (WT) in absence of PLM and complexes CD36 WT-PLM and MT-PLM Fig. 3a–c. The top principal components for model PC1 for CD36 WT in absence PLM, WT-PLM and MT-PLM are illustrated in Fig. 3b. In the top two eigenvectors, the most prominent motions observed are propeller loops of CD36 MT-PLM which are the most flexible in moving in the opposite direction. Particularly, in holo structure the propeller loops attached to the upper, moves leftwards or rightwards for CD36WT and it moves in the opposite direction upwards to downwards for the CD36 MT in presence of PLM. In contrast, the helix and coil at termination K164 has an expanding or contracting movements in the WT in the presence of PLM, which is the opposite in the WT only in absence of PLM. Inclusive, the mutation results in a change in the movement patterns of these secondary elements. We have visualized this movement of the top two eigenvectors Fig. 3c: they show the orientation and the volume of selected eigenvectors by each of the backbone atoms.





**Figure 3.** (a–c) Model and motion produced by principal component analysis for wild and mutant type of CD36 in presence and absence of PLM a function of time during 100,000 ps (100 ns). (a) Principal component analysis, PC2 vs PC1, PC2 vs PC3, PC3 vs PC1 and an eigenvalue rank plot. (b) Visualization of the motions along PC1. (c) Motion of two different modes by ANM Analyze, the red arrows show the direction and amplitudes of the movements occurred from first trajectory of holo and apo form in (gray) and (pink) color respectively. (d,e) Cross-correlation Analysis matrix of the fluctuations of the C $\alpha$  atoms of the residue from their average during 100 ns for wild and mutant type of CD36 in presence and absence of PLM. (d) Cross-correlation map for MD simulation. Cyan color indicates a positive correlation between residue fluctuations, while the pink color depicts a negative correlation. (e) Residue cross-correlations. CD36 is depicted by gray cartoon. Red and blue lines indicate correlated and anti-correlation motions.



**Figure 4.** Structure of the studied CD36 with their respective tunnels. (a) PLM-bound snapshot, with the crystal structure of the wild-type CD36 (CD36-CIDRa, PDB-ID: 5LGD) depicting fully formed tunnel (identified using CAVER 3.0 PyMOL plugin). The main tunnel is shown by the blue and green surface. (b) CHEXVIS was used to analyze snapshots from 50 and 100 ns molecular dynamics simulation for six simulations for wild and mutant type of CD36 in presence and absence of PLM. (c) The transport channel of the fatty acid in the apo and holo conformations were identified based on the PCA analysis using CHEXVIS for CD36 WT, CD36WT-PLM and CD36MT-PLM.



Simulations	Rank	Pt 1	Pt 2	Length	Bottle-neck	Straightness	Score
Wild CD36 (holo form)	1	2	7	59.978	4.059	0.903	0.844
Wild CD36-PLM (holo form)	1	3	4	55.923	4.243	0.876	0.852
CD36m-PLM (holo form)	1	1	9	62.427	3.592	0.821	0.795
Wild CD36 (apo form)	1	3	10	48.392	3.746	0.858	0.826
Wild CD36-PLM (apo form)	1	4	6	59.430	4.264	0.895	0.835
CD36m-PA (apo form)	1	5	7	75.344	3.754	0.769	0.868

**Table 1.** Cavity's profile as seen by the ChExVis program.

Although, there might be differences between the simulations with respect to the movement sampled, it is evident that this may suggest that the site K164 is able to settle the paradigm by reducing the movement of the loops to a reduce module.

**Covariance analysis and motions.** For a preferable comprehension of the association through the movement of the upper helix, coil and Lys164 residue, as well as the movements of the adjoining part, covariance analyses were executed during each of the separate simulation on the three groups CD36 WT, WT-PLM and MT-PLM. They were illustrated utilizing 3D system plots. As shown in Fig. 3d, positive correlations are mapped in light blue green color while negative correlations are mapped in pink color. The deeper color indicates more correlations (or anti-correlations).

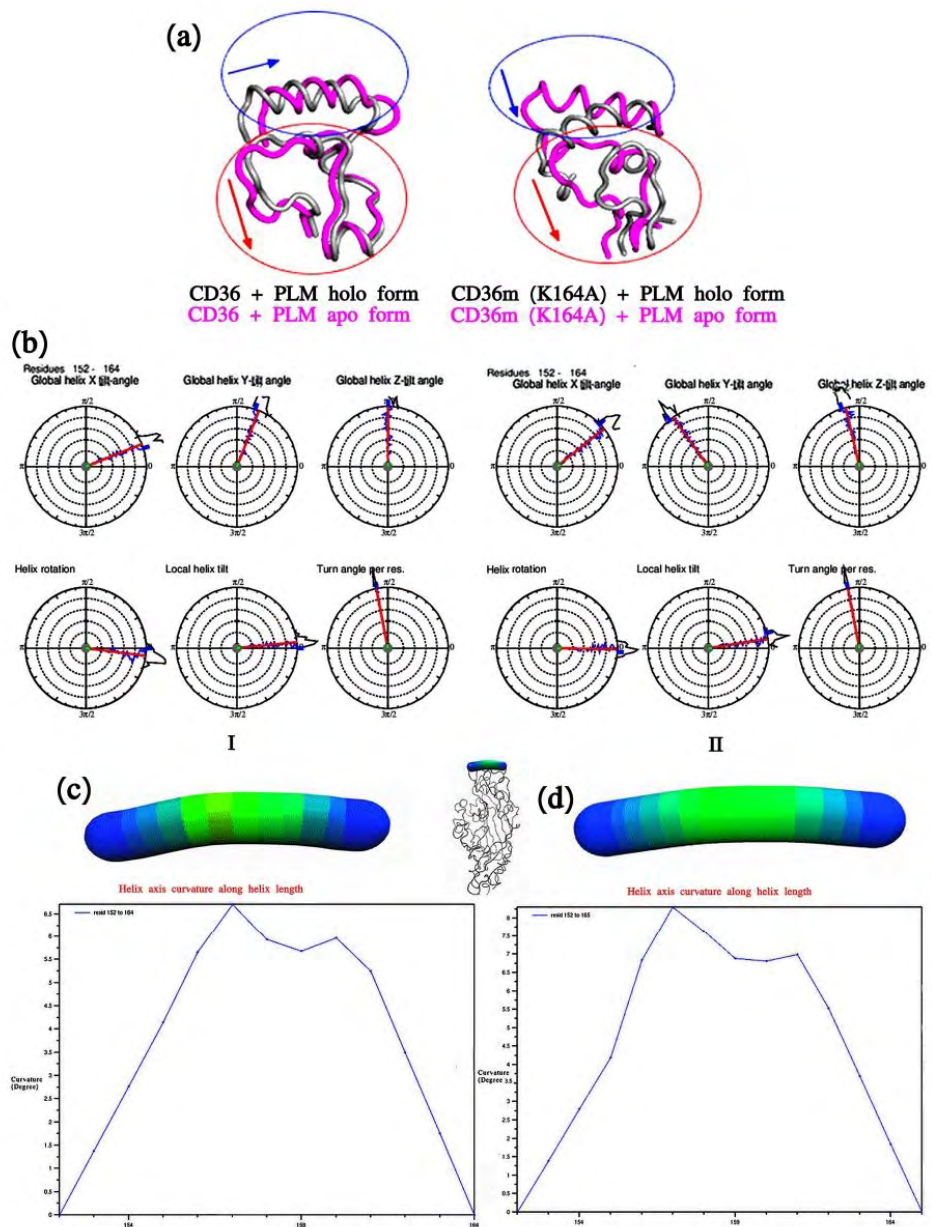
We detected a cluster with correlated motions in the middle the secondary structures and the adjoining zone in the active situation of the related groups. The correlated motions of these parts grow extremely in the CD36MT-PLM state Fig. 3e correlation and anticorrelation resented in red and blue color line. However, the mutation K164A increased the anticorrelated motions between the upper helix and C-terminal and N terminal (tunnel). In addition, strongly correlated motions were also observed within the beta sheets regions and helix C-terminal and N-terminal regions, which signal that the conjunction of these parts retired and proposed that the flexibility of the helix coil is in charge for the movements of the adjoining pathways. Besides, the movements of these two secondary structures are correlated with the movement of upper helix that is essential for the structural stability of CD36 structure for PLM transport.

In the next parts, we revealed cavities in the several conformations during the trajectory for the six MD. A more functional study has indicate the elaborated organizational connection in the midst the motions of the upper helix, and coils, along with the movements of the adjoining conduits. The channels in holo conformations of CD36 crystal structure.

**Transport channel analysis.** The channels in holo conformations of CD36 crystal structure with two co-crystal of PLM<sup>14</sup> were identified and analyzed using CAVER (Fig. 4a) in order to determine the associated motion between upper helix and the transport channel. To examine mutation effect K164A and active area of transport channel bottleneck amino acids were got by study the Channel Extraction and Visualization (CHEXVIS) during MD simulations this observed the conformation constituents Table 1.

CHEXVIS was used to study Scenes from 100 ns MD for CD36WT, CD36WT-PLM and CD36MT-PLM; Fig. 4b represents tunnel width, length, and bottleneck width after 50 ns and 100 ns. The width indicates the extreme flexible domain for fatty acid, the bottleneck can elucidate the hardness of absorption of fatty acid from receptor hole. A reduce bottleneck showed in CD36MT-PLM complex signals the problem of the receptor in absorption and transporting the fatty acid. From the above results, it can be concluded that K164A influence helix conformation and the ability of the site closing and binding and the transport of PLM. The transport channel of the fatty acid in the apo and holo conformations were identified based on the PCA analysis using CHEXVIS for CD36WT, CD36WT-PLM and CD36MT-PLM Table 1 and Fig. 4c. The width channel profiles in 2D of holo structure has enlarged, that demonstrate the unlocking of the channel, as indicate in Fig. 4c,IV,V,I (The bottle-neck 4.059, 4.243, 3.592 and Length 59.978, 55.923, 62.427 for CD36 WT, WT-PLM and MT-PLM channel respectively) In contrast, the channel in the apo conformation showed a decline in the width of channel especially at bottle-neck level of, as shown in Fig. 4c,I,II,III).The bottle-neck 3.746, 4.264, 3.754 and Length 48.392, 59.430, 75.344 for WT, WT-PLM and MT-PLM channel respectively. In addition, only the width of the channel for CD36 MT-PLM of holo form was very small in bottle-neck Fig. 4c,VI remains in the inactive state correspond to an inaccessible pathway for substrate. It also it indicate that the mutation influences the conduit unlocking-locking by modifying the movement of the upper helix and loops. However, an issue appears related the way the mutation influences the movement of the helix and loops.

**Analysis of properties of helix.** To determine the associated motion between the Lys164 residue and C-coil and the surrounding channels, a precedent research proposed that the upper helix plays a function in stability between the unlocking and locking structures of CD36. The K164A mutation based on the PCA analysis, both the coil and the helix of the wild-types and mutants switch in contrary orientations in apo and holo form Fig. 5a. In order to provide the detailed view of the observed transitions, we have analyzed the MD trajectories for CD36 WT-PLM and MT-PLM complexes by calculating the average geometrical parameters of upper helix using Simulaid software. K164 residues, therefore, directly contribute to its global helix that was turned towards



**Figure 5.** Properties of Protein helix during MD simulations. **(a)** Motion analysis based on the PCA analysis in apo (pink color) and holo (gray color) form of CD36 WT and MT in presence of PLM. **(b)** The geometrical parameters used for the analysis of the upper helix conformations. The values of the rotation angle calculated for the upper helices are plotted as a function of time for one of the MD simulations of CD36 WT (I) and MT (II) in presence of PLM. The angle is estimated with respect to the original inward facing structure. The blue ticks in the source and on the side of the circles measured the angle value for the first and last trajectory frames respectively; the red line corresponds to the mean angle value along the trajectory. **(c,d)** Analyze a static peptide's local angles along the length of the helix and heatmap color represented an angle-indicative graphics of CD36 wild type **(c)** and CD36 mutant type **(d)** in presence of PLM.



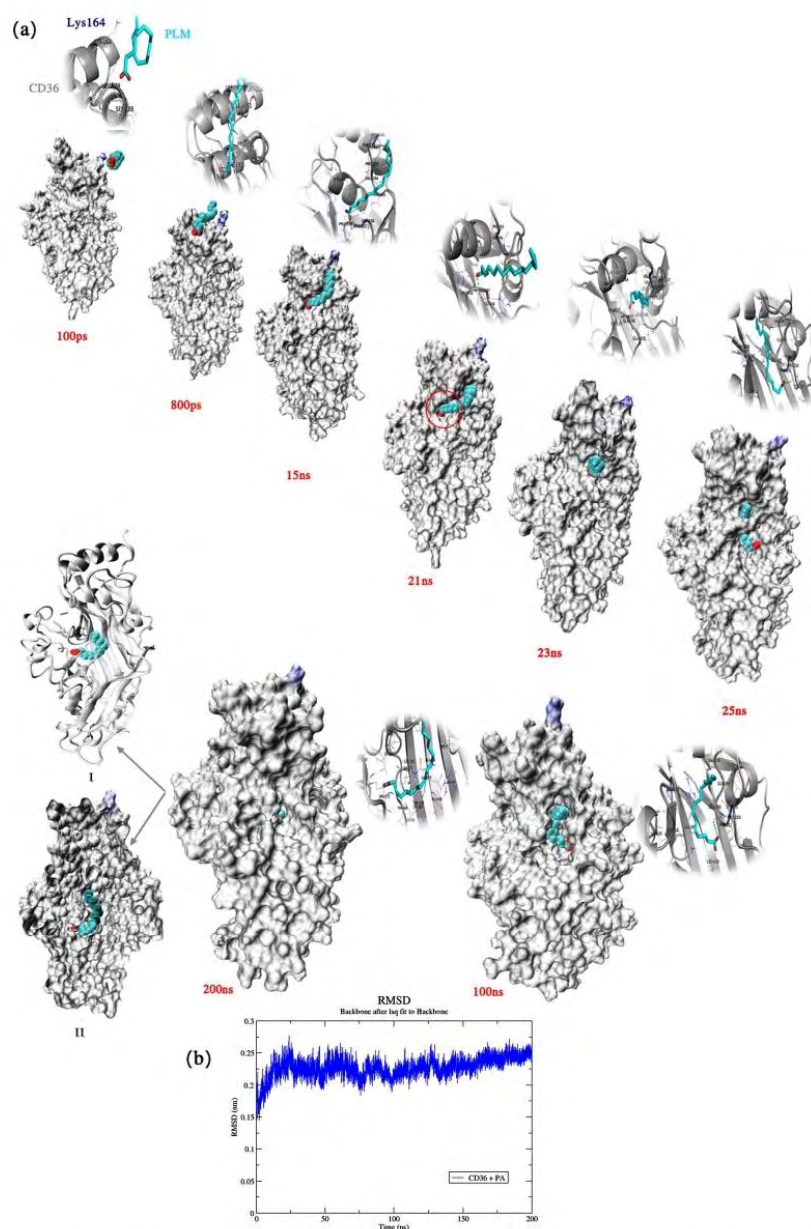
the right of Y axis Fig. 5b,I, while the mutation A164 mobility turned towards the left of Y and Z axis Fig. 5b,II. Lysine side chains, allowed the detection of their rotation in our simulations. As a result, it retains the helix rotation, local helix tilt and turn helix per residue.

To study the mechanism of the rotation in our simulations of the mutation on the upper helix of CD36 proteins, we concentrated on the structure analysis of upper helix using Bendix analyses it analyzed static peptide's local angles along the length of the helix showed in Fig. 5c,d. The result of the analyses of static peptide's local angles along the length of the helix had 152–164 residue and the helical angles were measured with 6.5 curvature degree for WT-PLM Fig. 5c in comparison with MT-PLM Fig. 5d the length of the helix 152–165 residue. And the helical angles were measured with 8 curvature degree. On the whole, our investigation of upper helix proposes that the upper helix at the MT-PLM mutate situation of CD36 receptor tends to be wrapped, that influences the movement of the helix and loops of opening fatty acid transport site as revealed by the covariance analysis.

**Portals PLM transport and pathway.** In the X-ray structure of CD36-complex, the first PLM occupied the channel connecting the helical portal region and the inner cavity thus<sup>14</sup>, if the helical portal is the only choice to enter the cavity, does it exist an alternative portal that might allow the direct exchange between PLM? But there was no direct evidence showing the feasibility of such a process. In order to evaluate the feasibility of the structural dynamics underlying this hypothesis was directly applied to only PLM in ligated protein complex. A successful penetration of PLM from principal portal was observed Fig. 6.

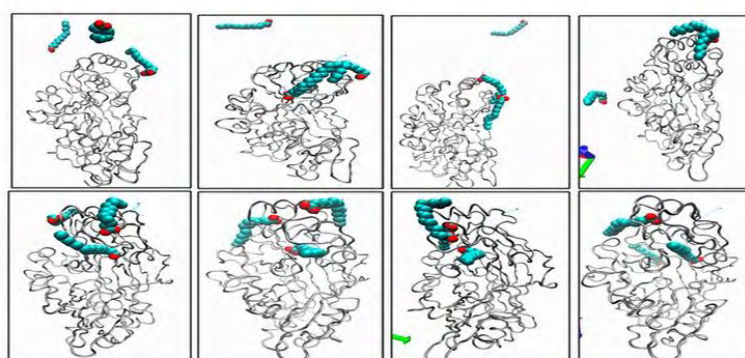
Figure 6 presents snapshots at 100 ps, 800 ps, 15 ns, 21 ns, 23 ns, 25 ns, 100 ns, 200 ns presenting the pathway of PLM transport at the CD36 taken during a simulation, where one palmitic acid (PLM) were randomly located in the simulation box together with the crystal CD36 (PDB code: 5LGD) and records the RMSD calculated for the whole system as it evolves with time. At about 100 ps after the initiation of the simulation Fig. 6a the PLM hydrophobic tail joined the Lys164 residue first one colored blue, due to hydrophobic and the electrostatic interaction to attach to the protein. 800 ps later, the tight interaction between PLM and Lys164, stimulated the conformation change caused by partially penetrating from the portal region initially. Then, it tried to slide over a loop dissociated from the gap between a loop and a-helix Fig. 6a (red circle) by the carboxylic part of a PLM molecule in the inner cavity of the protein. 15–150 ns later the PLM had accessed the inner channel and finished its way till the termination of the simulation. During the 200 ns/time, the FA had approached the central inner cavity protein. Residues, which were commonly encountered by the Palmitic acid during expulsion, were important for the constitution of individual portals. On the starting protein structure of our MD simulations, the constituting residues, identifying these residues, that are highlighted Fig. 6a are Arg63, Phe65, leu140, Ala143, Ala144, Ile148, Leu161, Ser162, Lys164, Leu187, Leu188, Leu189, Pro191, TYR192, Pro193, Val194, Thr195, Val198, Thr197, Gly199, Leu200, Phe201, Pro203, Tyr204, Thr207, Ser268, Ser269, Asp270, Ile271, Arg273, Ile330, Lys334, Glu335, Gly336, Arg337, Pro338, Ser339, Val339, Lys385, Leu387, Val389, Leu343, Ile341, Thr369, Leu391, Leu416, Glu418. Small openings at the portal regions were closed in the starting structure, however, certain parts of portal (e.g. Arg337, Thr195, Val194 at loop and Ala144, Ile148 at upper helical) showing the potentiality of creating an opening at this region. Since the PLM molecule bind in to Arg337 was close to this portal, so it would be the most favorable route. This portal was the most recorded one in multiple repetitions of MD simulations; as a result this portal is believed to be the only choice for Arg337. We therefore conducted MD simulations of 200 ns on wild type in presence of one PLM as shown in the RMSD profile Fig. 6b. The plots of RMSD revealed to be stable during simulation. The result indicates that Lys164 residue which provides the structural basis of forming an opening at the region of the principal portal for the dissociation of palmitic acid molecules from this path.

**Fatty acid transport rates.** Here we described the outcome of MD simulations, using high concentrations of palmitate molecules. In the X-ray structure of CD36 complex, two PLM crystallized as described elsewhere occupied the channel connecting the portal area<sup>14</sup>. Consequently, there were no hypothetical proposals of how it would absorb from the receptor and even the maximal rates of palmitic uptake. To determine the rate of the fatty acid transported via the CD36 receptor and it existence other intermediate portals were entry portals using the same MD parameters with those for one PLM. Figure 7a presents snapshots at 100 ps, 50,100, 150, and 200 ns taken during each simulation, which depicted the pathway, transport mode and rate of three PLM to the portal CD36 receptor. In this simulation, 200 ns long, a crystal human CD36 (PDB code 5LGD) was equilibrated with three palmitic acid (PLM) that were placed in the simulation box at random locations outside. Figure 7b records the RMSD, RMSF, Rg calculated for the whole systems for three PLM outside and three PLM (two inside and one outside) as it evolves with time simulation. At the beginning we started with simulation (5 ns) Fig. 7a, the first PLM approximated to Lys164 after, a second PLM molecule continued and the third molecule moved away of the protein range electrostatic attraction of one of the carboxylates to the positive charge of the Lys164 amine side chain. 100 ns later, the first PLM moved towards the central portal zone due to hydrophobic reaction and second continued and one single PLM into the interior hole of the receptor. The third PLM had stayed in a liberate structure until the stop of the simulation. So, it formed an dimmer in conflict connected to the receptor. During this time, the liberate PLM had approached the receptor a shortly, but none of them led to the creation of a stable structure. To ascertain that the encounter between the protein and the two PLM was not accidental; we repeated the simulation using the toad crystal structure (PDB Id: 5LGD) and the pre-formed of two PLM inside and one outside molecules. We found out that the PLM outside was not absorbed till the simulation was terminated at 40 ns. As seen in Fig. 7 after the results of the dynamics of the system with three PLMs that the protein had transported two PLMs one by one to confirm these results we determined that the encounter between the protein and the two PLM was not accidental. Indeed, the simulation was repeated using the toad the same crystal structure and a pre-formed of two PLM inside and one outside the molecules. The PLM outside was dissociated

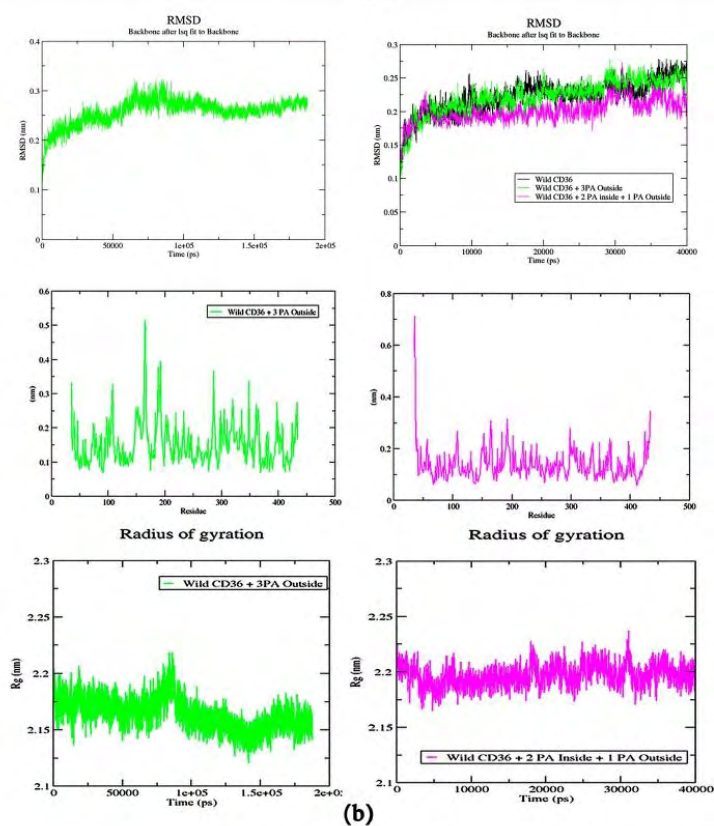


**Figure 6.** A snapshots representing the pathway of palmitic acid al transport at the portal of the CD36. (a) The snapshots correspond with 100 ps, 800 ps, 15 ns, 21 ns, 23 ns, 25 ns, 100 ns, 200 ns with one FA adsorbed to the protein. (b) The RMSD of the system, as they evolve with time. A is representation of Fatty acids at the starting point of the simulation 200 ns.





(a)



(b)

**Figure 7.** (a) Snapshots representing of Fatty acid transport rates, snapshots at 100 ps, 50, 100, 150, and 200 ns taken during each simulation, which depicted the pathway, transport mode and rate of three PLM to the portal CD36 receptor. (b) MD analyses of three PLM (three put outside two PLM put inside (green color) and one PLM outside (pink color)) Uptake across the wild type CD36 receptor. RMSD plot, Backbone residual fluctuations (RMSF) plot, time evolution of radius of gyration (Rg).

till the simulation was terminated at 18 ns. Another system had to be carried out to put two PLMs inside and one outside after the simulation of the third PLM always remained outside and according to the RMSD, RMSF and radius of gyration (Rg) analysis it was always stable by contribution to the first simulation three PLM ext. The fluctuations of protein structure variation, in the time of the absorption from the three PLM, was observed in Fig. 7b (The green and pink curves respectively).

The principal structure fluctuation, as expected, occurred at the proximity of the specific portal. For the green line, the augmentation of RMSF at the upper helix region and coil was thought to be affected by the amino acids impacts of the starting tries to exit from this area. The domain with the growing RMSF, immediately participated by the ligand entrance from the gateway, K164 helix I loop, surrounding the bottom of the cavity. The outside absorption of the three PLM from this gateway had imposed a structure conformational modification contrariwise of the system two PLM one inside and one outside. Clearly, the two PLM are the most preferred to CD36 structure stabilization. Our CD36 structure model therefore supported a model in which fatty acids move through the cavity, identifying just one main possible entrance and favorable positions through which translocating ligands moved as they went towards the membrane proximal exit and the plasma membrane.

## Conclusion

In our research we have applied MD simulations to obtain a picture of fatty acid uptake rates model systems via CD36 at a molecular level. Models which are consisting of palmitic acid and CD36 receptor were investigated.

The aim of our study was to gain more understanding of the role of Lys164 residue in Fatty acid uptakes and to find out the mechanism of the functioning of the complex CD36/Fatty acids. From the different simulations done we have obtained a picture of the CD36 wild and mutant K164 in the presence and in the absence of fatty acid PLM. These simulations reveal both the conformational changes induced by Lys164 residue and the altered interactions induced by the mutagenesis of surface lysine, adversely affecting protein folding. All this was done in accordance with numerous investigations, where the results have shown that mutation K164A has decreased the productivity of the function of the protein transportation. Our MD analysis demonstrates the mode of global uptake Fatty acid transport via CD36 receptor. The parameters which were calculated from the theoretical work are in good qualitative agreement with the experimental results. The results of this study have provided the confidence to extend our fatty acid uptake rate model for further MD studies.

## Computational methods

**Preparation of wild type and mutant CD36 protein for MD.** The tridimensional structures of virtual mutant K164A ectodomain of CD36 was predicted by using SwissModel server to construct the homology model based on the crystal structure CD36-CIDRa (PDB ID: 5LGD) which is obtained from the Protein Data Bank, where all ligands and water molecules in the protein crystal structure were deleted<sup>26,27</sup>. The entire sequence of human CD36 was taken from the UniProt Database (<http://www.uniprot.org>) with accession number P16671. The quality of the homology model was evaluated by Ramachandran plot<sup>28</sup>, and its overall quality factor ERRAT<sup>29</sup> plot and PROVE produced by the SAVS server v6.0 (<https://saves.mbi.ucla.edu/>).

**Molecular dynamics fatty acid transport protocols.** All procedures for MD simulations led out for 100, 200 ns using the GROMACS program, version 4.5.5, using the GROMOS56A1 force field<sup>30,31</sup> were simulated under fully periodic boundary conditions. All electrostatic reactions were processed including Particle Mesh Ewald (PME). So every bonds between hydrogen molecules was recondition utilizing the LINCS algorithm<sup>32</sup>. Water organization was constrained using rigid SETTLE algorithm<sup>33</sup> and non-bonded reactions, a cut-based is used, which allowed us to use an integration step of 2 fs. The ligand topology files, as needed in GROMACS, were generated from the PRODRG server<sup>34</sup> which was an automated server for topology generation (<https://www.sites.google.com/site/vanaaltenlab/>). The complex was solubilized using SPC in water and balanced by the raise of K<sup>+</sup> and Cl<sup>-</sup> ions in a cubic and dodecahedron box.

Five different simulations were carried (as summarized in Fig. 1a). In the three first simulations MD, the wild and mutant type protein was simulated in the presence and absence of a single palmitic acid, located near its surface. In the last two simulations, three palmitic acids were randomly put in the simulation cell and the last simulation (three PLM put outside) and (two PLM put inside and one PLM outside) in protein surface of wild type.

MD simulation of each complex was performed in five consecutive stages: (1) energy minimization, (2) NVT heating, (3) NPT equilibration, (4) NPT pre-production simulation, and (5) production simulation. Besides, the time step for the simulations was 2 fs. All MD calculations for period 100 ns and 200 ns. The RMSD, RMSF, Rg, and SASA calculated using GROMACS module. Furthermore, the evolution of the secondary structure was followed using the DSSP and VMD programs<sup>35,36</sup>. Visual depiction was taken out employing PyMOL and VMD.

**Principal component analysis and PCA cross correlation analysis.** The principal component analysis (PCA) and cross-correlation were performed using the bio3D package in R and the anisotropic network model (ENM) was generated with PyANM plugin in PyMOL<sup>37,38</sup>. Once a sufficient MD production runs had been completed, the most important component analysis is used to analyze the results from the MD simulations. PCA became an efficient and an excellent tool for the global extract motions and the correlated structural dynamics across many biological systems.

**Channel profiles analysis of the wild and mutant K164A type CD36 receptor.** CHEXVIS was used to analyze visual framework from 100 ns MD simulation of wild type (WT) and mutant K164A (MT) CD36 receptor in both the presence and absence of PLM, which revealed the structural details of the tunnel which are



the width, length, and bottleneck distance of the tunnel<sup>39</sup>. We used CAVER 3.0 35 to provide detailed characteristics of crystal transport pathways<sup>40</sup>.

**Analysis of properties of helix.** Analysis of the rotational movements of upper helice was performed with help of the TRAJELIX tool of the Simulaid package (<https://mezeim01.u.hpc.mssm.edu/simulaid/>) Earlier versions of Simulaid (IRIS GL GRAPHICS VERSION) and VMD Bendix Plugin, Version 1.1<sup>41–43</sup>.

Received: 18 June 2021; Accepted: 21 September 2021

Published online: 01 December 2021

## References

- Silverstein, R. L. & Febbraio, M. CD36, a scavenger receptor involved in immunity, metabolism, angiogenesis, and behavior. *Sci. Signal.* **2**, re3. <https://doi.org/10.1126/scisignal.272re3> (2009).
- McArthur, M. J. *et al.* Cellular uptake and intracellular trafficking of long chain fatty acids. *J. Lipid Res.* **40**, 1371–1383. [https://doi.org/10.1016/S0022-2275\(20\)33379-4](https://doi.org/10.1016/S0022-2275(20)33379-4) (1999).
- Magnetto, S. *et al.* CD36 mediates binding of soluble thrombospondin-1 but not cell adhesion and haptotaxis on immobilized thrombospondin-1. *Cell Biochem. Funct.* **16**, 211–221. [https://doi.org/10.1002/\(SICI\)1099-0844\(199809\)16:3%3c211::AID-CBF788%3e3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1099-0844(199809)16:3%3c211::AID-CBF788%3e3.0.CO;2-Z) (1998).
- Magwenzi, S. *et al.* Oxidized LDL activates blood platelets through CD36/NOX2-mediated inhibition of the cGMP/protein kinase G signaling cascade. *Blood* **125**, 2693–2703. <https://doi.org/10.1182/blood-2014-05-574491> (2015).
- Nicholson, A. C., Febbraio, M., Han, J., Silverstein, R. L. & Hajjar, D. P. CD36 in atherosclerosis. The role of a class B macrophage scavenger receptor. *Ann. N. Y. Acad. Sci.* **902**, 128–131 (2000).
- Febbraio, M. *et al.* Targeted disruption of the class B scavenger receptor CD36 protects against atherosclerotic lesion development in mice. *J. Clin. Invest.* **105**, 1049–1056. <https://doi.org/10.1172/JCI9259> (2000).
- Silverstein, R. L. & Febbraio, M. CD36 and atherosclerosis. *Curr. Opin. Lipidol.* **11**, 483–491. <https://doi.org/10.1097/00041433-200010000-00006> (2000).
- Kuda, O. *et al.* CD36 protein is involved in store-operated calcium flux, phospholipase A2 activation, and production of prostaglandin E2. *J. Biol. Chem.* **286**(20), 17785–17795. <https://doi.org/10.1074/jbc.M111.232975> (2011).
- Dramane, G. *et al.* STIM1 regulates calcium signaling in taste bud cells and preference for fat in mice. *J. Clin. Invest.* **122**(6), 2267–2282. <https://doi.org/10.1172/JCI59953> (2012).
- Silverstein, R. L., Li, W., Park, Y. M. & Rahaman, S. O. Mechanisms of cell signaling by the scavenger receptor CD36: Implications in atherosclerosis and thrombosis. *Trans. Am. Clin. Climatol. Assoc.* **121**, 206–220 (2010).
- Rahaman, S. O., Zhou, G. & Silverstein, R. L. Vav protein guanine nucleotide exchange factor regulates CD36 protein-mediated macrophage foam cell formation via calcium and dynamin-dependent processes. *J. Biol. Chem.* **286**, 36011–36019. <https://doi.org/10.1074/jbc.M111.265082> (2011).
- Wang, D. D., Li, W. G. & Zhang, Y. Y. Association of CD36 expression and polymorphism with serum biochemical indices in Cherry Valley duck. *Genet. Mol. Res.* <https://doi.org/10.4238/gmr.15027994> (2016).
- Love-Gregory, L. & Abumrad, N. A. CD36 genetics and the metabolic complications of obesity. *Curr. Opin. Clin. Nutr. Metab. Care.* **14**(6), 527–534. <https://doi.org/10.1097/MCO.0b013e32834bbac9> (2011).
- Hsieh, F. L. *et al.* The structural basis for CD36 binding by the malaria parasite. *Nat. Commun.* **7**(1), 1–11. <https://doi.org/10.1038/ncomms12837> (2016).
- Gu, X. *et al.* The efficient cellular uptake of high density lipoprotein lipids via scavenger receptor class B type I requires not only receptor-mediated surface binding but also receptor-specific lipid transfer mediated by its extracellular domain. *J. Biol. Chem.* **273**(41), 26338–26348. <https://doi.org/10.1074/jbc.273.41.26338>. Erratum in: *J. Biol. Chem.* **273**(52), 35388 (1998).
- Demers, A. *et al.* Identification of the growth hormone-releasing peptide binding site in CD36: A photoaffinity cross-linking study. *Biochem. J.* **382**(Pt 2), 417–424. <https://doi.org/10.1042/BJ20040036> (2004).
- Frieda, S., Pearce, A., Wu, J. & Silverstein, R. L. Recombinant GST/CD36 fusion proteins define a thrombospondin binding domain. Evidence for a single calcium-dependent binding site on CD36. *J. Biol. Chem.* **270**(7), 2981–2986. <https://doi.org/10.1074/jbc.270.7.2981> (1995).
- Neculai, D. *et al.* Structure of LIMP-2 provides functional insights with implications for SR-BI and CD36. *Nature* **504**(7478), 172–176. <https://doi.org/10.1038/nature12684> (2013).
- Marie, T. *et al.* Assessment of key amino-acid residues of CD36 in specific binding interaction with an oxidized low-density lipoprotein. *Biosci. Biotechnol. Biochem.* **77**, 1134–1137. <https://doi.org/10.1271/bbb.130072> (2013).
- Kuda, O. *et al.* Sulfo-N-succinimidyl oleate (SSO) inhibits fatty acid uptake and signaling for intracellular calcium via binding CD36 lysine 164: SSO also inhibits oxidized low density lipoprotein uptake by macrophages. *J. Biol. Chem.* **288**(22), 15547–15555. <https://doi.org/10.1074/jbc.M113.473298> (2013).
- Coort, S. L. *et al.* Sulfo-N-succinimidyl esters of long chain fatty acids specifically inhibit fatty acid translocase (FAT/CD36)-mediated cellular fatty acid uptake. *Mol. Cell Biochem.* **239**(1–2), 213–219 (2002).
- Baillie, A. G., Coburn, C. T. & Abumrad, N. A. Reversible binding of long-chain fatty acids to purified FAT, the adipose CD36 homolog. *J. Membr. Biol.* **153**(1), 75–81. <https://doi.org/10.1007/s002329900111> (1996).
- Nicholls, H. T. *et al.* Hematopoietic cell-restricted deletion of CD36 reduces high-fat diet-induced macrophage infiltration and improves insulin signaling in adipose tissue. *Diabetes* **60**(4), 1100–1110. <https://doi.org/10.2337/db10-1353> (2011).
- Wlodawer, A. Stereochemistry and validation of macromolecular structures. *Methods Mol. Biol.* **1607**, 595–610. [https://doi.org/10.1007/978-1-4939-7000-1\\_24](https://doi.org/10.1007/978-1-4939-7000-1_24) (2017).
- Kar, N. S., Ashraf, M. Z., Valiyaveetil, M. & Podrez, E. A. Mapping and characterization of the binding site for specific oxidized phospholipids and oxidized low density lipoprotein of scavenger receptor CD36. *J. Biol. Chem.* **283**(13), 8765–8771. <https://doi.org/10.1074/jbc.M709195200> (2008).
- Waterhouse, A. *et al.* SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**(W1), W296–W303. <https://doi.org/10.1093/nar/gky427> (2018).
- Lovell, S. C. *et al.* Structure validation by Alpha geometry: Phi, psi and Cbeta deviation. *Proteins* **50**(3), 437–450. <https://doi.org/10.1002/prot.10286> (2003).
- Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7**, 95–99. [https://doi.org/10.1016/s0022-2836\(63\)80023-6](https://doi.org/10.1016/s0022-2836(63)80023-6) (1963).
- Colovos, C. & Yeates, T. O. Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Sci.* **2**(9), 1511–1519. <https://doi.org/10.1002/pro.5560020916> (1993).
- Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25. <https://doi.org/10.1016/j.softx.2015.06.001> (2015).

31. Scott, W. *et al.* The GROMOS biomolecular simulation program package. *J. Phys. Chem. A* **103**, 3596–3607. <https://doi.org/10.1021/jp984217f> (1999).
32. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comp. Chem.* **18**, 1463–1472 (1997).
33. Miyamoto, S. & Kollman, P. A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comp. Chem.* **13**, 952–962. <https://doi.org/10.1002/jcc.540130805> (1992).
34. Van Aalten, D. M. *et al.* A program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules. *J. Comput. Aided Mol. Des.* **10**(3), 255–262. <https://doi.org/10.1007/BF00355047> (1996).
35. Kabsch, W. & Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**(12), 2577–2637. <https://doi.org/10.1002/bip.360221211> (1983).
36. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**(1), 33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5) (1996).
37. Skjærven, L., Yao, X. Q., Scarabelli, G. & Grant, B. J. Integrating protein structural dynamics and evolutionary analysis with Bio3D. *BMC Bioinform.* **15**(1), 399. <https://doi.org/10.1186/s12859-014-0399-6> (2014).
38. Yuan, S., Chan, H. S. & Hu, Z. Using PyMOL as a platform for computational drug design. *Comput. Mol. Sci.* **7**(2), e1298. <https://doi.org/10.1002/wcms.1298> (2017).
39. Masood, T. B., Sandhya, S., Chandra, N. & Natarajan, V. CHEXVIS: A tool for molecular channel extraction and visualization. *BMC* **16**, 119. <https://doi.org/10.1186/s12859-015-0545-9> (2015).
40. Chovancova, E. *et al.* CAVER 3.0: A tool for the analysis of transport pathways in dynamic protein structures. *PLoS Comput. Biol.* **8**(10), e1002708. <https://doi.org/10.1371/journal.pcbi.1002708> (2012).
41. Mezei, M. & Filizola, M. TRAJELIX: A computational tool for the geometric characterization of protein helices during molecular dynamics simulations. *J. Comput. Aided Mol. Des.* **20**(2), 97–107. <https://doi.org/10.1007/s10822-006-9039-1> (2006).
42. Mezei, M. Simulaid: A simulation facilitator and analysis program. *J. Comput. Chem.* **31**(14), 2658–2668. <https://doi.org/10.1002/jcc.21551> (2010).
43. Dahl, A. C., Chavent, M. & Sansom, M. S. Bendix: Intuitive helix geometry analysis and abstraction. *Bioinformatics* **28**(16), 2193–2194. <https://doi.org/10.1093/bioinformatics/bts357> (2012).

### Author contributions

J.A. conducted the MD simulations, performed comparative analysis and wrote the manuscript. All authors analysed the simulation results. C.E., R.E. and A.I. supervised the research.

### Competing interests


The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to J.A.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021

**CHAPITRE III: *In silico* Method for Design and Discovery a New Compound as Potential of the Fat-taste Analogues.**

### III.1. Résumé de l'article

#### **Développement d'une méthode d'analyse *in silico* du mécanisme de signalisation des acides gras via le récepteur CD36 et la conception de nouveaux analogues phytochimiques naturels du goût gras.**

L'obésité reste l'un des enjeux majeurs de santé publique du XXI<sup>e</sup> siècle, en raison de sa progression rapide dans le monde et des effets délétères qui lui sont généralement associés (diabète de type 2, troubles vasculaires, hypertension, cancer, maladies neurodégénératives) ((Flegal K *et al.*, 2010; (Mokdad AH *et al.*, 1999)). Bien que les origines de cette épidémie soient clairement multifactorielles, le bouleversement récent de nos habitudes alimentaires contribue largement à ce phénomène. L'une des conséquences d'une offre alimentaire abondante est de faciliter la consommation d'"aliments plaisir", généralement riches en sucres et en graisses. En effet, ces aliments s'avèrent très attractifs en raison de la satisfaction hédonique immédiate qu'ils procurent. Cette observation suppose l'existence d'une détection précoce et sélective des caractéristiques physicochimiques de ces aliments dès leur mise en bouche (Zuccotti G *et al.*, 2013). Cette détection orale dépend de l'efficacité d'un système sensoriel. L'identification récente de récepteurs lipidiques (par exemple CD36) au niveau des arrondissements gustatifs chez les rongeurs et chez l'homme, dont l'absence partielle ou totale a détecté les choix alimentaires, suggère que les lipides alimentaires peuvent également être perçus par le goût (Degrace-Passilly P *et al.*, 2012). Nous savons que les personnes obèses ont tendance à préférer les aliments riches en matières grasses. CD36 est un candidat plausible en tant que capteur lipidique gustatif, également connu sous le nom de transporteur d'acides gras. À ce jour, le récepteur cristallographique CD36 est capturé dans la structure à l'état ouvert sans extrémité N-terminale et c-terminale (TM).

Les modificateurs de goût gras sont les composés qui peuvent être synthétisés chimiquement ou purifiés à partir de matériaux biologiques et reproduisent le principe du goût gras en se liant aux récepteurs du goût gras (Shanmugamprema *et al.*, 2020). Très récemment, démontre que l'agoniste TUG891 se lie aux récepteurs linguaux et déclenche son axe langue-cerveau-intestin pour moduler la préférence gustative du gras (Murtaza *et al.*, 2020).

L'objectif de notre étude est la mise en place un protocole pour de déterminer le mécanisme de la régulation allostérique (démérisation) et l'activation de la signalisation CD36 d'une part, et la conception et le criblage statique et dynamique d'analogues de composés phénoliques pour nouvelle approche thérapeutique potentielle contre l'obésité.



Dans ce travail, nous abordons l'activation de CD36 en décrivant d'abord un mécanisme de liaison avec l'acide palmitique à l'aide d'une étude de simulations moléculaire. Ensuite, nous avons réalisé un aperçu des possibles interactions de dimérisation contribuant à la communication allostérique entre le site K164 et les domaines TM1 et TM2. Enfin, nous avons développé 15 analogues de composés phénoliques activant le CD36, car une signalisation alternative via les analogues pourrait éviter les effets secondaires négatifs caractérisant les médicaments actuellement approuvés, tels que l'hypoglycémie et la prise de poids.

Les résultats de ce travail ont montré que l'acide palmitique induit une dimérisation contribuant à la communication allostérique entre le site K164 et les domaines TM1 et TM2. La conception et par la suite le criblage statique de 15 molécules, en utilisant l'amarrage moléculaire, a confirmé un score élevé pour trois analogues (CIN4, COU5, CAF1). L'analogue CIN4 a montré, par la suite, par criblage dynamique, une dimérisation des 2 domaines TM1 et TM2 qui ressemble à l'acide palmitique.

Ce chapitre est présenté sous forme d'un article soumis dans *Journal of chemical information and modeling* (Akachar *et al.*, 2022). Dans cette publication, on a réalisé tout le travail d'analyse bioinformatique et de rédaction du manuscrit.

### III.2. Article original

#### ***In silico* Method for Design and Discovery a New Compound as Potential of the Fat-taste Analogues.**

Jihane Akachar<sup>1</sup>, Catherine Etchebest<sup>2</sup>, El Mehdi Bouricha<sup>1</sup>, Rachid El Jaoudi<sup>1</sup> and Azeddine Ibrahimi<sup>1</sup>

<sup>1</sup> Laboratory of Biotechnology (MedBiotech), Rabat Medical and Pharmacy School, Mohammed V University in Rabat, Rabat, Morocco.

<sup>2</sup> UMR\_S 1134, DSIMB, Université Paris Diderot, Sorbonne Paris Cite, Paris, France,

#### **Abstract**

Current studies have suggested that extreme intake of dietary fat is associated with obesity. Some obese individuals have been reported to exhibit high thresholds for the gustatory detection of lipids via lipid receptors, such as cluster of differentiation 36 (CD36), involve Ca<sup>2+</sup> signaling. Our examination of fatty acid signaling mechanism via the CD36 receptor by molecular modeling and dynamics revealed a new class of phytochemical analogues as “fat taste modifiers” that would bring no calories and could be used to decrease the fat-eating behavior in obese and diabetic-obese patients. This study confirms a high threshold of gustatory fat detection *in silico*.

**Key words:** Obesity, Lipids, Taste, CD36, Molecular dynamics,

#### **Introduction**

Obesity remains one of the major public health challenges of the twenty-first century because of its rapid progression worldwide and the deleterious effects which are generally associated with it (type 2 diabetes, vascular disorders, hypertension, cancer, neurodegenerative) [1,2]. Although the origins of this epidemic are clearly multifactorial, the recent upheaval in our eating habits contributes significantly to this phenomenon. One of the consequences of an abundant food supply is to facilitate the consumption of “pleasure foods”, generally rich in sugars and fats. In fact, these foods turn out to be highly attractive because of the immediate hedonic satisfaction they provide. This observation assumes the existence of early and selective detection of the physicochemical characteristics of these foods as soon as they are put in the mouth [3]. This oral detection depends on the efficiency of a sensory system. The



recent identification of lipid-receptors (eg CD36) at the level of taste boroughs in rodents and in humans, the partial absence of which or total affects food choices, suggests that dietary lipids may also be perceived through taste [4]. We know that obese people tend to prefer foods rich in fat. CD36 is a plausible candidate as the gustatory lipid sensor, which is also known as the fatty acid transporter. The CD36 amino acid sequence exists as a transmembrane glycoprotein with an extracellular pocket structure between the cytoplasmic amino and carboxyl terminal tails [5], allowing the transduction of external signal into the cell [6]. It has a nanomolar-range affinity [7] to a range of lipid-based ligands, such as lipoprotein, apoptotic cells, and LCFAs [8]. CD36 has been detected in human foliate and circumvallate papillae [9]. Furthermore, CD36 gene inactivation abolishes the preference for long chain fatty acids (LCFAs) in rats [10,11] without affecting the sensitivity to sweet or bitter taste [10]. CD36 could be the target of those taste modulators which believed to elicits the taste sensation without caloric value and possibly used against the dietary-induced obesity. The fat taste modifiers are the compounds that can be either synthesized chemically or purified from biological materials and duplicate the principle of fat taste by binding to fat taste receptors [12]. Very recently, demonstrates that the agonist TUG891 binds to the lingual receptors and triggers its tongue-brain-gut axis to modulate the fat taste preference[13].

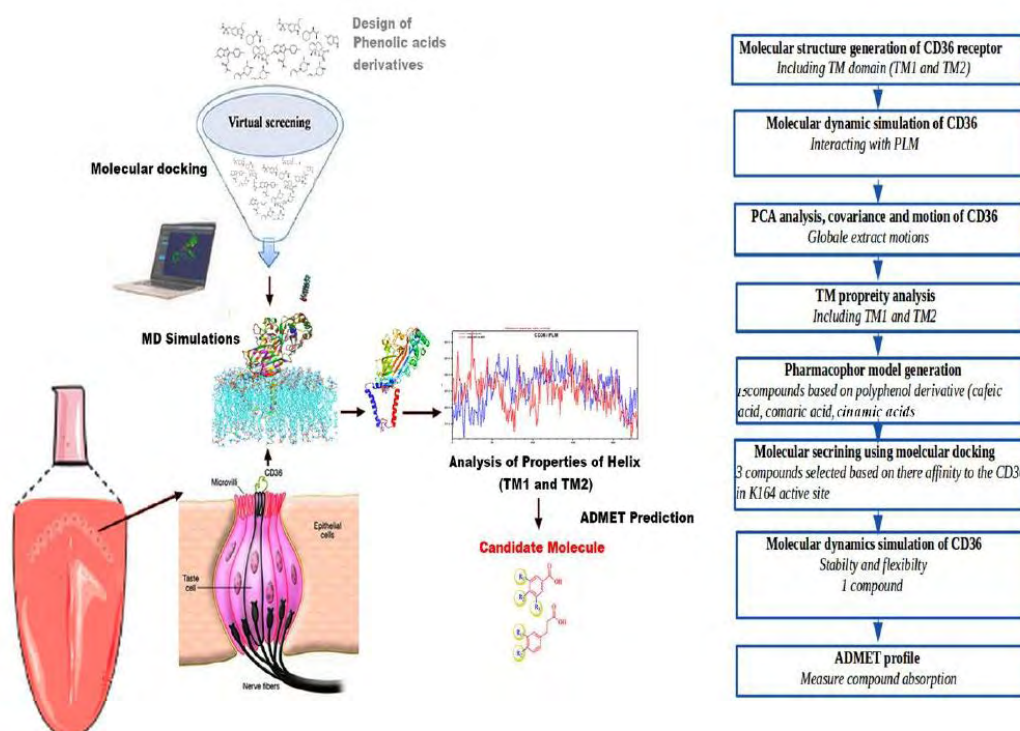
Phenolic acids are types of aromatic acid compounds which are widely distributed in plants and are abundant in fruits, vegetables, and beverages. The great interest in phenolic acids is associated with their high potential for food preservation (antioxidant and antimicrobial activity) and, last but not least, high therapeutic potential [14]. In particular, p-coumaric, caffeic, ferulic, and cinnamic acids have been highlighted due to their relevant biological activities [15,16]. This study explores the developement of new method *in silico* for gustatory mechanisms of fatty acid chemoreception (fat taste) and how these mechanisms may be involved in the report-ed associations between fat taste threshold, weight gain and obesity.

Here we address the CD36 activation first describing a putative binding mechanism for the Fatty acid complex using molecular dynamics (MD) simulations. Then, insights on possible demirisation interactions contributing to the allosteric communication between site K164 and TM1 and TM2 from MD simulations performed on different CD36/PLM complexe. Finally, the development of phenolic compounds analogues activating CD36 represent a new potential therapeutic approach against obesity as the activation of its alternative signaling may avoid the negative side effects characterizing the currently approved drugs, such as hypoglycemia and weight gain. To date, the structure-based design of new compounds can

exploit four crystallographic CD36 complexes. All these structures the receptor is captured in the open state. Overall, this study produces mechanistic working hypothesis on the allosteric regulation and activation of CD36.

### Computational methods

For designing new class of molecules as ‘fat taste modifiers’ 15 phytochemical analogues (cinnamic acid, p-coumaric acid, and caffeic acid) via lipid receptor CD36, the method adopted in this computational study is illustrated in [Figure 1](#). PLM which was used as a control in molecular modeling, docking and molecular dynamics simulations to predict the binding modes and estimate their binding affinity with and lipid receptor CD36.



**Figure 1.** Outline of computational method adopted in this study.

## **Study of Molecular Mechanisms of Fat Taste Perception**

### **Molecular Modeling**

The tridimensional structures of N-terminal region TM1 (from 1 to 160) and C-terminal TM2 (400 to 441) transmembran domain of CD36 was predicted by using Modeld softwer to construct the homology model based on the crystal structure CD36-CIDR $\alpha$  (PDB ID: 5LGD) which is obtained from the Protein Data Bank, where all ligands and water molecules in the protein crystal structure were deleted (17,18). The entire sequence of human CD36 was taken from the UniProt Database ([http:// www. unipr ot. org](http://www.uniprot.org)) with accession number P16671. The quality of the homology model was evaluated by Ramachandran plot (19), and its overall quality factor ERRAT2 (20) plot and PROVE produced by the SAVS server v6.0 ([https:// saves. mbi. ucla. Edu/](https://saves.mbi.ucla.edu/)) and (<https://zlab.umassmed.edu>). The structure of phenolic compounds cinnamic acid, p-coumaric acid, and caffeic acid which is obtained from the PubChem databases (21) and the software MarvinSketch (Marvinsketch, 2017) used to edit and draw molecules (22).

### **Molecular docking study of PLM with CD36 receptor**

To position the fatty acid (PLM) in the binding site (K164) (23) of the CD36 outward facing conformer, molecular docking was performed with the Autodock vina 4.2.6 software (24). The input files for docking calculations were obtained using mgltools 1.5.6 software. Gasteiger partial charges were assigned for all the atoms of the system by default. Conformational search was performed using the Lamarckian genetic algorithm (LGA) with a population size of 300 randomly placed individuals and maximum number of  $2.5 \cdot 10^7$  energy evaluations.

### **Study of Molecular Mechanisms of Fat Taste Perception by Molecular Dynamics Simulation**

Lipid receptor CD36 structure was inserted into palmitoyl-oleoyl phosphatidyl choline (POPC) bilayer using the CHARMM-GUI tool (25). The resulting system was explicitly solvated using the TIP3P water model (26) and neutralized by the addition of counterions of K<sup>+</sup> and Cl<sup>-</sup> ions at concentration of 0.15 M. MD simulations leaded out for 100 ns using the GROMACS program (27), version 4.5.5, using the CHARMM36 (28) force field, were simulated under fully periodic boundary conditions. All electrostatic reactions were processed including Par-ticle Mesh Ewald (PME). So every bonds between hydrogen molecules was



recondition utilizing the LINCS algorithm<sup>32</sup>. Water organization was constrained using rigid SETTLE algorithm (29) and non-bonded reactions, a cut-based is used, which allowed us to use an integration step of 2 fs. The ligand topology files, as needed in GROMACS, were generated from the CGenFF server (The official CHARMM General Force Field server) (30) which was an automated server for topology generation (<https://cgenff.umaryland.edu/>). The complex was solubilized using SPC in water and balanced by the raise of K<sup>+</sup> and Cl<sup>-</sup> ions in a cubic and dodecahedron box. In the simulations MD, the protein was simulated in the presence of a single palmitic acid, located near its surface. In MD simulation of each complex was performed in five consecutive stages: (1) energy minimization, (2) NVT heating, (3) NPT equilibration, (4) NPT pre-production simulation, and (5) production simulation. Besides, the time step for the simulations was 2 fs. All MD calculations for period 100 ns. The RMSD, RMSF, Rg, and SASA calculated using GROMACS module. Furthermore, visual depiction was taken out employing PyMOL and VMD (31,32).

#### **Principal component analysis PCA, Covariance Analysis and Motions of CD36/PLM complexe receptor**

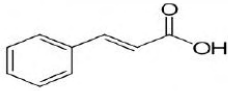
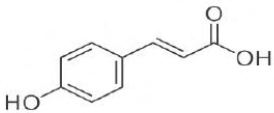
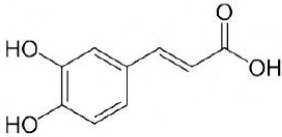
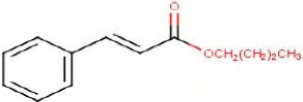
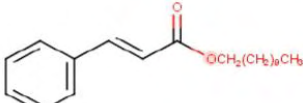
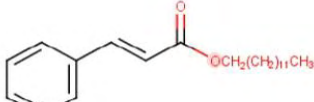
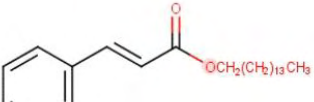
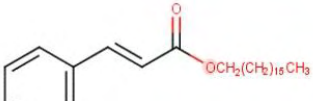
The principal component analysis (PCA) and cross-correlation were performed using the bio3D package in R (33). Once a sufficient MD production runs had been completed, the most important component analysis is used to analyze the results from the MD simulations. PCA became an efficient and an excellent tool for the global extract motions and the correlated structural dynamics across many biological systems.

#### **Analysis of Properties of Transmembranaire TM1 and TM2 Helix of CD36 receptor with PLM**

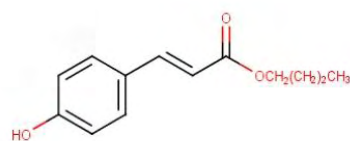
To analyzes the bending and rotational movements of helices TM1 and TM2 of CD36 in static structures and dynamic trajectories was performed with help of of the TRAJELIX tool of the Simulaid package (<https://mezeim01.u.hpc.mssm.edu/simulaid/>) Earlier ver-sions of Simulaid (IRIS GL GRAPHICS VERSION) (34) and the VMD Bendix Plugin, Version 1.1 43.

### **Vertuel secrening of potontial ligands (fat taste modifiers) molecular docking**

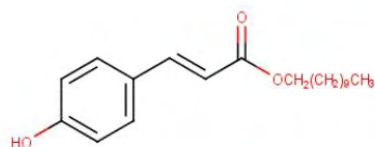
The structure of phenolic compounds cinnamic acid, p-coumaric acid, and caffeic acid which is obtained from the PubChem databases (21) and the software MarvinSketch (Marvinsketch, 2017) (22) (Table 1) used to edit and draw molecules in 2D. Were minimized and optimized by using AutoDock Tool (ADT). To position the all new phenols analogues in the binding site (K164) (23) of the CD36 outward facing conformer, molecular docking was performed with the Autodock vina 4.2.6 software (24). The input files for docking calculations were obtained using mglttools 1.5.6

Compound	Structure
<b>Cinnamic acid</b>	
<b>Coumaric acid</b>	
<b>Caffeic acid</b>	
<b>CIN1: C4</b>	
<b>CIN2: C11</b>	
<b>CIN3: C13</b>	
<b>CIN4: C15</b>	
<b>CIN5: C17</b>	

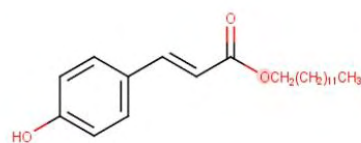
**COU1: C4**



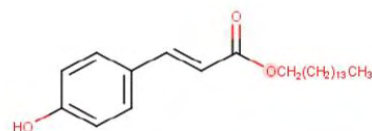
**COU2: C11**



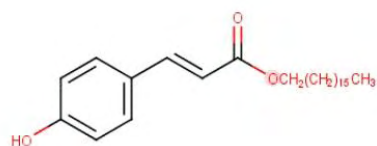
**COU3: C13**



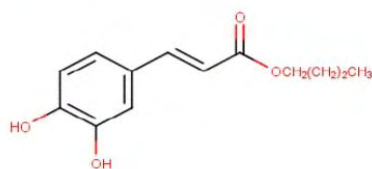
**COU4: C15**



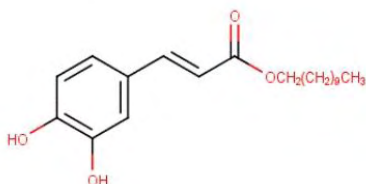
**COU5: C17**

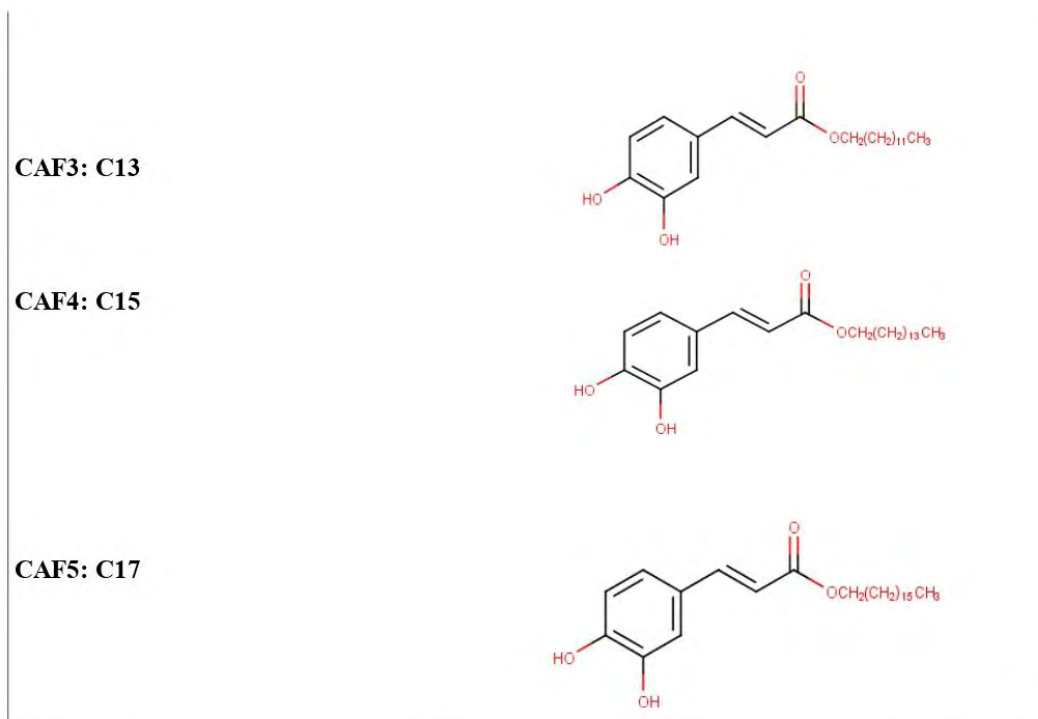


**CAF1: C4**



**CAF2: C11**





**Table 1.** Chemical structures substituent for the design new potential ligands derived from phenolic acids (cinamic acids, coumaric acid, caffeic acid) used in this study as analogues fat taste.

Software. The docking scores were resulted in the generated .log files. Gasteiger partial charges were assigned for all the atoms of the system by default. Conformational search was performed using the Lamarckian genetic algorithm (LGA) with a population size of 300 randomly placed individuals and maximum number of  $2.5 \cdot 10^7$  energy evaluations. The output docking scores were defined as affinity binding (Kcal/mol). The ligands protein interactions were created by using PyMOL (version 1.6) and VMD. (31,32).

### **Study of Molecular Mechanisms of new fat taste analogues by Molecular Dynamics**

#### **Simulation**

The most realistic scenario of protein-ligand interactions cannot be replicated by docking experiments. To better understand the stability of the complexes between new ligands designs and S CD36 receptors, molecular dynamics simulations (MDS) were performed for 5000 ps using the desmond module of Schrödinger software 2017-4 (35). The OPLS3e force field parameters were used in all simulations. A three-step process was followed which included system builder, energy minimization and molecular dynamics simulation. The structures of



all new ligands docking complexes with CD36 were solvated in an orthorhombic boundary box using system builder by choosing the TIP3P as water model. The system charge was neutralized by adding Sodium (Na<sup>+</sup>) and chloride (Cl<sup>-</sup>) ions. The SHAKE algorithm was adopted for bond geometry constraints, and the cutoff radius parameter of 10 Å was chosen in Coulomb interactions. The electrostatic interactions were produced by on the particle mesh Ewald method (PME). The prepared systems in each run were subjected to energy minimization and equilibrated via an NPT ensemble using Berendsen thermostat at a temperature of 300K and 1 bar pressure. Finally, a total of 2000 ps of production MD was performed for each complex and the trajectories were saved every 4.8ps. The stability of the complexes throughout the simulations was assessed by the root mean square deviation (RMSD) and the root mean square fluctuation (RMSF) metrics implemented in the Desmond MD package.

#### **Analysis of Properties of Transmembranaire TM1 and TM2 Helix of CD36 receptor with new fat taste analogues**

To analyzes the bending and rotational movements of helices TM1 and TM2 of CD36 in static structures and dynamic trajectories was performed with help of VMD Bendix Plugin, Version 1.1 (36).

#### **Analysis of the physicochemical properties of the candidate fat taste modifiers**

The physicochemical properties of new fat taste analogues was assessed through the prediction of adsorption, distribution, metabolism and excretion (ADME) properties using the online server preADMET (<http://preadmet.bmdrc.org/>).<sup>26</sup> In this study we chose the properties human intestinal absorption (HIA), in vitro Caco-2 (colon adeno carcinoma) cell permeability, in vitro MDCK (Madin–Darby canine kidney) cell permeability, in vitro plasma protein binding and in vivo BBB (blood–brain barrier) penetration to be predicted in this program.<sup>(37)</sup>

## Results and Discussion

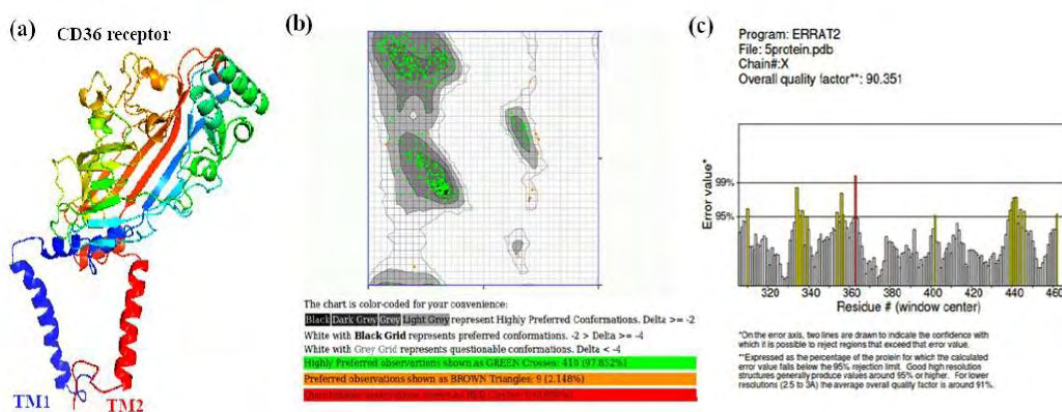
Computational methods can be used as a reply to many types of issues. Here we used molecular modeling, molecular docking and MD simulations to determine the mechanism of Fat Taste Perception via lipid receptor CD36 and desing a new candidate fat taste modifiers.

Figure 1 illustrates an overview of the current method.

### Structural studies of Mechanisms of Fat Taste Perception

#### Molecular Moldeing CD36 complet TM domain (TM1 and TM2)

The three dimensional (3D) structure of CD36 including TM domain (TM1 and TM2) predicted by using Modeled softwer to construct the homology model based on the crystal structure CD36-CIDR $\alpha$  (PDB ID: 5LGD). The quality of the homology model was evaluated by Ramachandran plot, the overall quality factor ERRAT2 plot produced by the SAVS server v6.0 (<https://saves.mbi.ucla.edu/>). The stereo chemical properties of the modeled demonstrates that a majority of its residues are in favored region (97.85%) and allowed region (2.14%) of the Ramachandran plot, that indicates reasonable quality of this homology model. Moreover, the ERRAT2 score was 90.35. We observed that the currently modeled structure have structural qualities that are acceptable 24 (Figure 2).



**Figure 2** : Molecular Moldeing CD36 complet : (a) 3D Cartoon presentation of the CD36 with TMs (TM1 and TM2) structure, (b) Ramachandran plot calculations for CD36 protein. (c) The 3D profile verified results of the CD36 model using ERRAT program

### **Molecular docking study of PLM with CD36 receptor**

Docking analysis was dedicated to predict the binding modes and affinity of PLM on the K164 CD36 fatty acid binding domain (Table 2). The docking scores for PLM were  $-3.00$  kcal/mol. In addition, we studied the binding interactions of PLM found against CD36 to identify their activation mechanism. The surface of the binding site of CD36 protein is mostly surrounded by Arg63, Phe65, Leu140, Ala143, Ala144, Ile148, Leu161, Ser162, Lys164, Leu187, Leu188, Leu189, Pro191, Tyr192, Pro193, Val194, Thr195, Val198, Thr197, Gly199, Leu200, Phe201, Pro203, Tyr204, Thr207, Ser268, Ser269, Asp270, Ile271, Arg273, Ile330, Lys334, Glu335, Gly336, Arg337, Pro338, Ser339, Val339, Lys385, Leu387, Val389, Leu343, Ile341, Thr369, Leu391, Leu416, Glu418 residues (Akachar et al, 2021). These findings prompted us to perform MD simulations of the docked protein–ligand complexes to elucidate the dynamic behavior of PLM against CD36.

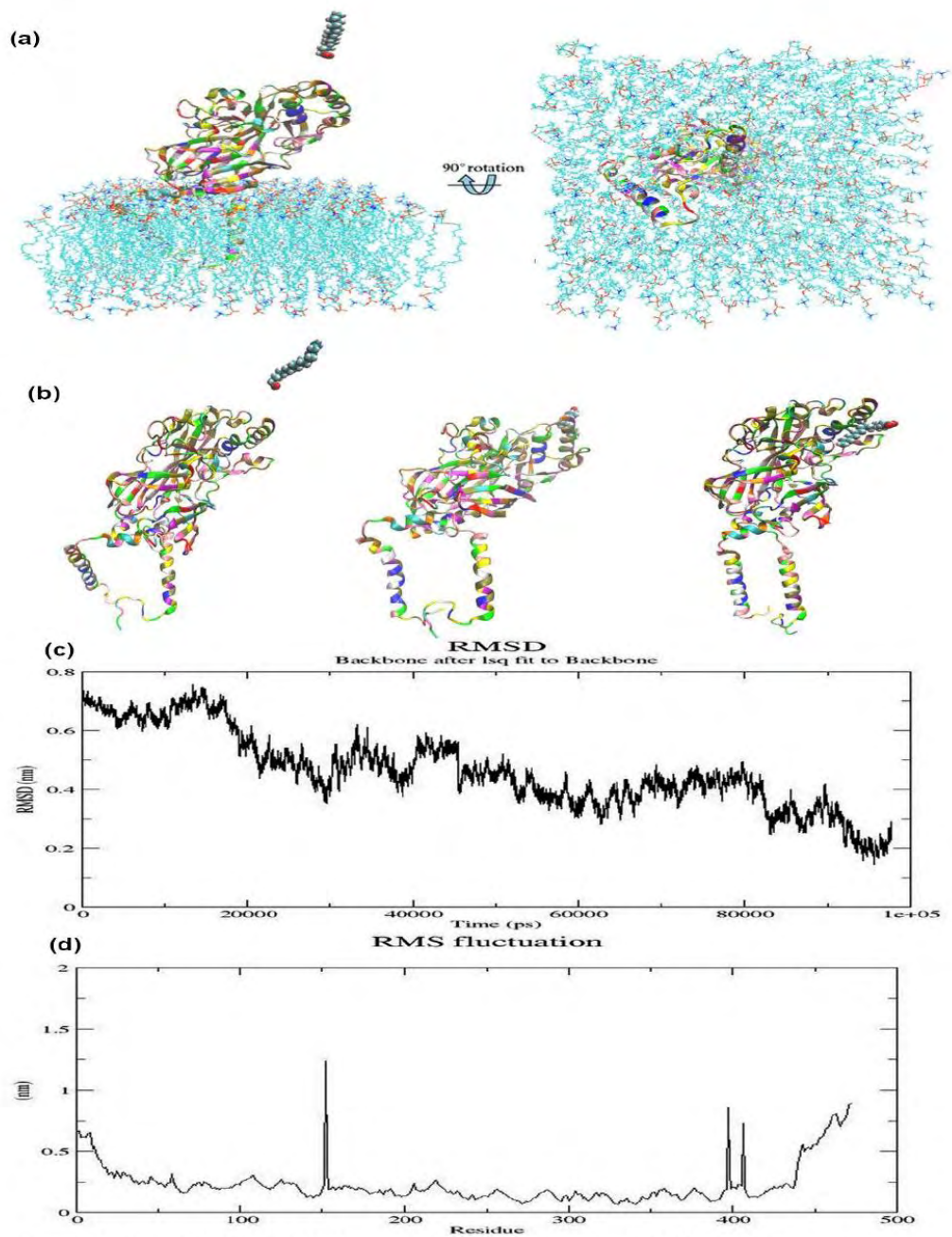
### **Global molecular dynamics of the simulations analysis of PLM with CD36 receptor**

Because protein flexibility represents a significant function in CD36 transport and signaling, we conducted MD simulations of 100 ns in presence of PLM. Also to study the link of TM1 and TM2 domain with PLM binding and signaling (Figure 3,a). After 100 ps the structure of CD36 with PLM, revealed that PLM is located above Lys164 and near the upper helix (Figure 3,b). This is a pattern of the transportation channel of CD36 receptor 25. CHO cells expressing mutated CD36 Lys-164 showed alterations in FA uptake and FA-induced calcium liberation indicating the importance of Lys-164 for both FA effects 20. We calculated and evaluated the three simulation systems: the structure stability of the proteins in the time of the MD and RMSD/backbone, which were considered relative to the first conformation, as we had observed in the RMSD representation (Figure 3,c). The plots of RMSD reflected equilibrium of the simulations that were signal of the global protein constancy. The peak RMSDs were at 0.70 nm for the simulation systems. The graph corresponded to CD36 in presence of PLM were illustrated with black bars. The CD36 in presence of PLM attained early stability due to convergence at 0.50 from 50 ns. The RMSF of C-alpha atoms was used to infer with the residue specific flexibility in order to study the fluctuation of every single residue. In other words, during MD simulation three RMSF plots for CD36 in presence of PLM were generated (Figure 3d).

For the black curve, a trimodal distribution at the regions close to the individual portal was observed for the complex CD36-PLM. The elevation of RMSF at the upper helix domain

and coil 1.40 nm and N-terminal region TM1 (from 1 to 160) and C-terminal TM2 (400 to 441) 0.6 , 0.9 nm respectively, was supposed to be produced by the amino acid lys 164 impacts of first assays to enter from these portals, which surround the bottom of the cavity. studies studies done in absence of





**Figure 3.** Study of Molecular Mechanisms of Fat Taste Perception by Molecular Dynamics Simulation. **(a)** : Orientation of CD36 proteins on membrane obtained with PPM modeling. **(b)** : Interactions of CD36 protein with PLM on results of molecular dynamics over simulation time. **(c, d)** : Analysis of all backbone atom root mean square deviations (RMSD) and Root mean square fluctuations (RMSF) of PLM/CD36 complex over 100 ns simulation time.

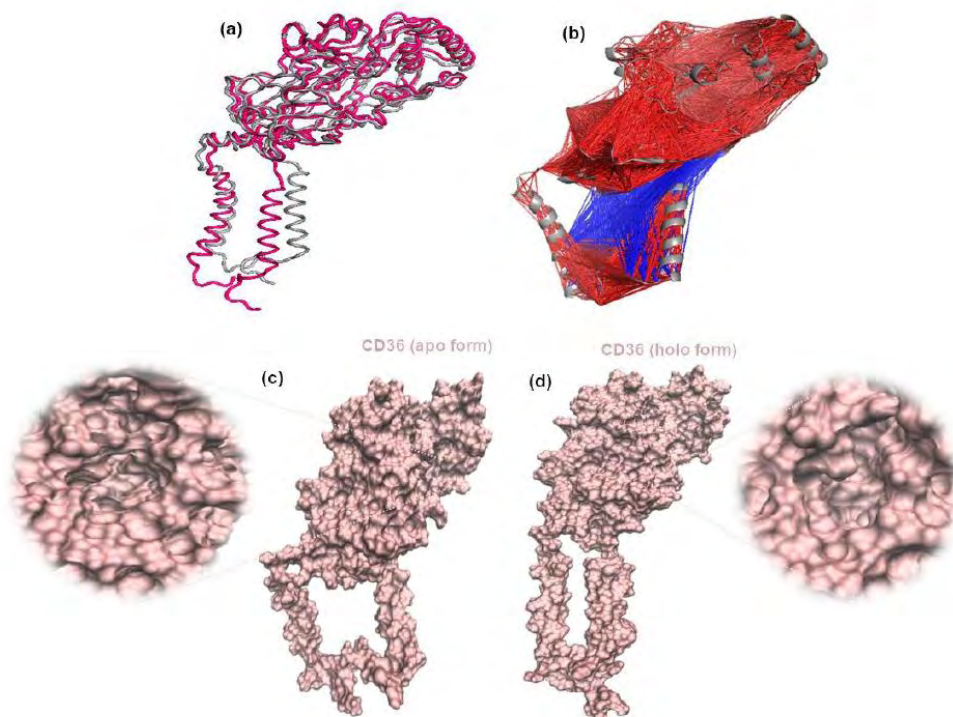
TM domain in simulation in presence PLM did not displayed a lot of displacement at the binding region Lys164 0.6 nm only (jihane ref). The difference in fluctuation patterns CD36 could be attributed to their subtle importance of Lys164 residues and structural features. The flexibility of the upper part of the helix/loop and TMs domain TM1 and TM2 is important for the Fatty acid signalisation. Clearly, the access of PLM from this portal requires protein conformational change in upper helix from Lys164, which might explain why it was the most preferred route in our multiple independent runs. To conclude the RMSF plot confirmed clearly that in the helix regions of the upper loop and TMs domain the fluctuations of the residues were more articulated.

### **Principal component analysis PCA, Covariance Analysis and Motions**

We use the Principal component analysis (PCA) in order to identify the motions and the direction of the flexibilities in the residues. The principal components analysis was applied to the backbone atoms for CD36-PLM complex (Figure 4a). The first outstanding featured motions during the simulation were analyzed through secondary structures which includes an opening-closing motion of transport tunnel. Our observation focused on the movements of the C-coil, A-helix and TM (TM1 and TM2) domain (Figure 4a). We found, that the flexible motions of TM domain structure in the distinct active states in CD36-PLM complex are quite dissimilar at the start and end of the simulation. The top principal components for model PC1 for CD36 in presence of PLM are illustrated in (Figure 4a). In this study we are interested to analyze the movement of the TM domain (TM1 and TM2) the most prominent motions observed are propeller TM2 of CD36 which are the most flexible in moving in the opposite direction. Particularly, in holo structure the pink color (Figure 4a) TM2 moves leftwards, and rightwards for moves in the opposite direction for apo form color grey. It is evident that this may suggest that the site K164 is able to settle the paradigm by activate the movement of the TM domain to activate cellular signalisation. For a preferable comprehension of the association through the movement of the TM domain, coil and Lys164 residue, as well as the movements of the adjoining part, covariance analyses were executed during simulation on the CD36 in presence of PLM. They were illustrated utilizing 3D system plots. As shown in Figure 4b, positive correlations are mapped in light blue green color while negative correlations are mapped in pink color. The deeper color indicates more correlations (or anti-correlations). We detected a cluster with correlated motions in the middle the secondary structures and the adjoining zone in the active situation of the related

groups. The correlated motions of these parts grow extremely in the TM2 (Figure 4b), The correlation and anticorrelation resented in red and blue color line. However, the K164A increased the anticorrelated motions between the upper helix and C-terminal and N terminal TM domain. In addition, strongly correlated motions were also observed within the beta sheets regions and helix C-terminal and N-terminal regions, which signal that the conjunction of these parts retired and proposed that the flexibility of the helix coil is in charge for the movements of the adjoining pathways. Besides, the movements of these two secondary structures are correlated with the movement of upper helix that is essential for the structural stability of CD36 structure for PLM transport and signaling. In the following parts, Figure 4c,d and d show the structure of CD36 on the surface to clearly see the difference between activation and inactivation of the CD36 receptor at the TM. The TM1 and TM2 domains in the holo conformations of CD36 are closer than the apo form. We have observed that the TM domain is getting closer, so we conclude that a signaling has been triggered.





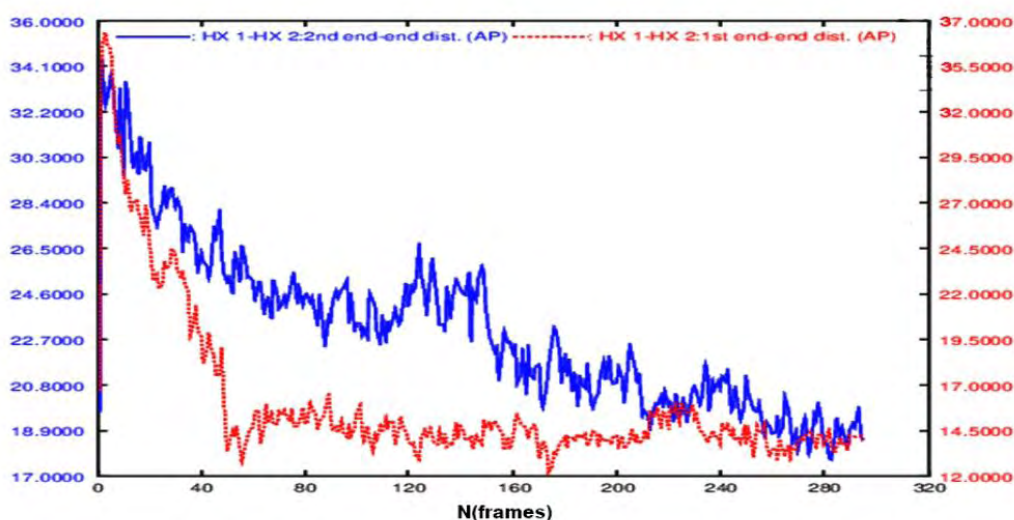
**Figure 4:** Model and motion produced by principal component analysis for CD36 in presence of PLM a function of time during 100,000 ps (100 ns). (a) Principal component analysis of CD36 and Visualization of the motions along PC1, the direction and amplitudes of the movements occurred from first trajectory of holo and apo form in (pink) and (gray) color respectively. (b) : Cross-correlation Analysis matrix of the fluctuations of the C  $\alpha$  atoms of the residue from their average during 100 ns of CD36 in presence of PLM. Residue cross-correlations. CD36 is depicted by gray cartoon. Red and blue lines indicate correlated and anti-correlation motions (c, d) Surface representation of CD36, The structure is represented in two different configurations, at the beginning (c) and at the end (d) of the simulation.

#### **Analysis of properties of helix Analysis TM1 and TM2 for PLM/CD36 complex**

Last study demonstrate the associated motion between the Lys164 residue and C-coil and the TMs domain. Analysis of changes in the global and local structural properties of transmembrane helices that may occur during MD simulations at the interface of dimerization may help understand the nature of the interaction, and its implication in receptor function. Evaluation of this contribution can be particularly important to understand mechanisms of signal transduction, with the ultimate goal of helping structure-based molecule bioactif



design. To enable quantitative assessments of possible changes in dynamic molecular models of CD36 we are currently carrying out extensive MD simulations of inactive and activated molecular models of in explicit lipid-water environment and presene of PLM. The results of these calculations, which will be reported elsewhere, are expected to provide new insights into the molecular determinants responsible for CD36 activation. Which suggests the organization of CD36 TM domain into two dimensional arrays of dimers with TM helices 1 and 2 involved in intradimeric contact. TRAJELIX used to reveal the nature of possible intermolecular changes at the interface of interacting alpha-helical membrane proteins during MD simulations. We carried out a relatively short MD simulation of the proposed TM1,2-TM2,1 dimeric model of CD36 with PLM using a new energy function for the implicit representation of the membrane, and analyzed the resulting trajectory of 100 ns with TRAJELIX program (Figure 5).



**Figure 5.** Analyse of the CD36 TMs dimerization, resulting of 100 ns with TRAJELIX program

An evaluation of the helix compression/extension during the MD simulation is provided in Figure III.6 (red and blue lines). Specifically, Figure III.6 shows that TM 1 and 2 change during the MD simulation, and they seem to undergo significant compression or extension values greater than 1 angstrom exactly 8 angstroms for TM1 and 1 angstrom for TM2 and show the total displacement TM1 towards the TM2 direction, and the displacement in the x, y and z directions of the TM helices in CD36. These movements are maximal for these helices during the 100 ns simulation. Finally, however, these suggested changes in the global and

local structural properties of CD36 TMs helices therefore result in a dimerization of two helices for signal transduction.

#### **Vertuel screening of potontial ligands molecular docking**

Most of new fat taste analogues showed good binding affinity against fatty acid binding site K164 of CD36 receptor (Table 2). The CIN4, COU5 and CAF1 showed the most potent binding affinity against CD36 than palmitique acid (PLM). The docking scores for new fat taste analogues pheare – 3.50 kcal/mol, and – 3.40 kcal/mol and – 3.60 Kcal/mol respectively and – 3.00 Kcal/mol for PLM. On the other side, CAF1 cafeic acid analogue acid are phenolic acid compounds and they showed favourably docking scores. Consequently, these docking scores exposed that most new phenolic analogues found high binding affinity against K164 better than PLM, and thus they would likely bind and activate the CD36 signaling. These findings prompted us to perform MD simulations of the docked protein–ligand complexes to elucidate the dynamic behavior of new fat taste analogues against CD36.

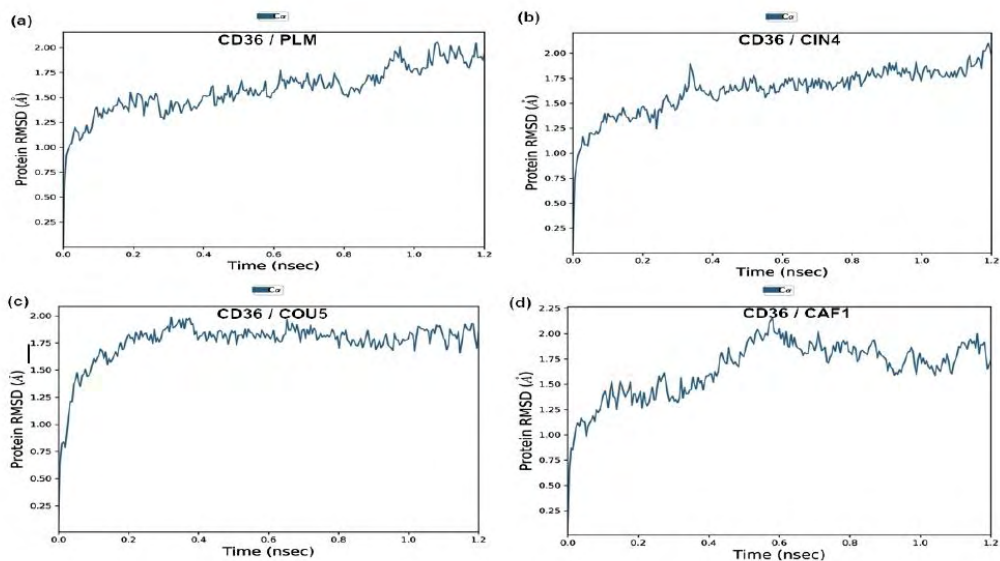
Compound	$\Delta G$ (kcal mol <sup>-1</sup> )
CIN1: C4	-3.4
CIN2: C11	-3.4
CIN3: C13	-3.4
CIN4: C15	-3.5
CIN5: C17	-2.5
COU1: C4	-3.1
COU2: C11	-3.3
COU3: C13	-2.4
COU4: C15	-3.0
COU5: C17	-3.4
CAF1: C4	-3.6
CAF2: C11	-3.5
CAF3: C13	-3.5
CAF4: C15	-3.2
CAF5: C17	-3.5
PLM	-3.0

**Table 2.** The binding affinity and dissociation constant of complexe CD36/new phytochimecal structures of cinnamic, caffeic and coumaric acid ester derivatives and palmitic acid CD36/(PLM) as control.

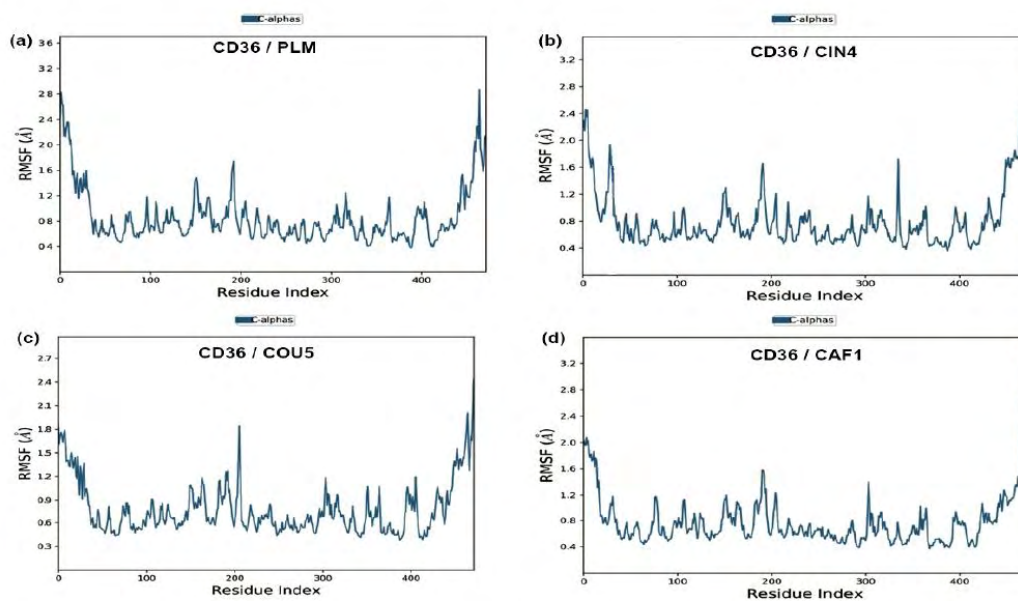
### **Study of Molecular Mechanisms of new ligand Fat Taste Perception by Molecular Dynamics Simulation**

Root mean square deviations RMSD is commonly used to determine the structural stability of all C  $\alpha$  -atoms of a system. Equilibrium vibration and low fluctuation of RMSD were noticed in the PLM/CD36 complex that than rose from PLM (1.20-1.75 Å), showing that the previous complex was greater stable at < 3 Å and induced smaller conformational changes longer time scales (Figure 6a), compared to other new compounds complex that than rose from CIN4 (1.20-1.75 Å), COU5 (1.50-1.75 Å) and CAF1 (1.20-1.75 Å) (Figure 6bcd).

In addition, RMSF or root mean square fluctuation was also calculated and a two-dimensional plot was generated (Figure 7). RMSF is important to define the flexible areas as of a system. Overall RMSF shows values were TMs and K164 that than rose (TM1 2.8, TM2 2.8, K164 1.6 Å) for PLM/CD36 complex structures (Figure 7a) compared to other new compounds/CD36 complex, that than rose (TM1 2.4, TM2 2.4, K164 1.4 Å) for CIN4/CD36 complex, (TM1 1.8, TM2 2.1, K164 1.2 Å) for COU5/CD36 complex, and (TM1 2.0, TM2 2.8, K164 1.2 Å) for CAF1/CD36 complex (Figure 7bcd). The structures indicating all compounds is stable overall, but fluctuate highly in TMs and K164, as. It is obviously appeared that CIN4 is near to the protein binding site and remain at the site during the MD, as compared to other compounds. This is also in accordance with the RMSD calculation, where little change deviations was observed for CD36 in complex with CIN4.



**Figure 6.** Analysis of all backbone atom root mean square deviations (RMSD) of all new drug designs (CIN4, COU5, CAF1 and PLM as control, simulation time with CD36 domain over 1.2 ns. (a): molecule PLM, (b): molecule COU5, (c): molecule CAF1 (d): molecule CIN4.



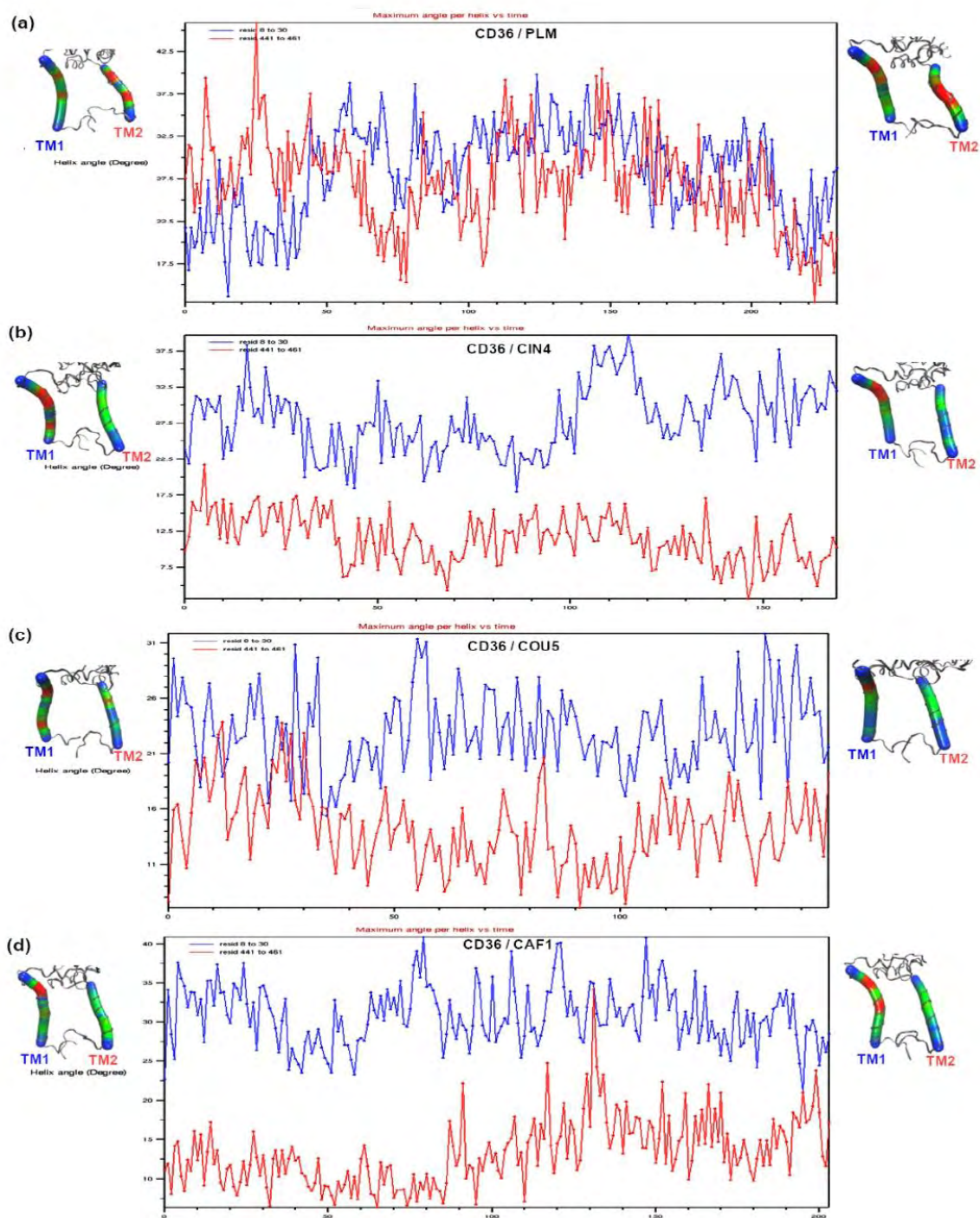
**Figure 7.** Root mean square fluctuations (RMSF) of the new molecules designs (CIN4, COU5, CAF1 and PLM as control) calculated for every residue with CD36 over 1.2 ns. (a): molecule PLM, (b): molecule COU5, (c): molecule CAF1 (d): molecule CIN4.



### **Analysis of properties of helix Analysis TM1 and TM2 for new fat taste anaogues/CD36 complex**

To study CD36 dimerization of the TM domain (TM1 and TM2) helix conformation of CD36 proteins over time, we concentrated on the structure analysis of upper helix using Bendix analyses it analyzed static peptide's local angles along the length of the helix. MultiPlot graph shows how the helix' maximum curvature curvature, per frame, changes over time showed in [Figure 8](#). This graph shows that the helix undergoes significant straightening, as the maximum angle per frame is reduced from a maximum degrees to minimum degrees.

The result of the analyses of static peptide's local angles along the length of TM1 and TM2 the helical angles were measured that than rose from TM1 (17.5-37.5) and TM2 (17.5-42) curvature degree for PLM ([Figure 8a](#)) in comparison with new compounds that than rose from TM1 (20-37.5), TM2 (5-17.5) for CIN4, TM1 (17-31), TM2 (2-22) for COU5 and TM1 (25-40), TM2 (2-25) ([Figure 8bcd](#)). And the helical angles were measured with 20 curvature degree. On the whole, our investigation of upper helix proposes that the upper helix at of CD36 receptor tends to be wrapped, that influences the movement of the helix, loops, and TMs of opening fatty acid transport site signalisation as revealed by the covariance analysis.



**Figure 8.** (a-d) Analysis of Properties of Helix angle (TM1 and TM2) CD36 during MD simulations in presence of new analogues CIN4, COU5 and CAF, PLM as a control.



### Analysis of the physicochemical properties of the candidate fat taste analogues

Pharmacokinetic studies and ADME properties are important conditions and major factors for the evaluation of bioavailability of compounds. Whatever compound is designed/isolated that is going to be a drug, it should satisfy the ADME properties and only then will it be approved as drug in clinical trials. To study the physicochemical properties the of new, fat taste analogues the online server preADMET (<http://preadmet.bmdrc.org/>). program was used. All the predicted ADME values are summarized in the [Figure 9](#).

PLM		CIN4		COU5		CAF1	
ID	Value	ID	Value	ID	Value	ID	Value
BBB	8.21885	BBB	18.0356	BBB	15.3812	BBB	0.794036
Buffer_solubility_mg_L	136.24	Buffer_solubility_mg_L	0.452101	Buffer_solubility_mg_L	0.141014	Buffer_solubility_mg_L	2303.15
Caco2	26.0735	Caco2	56.5975	Caco2	55.141	Caco2	21.86
CYP_2C19_inhibition	Inhibitor	CYP_2C19_inhibition	Inhibitor	CYP_2C19_inhibition	Inhibitor	CYP_2C19_inhibition	Inhibitor
CYP_2C9_inhibition	Inhibitor	CYP_2C9_inhibition	Inhibitor	CYP_2C9_inhibition	Inhibitor	CYP_2C9_inhibition	Inhibitor
CYP_2D6_inhibition	Non	CYP_2D6_inhibition	Non	CYP_2D6_inhibition	Non	CYP_2D6_inhibition	Non
CYP_2D6_substrate	Non	CYP_2D6_substrate	Non	CYP_2D6_substrate	Non	CYP_2D6_substrate	Non
CYP_3A4_inhibition	Inhibitor	CYP_3A4_inhibition	Inhibitor	CYP_3A4_inhibition	Inhibitor	CYP_3A4_inhibition	Inhibitor
CYP_3A4_substrate	Non	CYP_3A4_substrate	Non	CYP_3A4_substrate	Non	CYP_3A4_substrate	Non
HIA	98.297110	HIA	100.000000	HIA	96.036475	HIA	90.653936
MDCK	70.5083*	MDCK	67.2924*	MDCK	66.5858*	MDCK	9.89717
Pgp_inhibition	Inhibitor	Pgp_inhibition	Inhibitor	Pgp_inhibition	Inhibitor	Pgp_inhibition	Non
Plasma_Protein_Binding	100.000000	Plasma_Protein_Binding	100.000000	Plasma_Protein_Binding	100.000000	Plasma_Protein_Binding	84.461317
Pure_water_solubility_mg_L	1.70214	Pure_water_solubility_mg_L	0.00859510	Pure_water_solubility_mg_L	0.00442386	Pure_water_solubility_mg_L	203.724
Skin_Permability	-0.581535	Skin_Permability	-0.53228	Skin_Permability	-0.564253*	Skin_Permability	-1.63399
SKlogD_value	5.269190	SKlogD_value	8.517890	SKlogD_value	9.002000	SKlogD_value	3.198240
SKlogP_value	6.517190	SKlogP_value	8.517890	SKlogP_value	9.002000	SKlogP_value	3.198240
SKlogS_buffer	-3.274660	SKlogS_buffer	-5.899330	SKlogS_buffer	-6.455630	SKlogS_buffer	-2.011080
SKlogS_pure	-5.177970	SKlogS_pure	-7.620310	SKlogS_pure	-7.959090	SKlogS_pure	-3.064360

PLM		CIN4		COU5		CAF1	
ID	Value	ID	Value	ID	Value	ID	Value
algae_at	0.00256135	algae_at	0.0013679	algae_at	0.000610718	algae_at	0.0285745
Ames_test	mutagen	Ames_test	non-mutagen	Ames_test	non-mutagen	Ames_test	mutagen
Carcino_Mouse	negative	Carcino_Mouse	positive	Carcino_Mouse	positive	Carcino_Mouse	negative
Carcino_Rat	positive	Carcino_Rat	negative	Carcino_Rat	negative	Carcino_Rat	negative
daphnia_at	0.0115144	daphnia_at	0.000935654	daphnia_at	0.000475985	daphnia_at	0.123348
hERG_inhibition	low_risk	hERG_inhibition	medium_risk	hERG_inhibition	medium_risk	hERG_inhibition	medium_risk
medaka_at	0.000190998	medaka_at	1.8686e-006	medaka_at	5.30127e-007	medaka_at	0.021524
minnow_at	7.85971e-005	minnow_at	2.14123e-006	minnow_at	4.01734e-007	minnow_at	0.0159882
TA100_10RLI	negative	TA100_10RLI	negative	TA100_10RLI	negative	TA100_10RLI	negative
TA100_NA	negative	TA100_NA	negative	TA100_NA	negative	TA100_NA	negative
TA1535_10RLI	negative	TA1535_10RLI	negative	TA1535_10RLI	negative	TA1535_10RLI	negative
TA1535_NA	positive	TA1535_NA	negative	TA1535_NA	negative	TA1535_NA	positive

**Figure 9.** Pre-ADME and Toxicity prediction of new analogues CIN4, COU5 and CAF1, PLM as a control.

The results showed that all the compounds have a good HIA percentage, among them compound CIN4 showed the highest HIA of 100%. All the compounds showed medium Caco-2 cell permeability, among them tangeretin, CIN4 and COU5 showed satisfactory values of 56.5975, 55.141 nm s<sup>-1</sup>, respectively. The compound CAF1 have medium MDCK permeability, among them CIN4 and COU5 have the highest values of 67.2924 and 66.5858. Permeability can be classified into low (< 25 nm/sec) and mean (> 25 to 500 nm/sec). The

distribution properties are verified by the % plasma protein binding (%PPB) values, and are found 100%. Such inhibitors got the potential to trigger the pharmacological response as a large amount of free drug is available to interact with their receptors. In addition to changing the pharmacological response of molecule the PPB also modifies the renal excretion because only unbound drug is available for glomerular filtration thus increasing excretion and decreasing half-life. Toxicity of most potent five compounds has been predicted by using preADMET tool and results are listed in the [Figure 8](#). Data obtained from the toxicity prediction studies indicating that all the new analogues are found to be non-carcinogenic with their negative results for the carcino Rat model. In contrast to this, analogues have shown positive results against carcino rat model except for compound CAF1. Present study has indicated the medium level associated with all compounds against hERG inhibition

### **Conclusion**

Given the link between obesity risk and chronic fat overconsumption, it has become critical to understand the role of lipid-mediated regulation of feeding behaviour. CD36 play a significant role in dietary lipid perception that demonstrate in this study by *in silico* methods

The synthesis of novel “fat taste-modifiers” that would mimic as natural dietary fatty acids and, by binding to lingual CD36, trigger “fat-like taste” sensation. These agents, in principle, would bring no calories and can replace lipids in our daily meal and can be used in the management/treatment of obesity and its associated pathological conditions. The computational analysis approach through protein-protein docking and MD simulations appears to be an essential tool to identify new analogues fat taste.

### **Declarations**

#### **Author contribution statement**

J. Akachar: Conceived and designed the experiments; Performed the experiments; Wrote the paper.

E.M. Bouricha : Analyzed and interpreted the data.

R. El Jaoudi, A. Ibrahimi and C. Etchebest: Conceived and designed the experiments; Analyzed and interpreted the data.

#### **Funding statement**

This work was supported by National Funding from the Moroccan Ministry of Higher Education and Scientific Research (Covid-19 Program) to AI and a grant to AI from Institute of Cancer Research and the PPR-1 program to AI.

### **Declaration of interests statement**

The authors declare no conflict of interest.

### **Additional information**

No additional information is available for this paper.

### **Reference**

1. Flegal KM, Carroll MD, Ogden CL, Curtin LR. Prevalence and trends in obesity among US adults, 1999-2008. *JAMA*. 2010 Jan 20;303(3):235-41.
2. Mokdad AH, Serdula MK, Dietz WH, Bowman BA, Marks JS, Koplan JP. The spread of the obesity epidemic in the United States, 1991-1998. *JAMA*. 1999 Oct 27;282(16):1519-22.
3. Zuccotti G, Fabiano V, Dilillo D, Picca M, Cravidi C, Brambilla P. Intakes of nutrients in Italian children with celiac disease and the role of commercially available gluten-free products. *J Hum Nutr Diet*. 2013 Oct;26(5):436-44. doi: 10.1111/jhn.12026. Epub 2012 Dec 17. PMID: 23252793.
4. Degrace-Passilly P, Besnard P. CD36 and taste of fat. *Curr Opin Clin Nutr Metab Care*. 2012 Mar;15(2):107-11. doi: 10.1097/MCO.0b013e32834ff19c. PMID: 22248592.
5. Greenwalt DE, Lipsky RH, Ockenhouse CF, Ikeda H, Tandon NN, Jamieson G. Membrane glycoprotein CD36: a review of its roles in adherence, signal transduction, and transfusion medicine. *Blood* 1992;80:1105-15.
6. Martin C, Chevrot M, Poirier H, Passilly-Degrace P, Niot I, Besnard P. CD36 as a lipid sensor. *Physiol Behav* 2011;105:36-42.
7. Baillie A, Coburn C, Abumrad N. Reversible binding of long-chain fatty acids to purified FAT, the adipose CD36 homolog. *J Membr Biol* 1996;153:75-81.
8. Febbraio M, Hajjar DP, Silverstein RL. CD36: a class B scavenger receptor involved in angiogenesis, atherosclerosis, inflammation, and lipid metabolism. *J Clin Invest* 2001;108:785-91.
9. Simons PJ, Kummer JA, Luiken JJ, Boon L. Apical CD36 immunolocalization in human and porcine taste buds from circumvallate and foliate papillae. *Acta Histochem* 2011;113:839-43.
10. Laugerette F, Passilly-Degrace P, Patris B, Niot I, Febbraio M, Montmayeur J-P, et al. CD36 involvement in orosensory detection of dietary lipids, spontaneous fat preference, and digestive secretions. *J Clin Invest* 2005;115:3177-84.
11. Hichami A, Khan AS, Khan NA. Cellular and Molecular Mechanisms of Fat Taste Perception. *Handb Exp Pharmacol*. 2022;275:247-270. doi: 10.1007/164\_2021\_437. PMID: 33547589
12. Shanmugamprema D, Muthuswamy K, Subramanian G, Ponnusamy V, Krishnan V, Subramaniam S. Fat taste signal transduction and its possible negative modulator components. *Prog Lipid Res*. 2020 Jul;79:101035. doi: 10.1016/j.plipres.2020.101035. Epub 2020 May 8. PMID: 32437786.
13. Murtaza B, Hichami A, Khan AS, Shimpukade B, Ulven T, Ozdener MH, Khan NA. Novel GPR120 agonist TUG891 modulates fat taste perception and preference and activates tongue-brain-gut axis in mice. *J Lipid Res*. 2020 Feb;61(2):133-142. doi: 10.1194/jlr.RA119000142. Epub 2019 Dec 5. PMID: 31806728; PMCID: PMC6997603.
14. Kumar N, Goel N. Phenolic acids: Natural versatile molecules with promising therapeutic applications. *Biotechnol Rep (Amst)*. 2019 Aug 20;24:e00370. doi: 10.1016/j.btre.2019.e00370. PMID: 31516850; PMCID: PMC6734135.



15. El-Seedi HR, El-Said AM, Khalifa SA, Goransson U, Bohlin L, Borg-Karlson AK, Verpoorte R. Biosynthesis, natural sources, dietary intake, pharmacokinetic properties, and biological activities of hydroxycinnamic acids. *J Agric Food Chem.* 2012;60(44):10877–95.
16. Alam MA, Subhan N, Hossain H, Hossain M, Reza HM, Rahman MM, Ullah MO. Hydroxycinnamic acid derivatives: a potential class of natural compounds for the management of lipid metabolism and obesity. *Nutr Metab (Lond).* 2016 Apr 11;13:27. doi: 10.1186/s12986-016-0080-3. PMID: 27069498; PMCID: PMC4827240.
17. Waterhouse, A. et al. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46(W1),W296–W303. <https://doi.org/10.1093/nar/gky427> (2018).
18. Lovell, S. C. et al. Structure validation by Calpha geometry: Phi, psi and Cbeta deviation. *Proteins* 50(3), 437–450. <https://doi.org/10.1002/prot.10286> (2003).
19. Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. Stereochemistry of polypeptide chain configurations. *J Mol Biol.* 7,95–99. [https://doi.org/10.1016/s0022-2836\(63\)80023-6](https://doi.org/10.1016/s0022-2836(63)80023-6) (1963).
20. Colovos, C. & Yeates, T. O. Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Sci.* 2(9), 1511–1519. <https://doi.org/10.1002/pro.5560020916> (1993).
21. Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, Han L, He J, He S, Shoemaker BA, Wang J, Yu B, Zhang J, Bryant SH. PubChem Substance and Compound databases. *Nucleic Acids Res.* 2016 Jan 4;44(D1):D1202-13. doi: 10.1093/nar/gkv951. Epub 2015 Sep 22. PMID: 26400175; PMCID: PMC4702940.
22. ChemAxon (2017). MarvinSketch 17.1.2.0. Budapest: ChemAxon.
23. Jihane.Akachar J, Etchebest C, El Jaoudi R, Ibrahimi A. The computational analyses, molecular dynamics of fatty-acid transport mechanism to the CD36 receptor. *Sci Rep.* 2021 Dec 1;11(1):23207. doi: 10.1038/s41598-021-01373-5. PMID: 34853341; PMCID: PMC8636502.
24. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem.* 2010 Jan 30;31(2):455-61. doi: 10.1002/jcc.21334. PMID: 19499576; PMCID: PMC3041641.
25. Wu, E. L. et al. CHARMM-GUI Membrane builder toward realistic biological membrane simulations. *J Comp Chem* 35, 1997–2004 (2014).
26. 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 (1983).
27. Abraham, M. J. et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 1–2, 19–25 (2015).
28. Huang, W. C. et al. Exogenous expression of human SGLT1 exhibits aggregations in sodium dodecyl sulfate polyacrylamide gel electrophoresis. *Am. J. Transl. Res.* 5, 441–449, <https://doi.org/10.1073/pnas.1511698112> (2013).
29. Miyamoto, S. & Kollman, P. A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J Comp Chem* 13, 952–962 (1992).
30. Vanommeslaeghe K, MacKerell AD Jr. Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. *J Chem Inf Model.* 2012 Dec 21;52(12):3144-54. doi: 10.1021/ci300363c. Epub 2012 Nov 28. PMID: 23146088; PMCID: PMC3528824.
31. Yuan, S., Chan, H. S. & Hu, Z. Using PyMOL as a platform for computational drug design. *Comput. Mol. Sci.* 7(2), e1298. <https://doi.org/10.1002/wcms.1298> (2017).
32. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* 14(1), 33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5) (1996).

33. Skjærven, L., Yao, X. Q., Scarabelli, G. & Grant, B. J. Integrating protein structural dynamics and evolutionary analysis with Bio3D. *BMC Bioinform.* 15(1), 399. <https://doi.org/10.1186/s12859-014-0399-6> (2014).
34. Mezei, M. & Filizola, M. TRAJELIX: A computational tool for the geometric characterization of protein helices during molecular dynamics simulations. *J. Comput. Aided Mol. Des.* 20(2), 97–107. <https://doi.org/10.1007/s10822-006-9039-1> (2006).
35. S. Release, 3: Desmond Molecular Dynamics System, DE Shaw research, New York, NY, Maestro-Desmond Interoperability Tools, Schrodinger, New York, NY, 2017, 2017.
36. Dahl, A. C., Chavent, M. & Sansom, M. S. Bendix: Intuitive helix geometry analysis and abstraction. *Bioinformatics* 28(16), 2193–2194. <https://doi.org/10.1093/bioinformatics/bts357> (2012).
37. Kovačević SZ, Jevrić LR, Podunavac Kuzmanović SO, Lončar ES. Prediction of In-silico ADME Properties of 1,2-O-Isopropylidene Aldohexose Derivatives. *Iran J Pharm Res.* 2014 Summer;13(3):899-907. PMID: 25276190; PMCID: PMC4177650.

**CHAPITRE IV: In silico structural mecanisme  
analysis of COVID-19 induced Fatty Acid  
Taste Impairments via the CD36 Receptor**



## IV.1.Résumé de l'article

### **Analyse *in silico* du mécanisme structurel des troubles du goût des acides gras via le récepteur CD36 lors de l'infection à la COVID-19**

Lors de la récente épidémie du coronavirus 2019 (COVID-19), l'Organisation mondiale de la santé (OMS) a défini cinq variantes préoccupantes (COV) du coronavirus respiratoire aigu sévère 2 (SRAS-CoV-2) : Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), Delta (lignées B.1.617.2 et AY) et Omicron (à l'origine B.1.1.529, puis reclassé en lignées BA). Ces derniers sont toujours un problème de santé public qui a causé de nombreux décès à l'échelle mondiale (Tracking., 2022). Les symptômes les plus courants observés chez les patients atteints de la COVID-19 comprenaient la fièvre, la toux, la diarrhée, la rhinorrhée, le mal de gorge et l'essoufflement, bien que certains patients aient développé une pneumonie sévère, un syndrome de détresse respiratoire aiguë et une défaillance multiviscérale entraînant la mort. Parallèlement à la propagation de la pandémie dans le monde entier, et à l'augmentation du nombre d'individus infectés, les dysfonctionnements olfactifs et gustatifs ont commencé à être connus et ils sont de plus en plus signalés chez un grand nombre de patients ((Parma *et al.*, 2020) ;(Santos *et al.*, 2021)). Le dysfonctionnement du goût a été signalé, parmi les symptômes les plus courants de la COVID-19, chez environ 65 % à 70 % des patients atteints d'une maladie légère à modérée présentant une altération chimiosensorielle pendant la phase aiguë de la COVID-19 ((Spinato G *et al.*, 2020) ;(Lechien JR *et al.*, 2020) ;(Vaira LA *et al.*, 2020)). Récemment, dans une série de 81 sujets testés positifs pour le variant Omicron du SARS-CoV-2, l'altération de l'odorat et du goût a été autodéclarée par 12 % et 23 % des patients, respectivement (Press Release.,2022).L'identification récente de plusieurs récepteurs lipidiques a montré que le CD36 est un candidat plausible en tant qu'un capteur lipidique gustatif, également connu sous le nom de transporteur d'acides gras. La séquence d'acides aminés CD36 existe sous la forme d'une glycoprotéine transmembranaire avec une structure de poche extracellulaire entre les queues amino et carboxyle cytoplasmiques (Greenwalt DE *et al.*, 1992), permettant la transduction du signal externe dans la cellule (Martin C *et al.*, 2011). Il a une affinité nanomolaire (Baillie A *et al.*,1996) pour une gamme de ligands à base de lipides, tels que les lipoprotéines, les cellules apoptotiques et les LCFA (Febbraio M *et al.*, 2001). Le CD36 a été détecté dans les papilles foliées et circonvallées humaines (Simons *et al.*, 2011).

Le but de notre étude était d'explorer par modélisation moléculaire et simulations dynamiques (MD) le dysfonctionnement gustatif de la chimioréception des acides gras spécifiquement le "goût de gras" des patients testés positifs au COVID-19. La méthodologie de l'étude a consisté à l'analyse de l'activation et la désactivation de CD36 en décrivant d'abord le mécanisme de liaison structurelle

putatif avec le coronavirus précédent SARS-CoV et avec le SARS-CoV-2 non mutant et son variant Omicron, et en comparant, par la suite, le complexe d'acides gras/ CD36. En outre, nous avons également utilisé la dynamique moléculaire pour étudier la dimérisation des parties transmembranaires (TM1 et TM2) de CD36, afin d'identifier le mécanisme structurel spécifique de dysfonctionnements du système de signalisation d'acide gras.

Les résultats de cette étude ont montré une affinité élevée et une stabilité du domaine RBD du SARS-CoV-2 non mutant au récepteur CD36 par rapport à celui du SRAS-CoV et du variant Omicron du SARS-CoV-2. L'étude des angles locaux du peptide de TM1 et de TM2 a révélé une intense dimérisation après fixation de SARS-CoV-2 entraînant un blocage de la signalisation du site de transport des acides gras. En conclusion, le CD36 représente une cible puissante pour le développement de peptides épitopes dérivés, ce qui devrait conduire à la découverte et conception d'un nouveau groupe thérapeutiques ciblant le au dysfonctionnement du goût par le Covid 19 et neutralisant le SRAS-CoV-2.

Ce chapitre est présenté sous forme d'un article soumis au Journal "Spring Journal of 3Biotech" (Akachar *et al.*, 2022). Dans cette publication, on a réalisé tout le travail d'analyse bioinformatique et de rédaction des manuscrits.

## IV.2. Article original

### **In silico structural mechanism analysis of COVID-19 induced Fatty Acid Taste Impairments via the CD36 Receptor**

Jihane Akachar<sup>1</sup>, El Mehdi Bouricha<sup>1</sup>, Rachid El Jaoudi<sup>1</sup> and Azeddine Ibrahim<sup>1</sup>

<sup>1</sup> Laboratory of Biotechnology (MedBiotech), Rabat Medical and Pharmacy School, Mohammed V University in Rabat, Rabat, Morocco.

#### **Abstract**

Taste dysfunctions are a common side effect of post-viral illness and known to impact quality of life and health status. Evidence is emerging that taste and smell loss are common symptoms of Covid-19 that may emerge and persist long after initial infection. The aim of the present study was to analysis Covid-19 variants (SARS-CoV-2, Omicron) signaling and alteration mechanism to fatty taste via taste bud lipid receptors like CD36 using molecular modeling and dynamics. In this paper, protein-protein docking was utilized for identifying the critical residues in CD36 which are involved in the binding with SARS-CoV, SARS-CoV-2, Omicron Spike receptor binding domain (RBD). The results of binding free energies showed a high affinity of SARS-CoV-2 RBD to CD36 receptor. Molecular docking and molecular dynamics simulations of CD36 showed favorable binding features and stability with SARS-CoV-2 RBD and therefore can be further considered as potential candidates in future anti-SARS CoV-2 drug discovery studies.

#### **Introduction**

The recent ongoing outbreak of coronavirus disease 2019 (COVID-19), the World Health Organization (WHO) defined five severe acute respiratory coronavirus 2 (SARS-CoV-2) variants of concern (VOC): Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), Delta (B.1.617.2 and AY lineages), and Omicron (originally B.1.1.529, then reclassified into BA lineages) is still a clinical problem which has caused widespread. The most common symptoms observed in COVID-19 patients included fever, cough, diarrhea, rhinorrhea, sore throat, and loss of taste and smell. (1)

Throat and shortness of breath, though some patients developed severe pneumonia, acute respiratory distress syndrome, and multiple organ failure resulting in death<sup>1</sup>. Moreover, in parallel with the pandemic spread to more countries and with the escalation in the numbers of infected individuals, olfactory and gustatory dysfunctions began to be known to everyone, since have been increasingly reported in a great number of patients (2,3). Taste dysfunction were consistently reported among the most common symptoms of COVID-19 with about 65%–70% of patients with mild-to-moderate disease experiencing a chemosensory impairment during the acute phase of the COVID-19. (4–6) Recently, in a series of 81 subjects tested positive for the SARS-CoV-2 Omicron variant, the impairment of the sense of smell and taste was self-reported by 12% and 23% of patients, respectively.(7 )

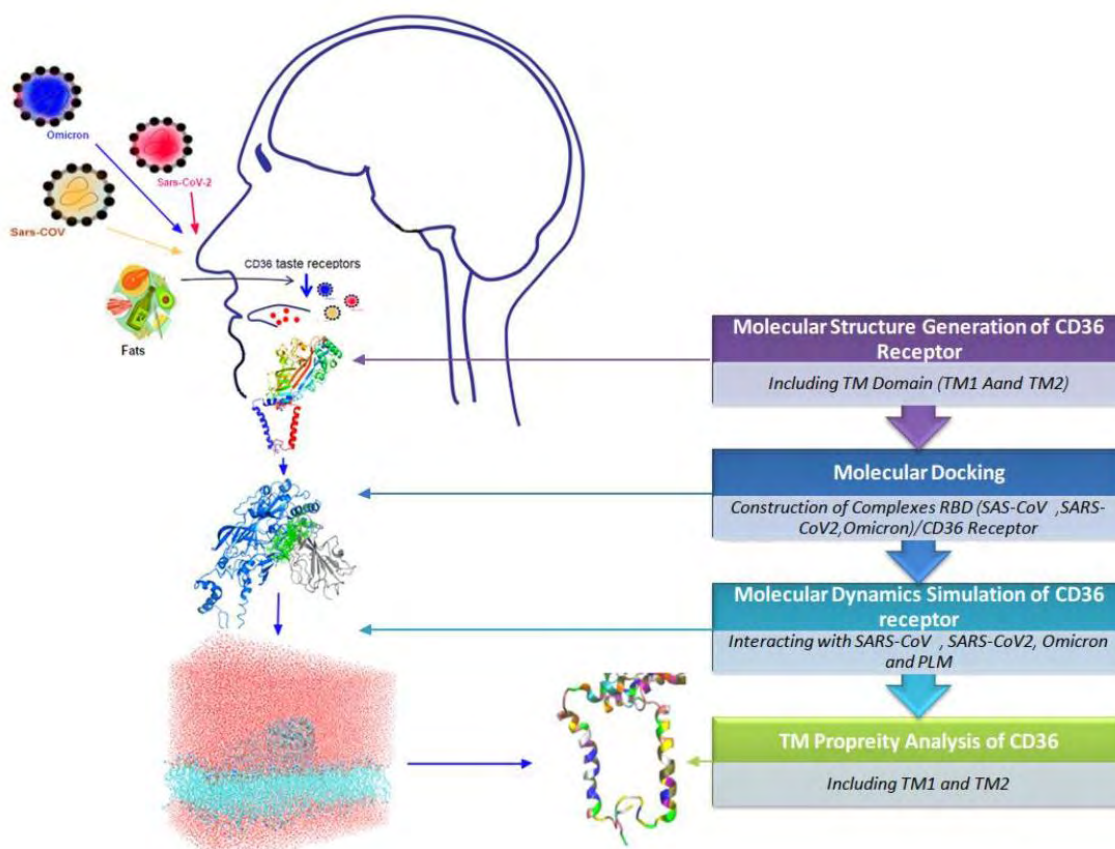
The recent identification of lipid-receptors (eg CD36) is a plausible candidate as the gustatory lipid sensor, which is also known as the fatty acid transporter. The CD36 amino acid sequence exists as a transmembrane glycoprotein with an extracellular pocket structure between the cytoplasmic amino and carboxyl terminal tails (8), allowing the transduction of external signal into the cell (9). It has a nanomolar-range affinity (10) to a range of lipid-based ligands, such as lipoprotein, apoptotic cells, and LCFAs (11). CD36 has been detected in human foliate and circumvallate papillae (12).

The aim of this study was to explore the gustatory dysfunction of fatty acid chemoreception specifically (fat taste) of patients tested positive for the COVID-19. Here we address CD36 activation and désactivation first describing a putative structural binding mechanism for a series of COVID-19 variants SARS-CoV-2, Omicron) and the previous coronavirus (SARS-CoV) to compare it with Fatty acid complex using supervised molecular modeling and dynamics (MD) simulations.

### **Computational methods**

For identifying a specific structural mechanism for the gustatory dysfunction against COVID-19. The workflow adopted in this computational study is illustrated in [Figure 1](#) and included screening of three targets COVID-19 (SARS-CoV-2, Omicron) and the previous coronavirus (SARS-CoV) as well as palmitic acid (PLM) which was used as a control in molecular docking and molecular dynamics simulations to predict the binding modes and estimate their binding affinity with SARS-CoV-2 RBD , SARS-CoV RBD and Omicron RBD.





**Figure 1.** Outline of computational mechanism of fat taste impairment via CD36 receptor. The methods included screening of three COVID 19 targets using molecular docking, MM-GBSA binding free energy estimation and molecular dynamics simulations.

### Data set

The crystallized structure files of The SARS-CoV-2 RBD [6M17], SARS-CoV RBD [2AJF], Omicron RBD [7QO7], CD36-CIDR $\alpha$  (PDB ID: 5LGD), were downloaded from the RCSB Protein Data Bank (PDB) (<https://www.rcsb.org/>), the PLM co-crystal to be included as a control. The initial structures were prepared by removing water and non-protein molecules and optimized in Discovery Studio (version 2.5.5) and PyMOl (version 1.6). The prepared structures were used as inputs for the molecular docking experiments.

### **Molecular Modeling CD36 complet**

The tridimensional structures of TMs (TM1 and TM2) transmembran domain of CD36 was predicted by using Modeld softwer to construct the homology model based on the crystal structure CD36-CIDR $\alpha$  (PDB ID: 5LGD) which is obtained from the Protein Data Bank, where all ligands and water molecules in the protein crystal structure were deleted (13,14). The entire sequence of human CD36 was taken from the UniProt Database ([http:// www.uniprot.org](http://www.uniprot.org)) with accession number P16671. The quality of the homology model was evaluated by Ramachandran plot (15), and its overall quality factor ERRAT2 (16) plot and PROVE produced by the SAVS server v6.0 ([https:// saves.mbi.ucla.edu/](https://saves.mbi.ucla.edu/)) and (<https://zlab.umassmed.edu>).

### **Construction of RBD (SARS-CoV, SARS-CoV2, Omicron)/CD36 receptor complexes by protein–protein docking**

The ZDOCK (version 2.1) (<http://zdock.umassmed.edu/>) protein-protein rigid body docking program based on the fast Fourier trans-form (FFT) correlation techniques (17-19) was used in this study to search globally for all possible binding configurations between the ligands (RBD of SARS-CoV-2 and SARSC-CoV and the lipid-receptor CD36. During protein-protein docking, the default parameters of ZDOCK were applied. The receptor was fixed and the ligand was sampled with an Euler angle interval of 15 throughout the rotational space, resulting in a total of 3600 rotations. For each rotation of the ligands, the final search was performed and only the top on the translational space using a grid spacing of 1.2 Å, translation with the highest configuration complementarily according to the scoring function in ZDOCK was kept. This yielded a total of 3600 putative binding modes of CD236 with SARS-CoV, SARS-CoV-2 and Omicron RBD domains. By default, the top 2000 putative CD36/CoV-2 RBD and CD36/CoV CD36/Omicron RBD binding configurations were retained in the final results of ZDOCK. These initial global search configurations from ZDOCK were further refined and rescored using the more accurate ITScore-PP method (20). The final docking poses were processed and analyzed using the tools of PRODIGY web server (<https://bianca.science.uu.nl/prodigy/>) (21,22). The graphical images were generated using PyMOL (version 1.6).



### **MM-GBSA binding free energy decomposition of the docked RBD (SARS-CoV, SARS-CoV2, Omicron)/CD36 receptor complexes**

Estimated free energies of binding of SARS-CoV-2 RBD/CD36, SARS-CoV RBD/CD36 and Omicron RBD/CD36 predicted complexes was calculated using the MM-GBSA decomposition module of the HawkDock server (<http://cadd.zju.edu.cn/hawkdock/>) (23) in order to analyze the individual energy contributions of each residue in the RBD/CD binding interface. All hydrogens and heavy atoms were added to the protein-protein complex using the tleap module in Amber16, and the ff02 force field. Afterward, the complex was optimized in vacuo by 2000 cycles of steepest descent and 3000 cycles of conjugate gradient minimizations. Finally, the polar desolvation energy was computed using the altered GB (GBOBC1) models. The outside and inside solute dielectric constants were set to 80 and 1, respectively. Graphpad Prisme program was then used to plot the graphs from the MM-GBSA computations.

### **Protein-protein molecular dynamics simulations of RBD (SARS-CoV, SARS-CoV2, Omicron)/CD36 receptor complexes**

The most realistic scenario of protein-protein interactions cannot be replicated by docking experiments. To better understand the stability of the complexes between SARS-CoV-2 RBD, SARS-CoV RBD and Omicron RBD with CD36 receptors, molecular dynamics simulations (MDS) were performed for 5000 ps using the desmond module of Schrodinger software 2017-4 (24).

The OPLS3e force field parameters were used in all simulations. A three-step process was followed which included system builder, energy minimization and molecular dynamics simulation. The structures of SARS-CoV-2, SARS-CoV and Omicron RBD docking complexes with CD36 were solvated in an orthorhombic boundary box using system builder by choosing the TIP3P as water model. The system charge was neutralized by adding Sodium (Na<sup>+</sup>) and chloride (Cl<sup>-</sup>) ions. The SHAKE algorithm was adopted for bond geometry constraints, and the cutoff radius parameter of 10 Å was chosen in Coulomb interactions. The electrostatic in-teractions were produced by on the particle mesh Ewald method (PME). The prepared systems in each run were subjected to energy minimization and equilibrated via an NPT ensemble using Berendsen thermostat at a temperature of 300K and 1 bar pressure. Finally, a total of 2000 ps of production MD was performed for each complex and the

trajectories were saved every 4.8ps. The stability of the complexes throughout the simulations was assessed by the root mean square deviation (RMSD) and the root mean square fluctuation (RMSF) metrics. Multiple sequence alignment and secondary structure estimation (SSE) were performed using ClustalW and the simulation interaction diagram (SID) implemented in the Desmond MD package.

#### **Analysis of Properties of Transmembrane TM1 and TM2 Helix of CD36 receptor**

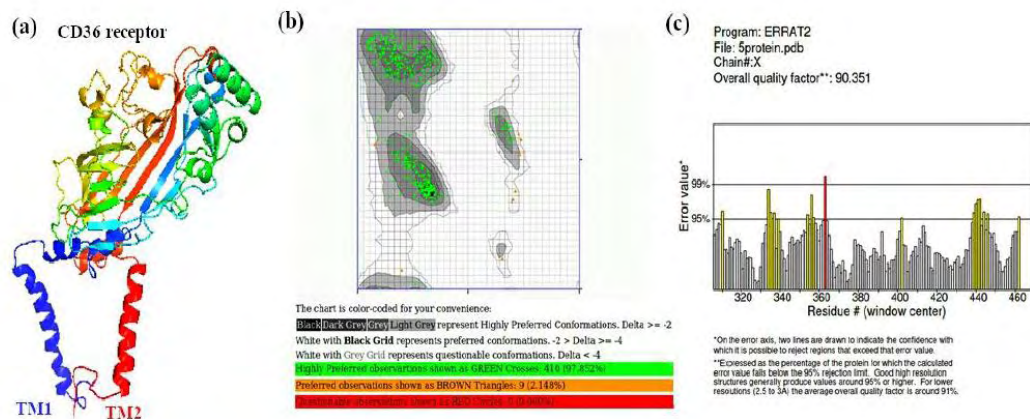
Analysis of the rotational movements of upper helix was performed with help of the VMD Bendix Plugin, Version 1.1 (25)

#### **Results and Discussion**

MD simulations can be used as a reply to many types of issues. Here we used MD simulations to determine a structural mechanism for the gustatory dysfunction against COVID-19. Via lipid receptor CD36. [Figure 1](#) illustrates an overview result of the applications of Molecular Dynamics Simulations, the timescale and the molecular motions accessible with the current simulation methods.

#### **Molecular Modeling CD36 complete TM domain**

The three dimensional (3D) structure of CD36 including TM domain (TM1 and TM2) predicted by using Modeled software to construct the homology model based on the crystal structure CD36-CIDR $\alpha$  (PDB ID: 5LGD). The quality of the homology model was evaluated by Ramachandran plot, the overall quality factor ERRAT2 plot produced by the SAVS server v6.0 (<https://saves.mbi.ucla.edu/>). The stereo chemical properties of the modeled demonstrates that a majority of its residues are in favored region (97.85%) and allowed region (2.14%) of the Ramachandran plot, that indicates reasonable quality of this homology model. Moreover, the ERRAT2 score was 90.35. We observed that the currently mutant modeled structure have structural qualities that are acceptable 24 ([Figure2](#)).

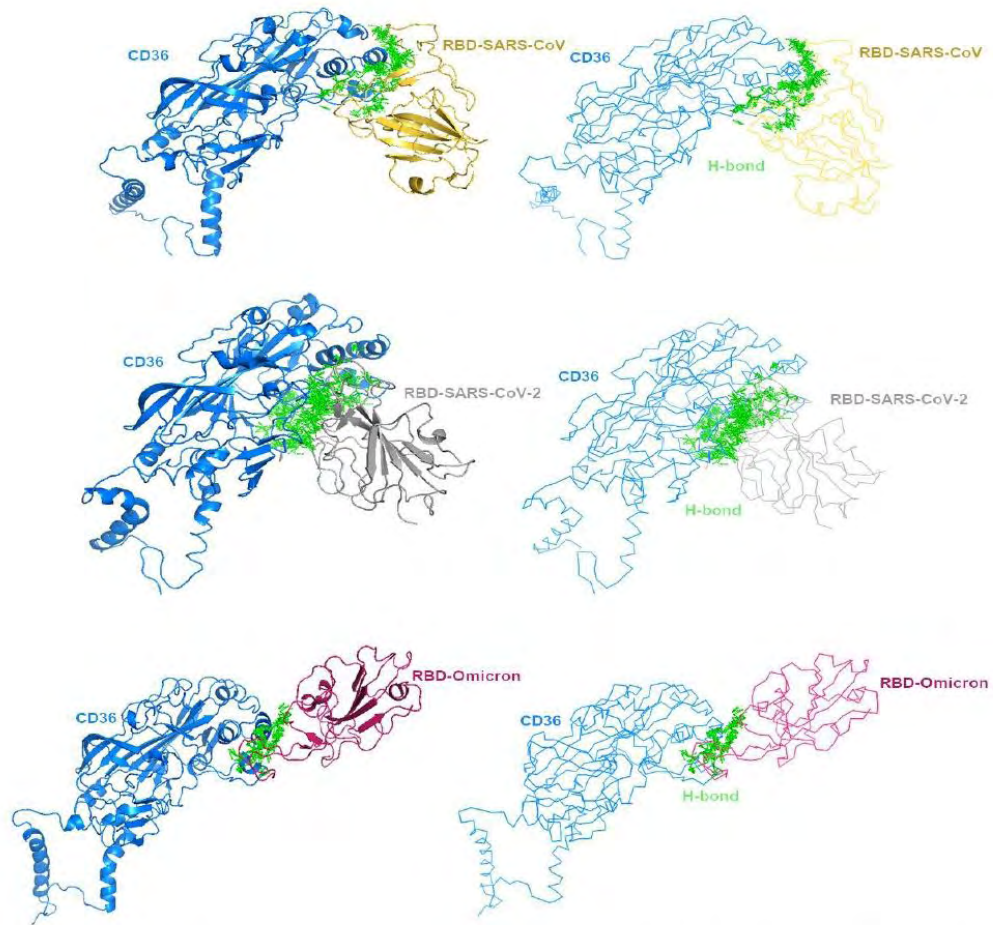


**Figure 2** : Molecular Modeling CD36 complet : (a) 3D Cartoon presentation of the CD36 with TMs (TM1 and TM2) structure. (b) Ramachandran plot calculations for CD36 protein. (c) The 3D profile verified results of the CD36 model using ERRAT program

### Molecular dockings studies of CD36 against the RBD of SARS-CoV1, SARS-CoV2 and Omicron

The molecular docking of SARS-CoV, SARS-CoV-2 and Omicron RBD against CD36 was performed using ZDOCK server. The top ranked 500 models were clustered into six groups with slightly different conformations. The binding affinity and MM-GBSA binding free energy of the docked complexes were calculated using PRODIGY and HawkDock webserver, respectively (Table 1). The complex of SARS-CoV-2 RBD with CD36 showed  $\Delta G$  value of 18.2 kcal mol, dissociation constant of  $4.3 \times 10^{-14}$  which are higher than those observed in the complexes with SARS-CoV and Omicron RBD. Moreover, the CD36 models presented a higher number of H-bonds in complex with SARS-CoV-2 RBD compared to SARS-CoV and Omicron RBD, suggests that this is a weak binding (Table 2, Figure 3).





**Figure 3.** Docking of SARS-CoV RBD (yellow cartoon) and SARS-CoV-2 RBD (grey cartoon) and Omicron (red cartoon) against CD36 receptor (blue color) using ZDOCK algorithm. Hydrogen bonds at the interface of two proteins are represented by the green lines. SARS-CoV-2 interacting with human CD36 were analyzed. The binding interaction between the RBD viral spikes is completely similar. The main interaction region of SARS-CoV-2 RBD with the CD36 receptor comprised 21 residues (Table 2), while only 11 residues were found in the interface of the SARS-CoV RBD and 8 residues for the Omicron RBD. In addition, as mentioned above, the existence of the two loops in the binding domain of CD36 is likely to induce a stabilizing effect on its interaction with the RBD of SARS-CoV-2. Figure 3 strongly suggest that these loops may enhance the electrostatic interactions between the SARS-CoV-2 RBD and the CD36 receptor.

Targets	$\Delta G$ (kcal mol <sup>-1</sup> )	Kd (M) at 25.0 °C	MM-GBSA $\Delta E$ binding
RBD-CoV	-11.0	9.0E-09	<b>-65.18</b>
RBD-CoV-2	-18.2	4.3E-14	<b>-128.37</b>
RBD-Omicron	-6.9	9.1E-06	<b>-30.66</b>

**Table 1.** Binding free energy (kcal.mol<sup>#1</sup>), binding affinity and dissociation constant of CD36, in complex with SARS-CoV, SARS-CoV2, Omicron RBD domain

Complex	H-bond residues CD36 receptor
RBD-CoV/CD36	Ile148, Tyr149, Asn151, Phe153, Val154, Met156, Ile157, Ser160, Pro191, Tyr192, Pro193
RBD-CoV-2/CD36	Ile148, Tyr149, Arg176, Glu177, Trp180, Gly181, Tyr182, Tyr192, Val194, Thr195, Phe201, Tyr202, Pro203, Asn205, Asn206, Thr207, Ala208, Lys231, Ser237, Tyr238, Glu335
RBD-Omicron/CD36	Tyr149, Asn151, Phe153, Val154, Met156, Ile157, Ser160, Tyr192
PLM/CD36	Arg63, Phe65, leu140, Ala143, Ala144, Ile148, Leu161, Ser162, Lys164, Leu187, Leu188, Leu189, Pro191, TYR192, Pro193, Val194, Thr195, Val198, Thr197, Gly199, Leu200, Phe201, Pro203, Tyr204, Thr207, Ser268, Ser269, Asp270, Ile271, Arg273, Ile330, Lys334, Glu335, Gly336, Arg337, Pro338, Ser339, Val339, Lys385, Leu387, Val389, Leu343, Ile341, Thr369, Leu391, Leu416, Glu418

**Table 2.** Hydrogen bond interactions at the SARS-CoV RBD-CD36 and SARS-CoV-1 RBD-CD36 and Omicron RBD-CD36 interfaces.

The residues present in these loops and showing direct interaction with the SARS-CoV-2 RBD are Ile148, Tyr149, Arg176, Glu177, Trp180, Gly181, Tyr182, Tyr192, Val194, Thr195, Phe201, Tyr202, Pro203, Asn205, Asn206, Thr207, Ala208, Lys231, Ser237, Tyr238, Glu335. As can be seen in [Table 2](#), the higher number of protein-protein contacts and the longer loops resulted in an increase in the binding affinities in SARS-CoV-2 RBD in comparison with SARS-CoV ([Table 1](#), [Figure 3,4](#)). Accordingly, these loops could have a main role in determining the host receptor specificity for the SARS-CoV-2 spike protein. These key residues of the CD36 receptor compared to the key residues binding to palmitic acid (PLM) Arg63, Phe65, leu140, Ala143, Ala144, **Ile148**, Leu161, Ser162, Lys164, Leu187, Leu188, Leu189, **Pro191**, **TYR192**, **Pro193**, **Val194**, **Thr195**, Val198, Thr197, Gly199, Leu200, **Phe201**, **Pro203**, Tyr204, **Thr207**, Ser268, Ser269, Asp270, Ile271, Arg273, Ile330, Lys334, **Glu335**, Gly336, Arg337, Pro338, Ser339, Val339, Lys385, Leu387, Val389, Leu343, Ile341, Thr369, Leu391,



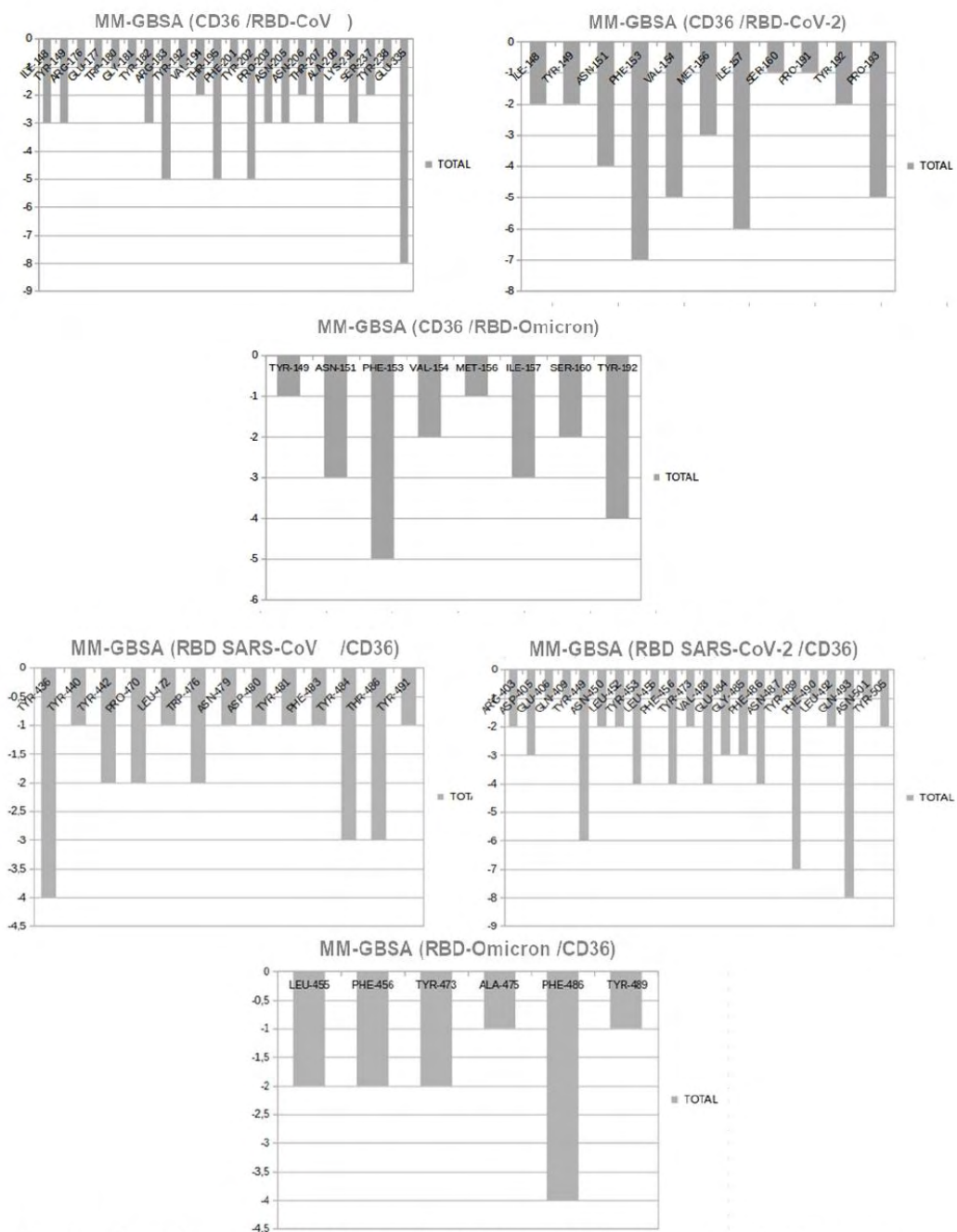
Leu416, Glu418 were used as reference simulations to determine a structural mechanism for the gustatory dysfunction.

Complex	CD36 receptor
<b>H-bond residues</b> RBD-CoV	Tyr436, Tyr440, Tyr442, Pro470, Leu472, Trp472, Trp476, Asn479, Asp480, Tyr481, Phe483, Thr486, Tyr491
<b>H-bond residues</b> RBD-CoV-2	Arg403, Asp405, Glu406, Gln409, Tyr449, Asn450, Leu452, Tyr453, Leu455, Phe456, Tyr473, Val483, Glu484, Gly485, Phe486, Asn487, Tyr489, Phe490, Leu492, Gln493, Asn501, Tyr505
<b>H-bond residues</b> RBD-Omicron	Leu455, Phe456, Tyr473, Ala475, Phe486, Tyr489

**Table 3.** Hydrogen bond interactions at the SARS-CoV RBD-CD36 and SARS-CoV-2 RBD-CD36 and Omicron RBD-CD36 interfaces.

### **Comparison of the binding free energies hydrogen bond residues using MM-GBSA of the lipid human CD36 receptor between SARS-CoV1, SARS-CoV2, Omicron RBD**

**Table 1** shows a comparison between the free energies of SARS-CoV-2 RBD, SARS-CoV and Omicron with lipid receptor CD36 using the MM-GBSA program in HawkDock server. The free energies value confirms the results of binding energy calculated with the PRODIGY web server. The complex RBD of SARS-CoV-2/CD36 shows a importantly strong binding free energy than that of SARS-CoV 65.18 and 30.66 kcal/mol for Omicron. The total of Hydrogen bond residues binding free energy that are -128.37 kcal/mol are 21 residues in complex SARS-CoV-2/CD36, compared with residues present in the SARS-CoV RBD/CD36 with 11 residues (**Figure 3,4**). These residues are worth considering as critical residues for the binding of RBD to a CD36. MM-GBSA binding free energy -128.37 kcal/mol.



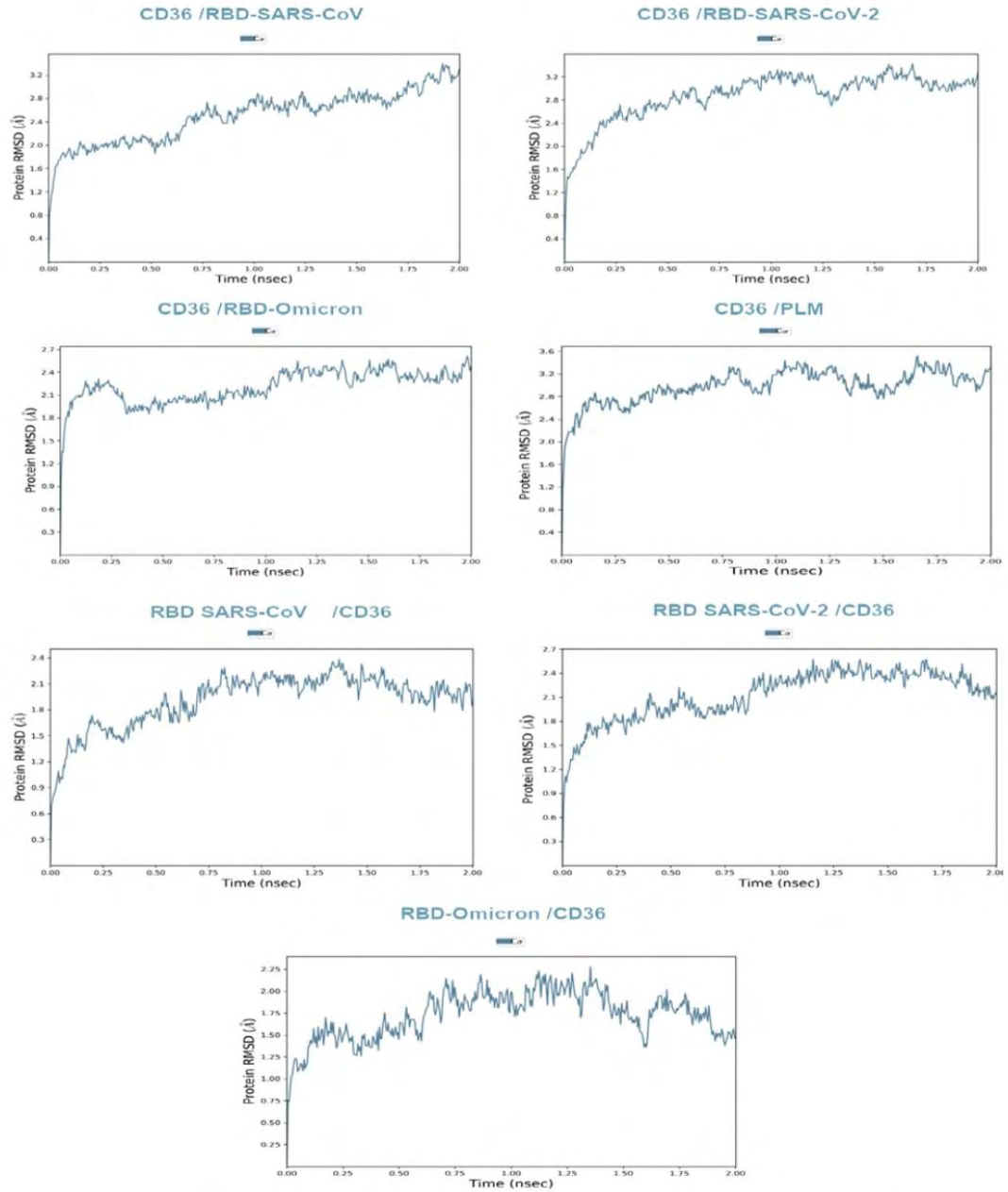
**Figure 4 .** Comparisons of binding free energies of Hydrogen bond residues of the CD36 receptor inside SARS-CoV, SARS-CoV2, Omicron RBD domain throughout the Docking using MM-GBSA HawkDock server.

### **Global molecular dynamics of the simulations analysis of the lipid human CD36 receptor between SARS-CoV1, SARS-CoV2, Omicron RBD**

This part of study aimed to understand the structure stability and mutating impact between the SARS-CoV, SARS-CoV-2 and Omicron RBD and the lipid receptor CD36 complexes. Previous studies showed low modifications in few residues of SARS-CoV, two main residues (479 and 487) that have undergone punctual mutations from civet to human, K479N, and S487T. These changes in the SARS-CoV-2 represent energetically favorable changes for the interaction with the receptor. MD simulations of the dynamic effects of SARS-CoV-2 RBD proteins and SARS-CoV with CD36, were analyzed with 2000 ps simulation period. Plot of root-mean-square deviation (RMSD) were produced by molecular dynamics (MD) of the complexes. The backbones RMSD value of the two complexes were stable at  $< 3 \text{ \AA}$  after a gradual rise around of 2 ns (Figure 5).

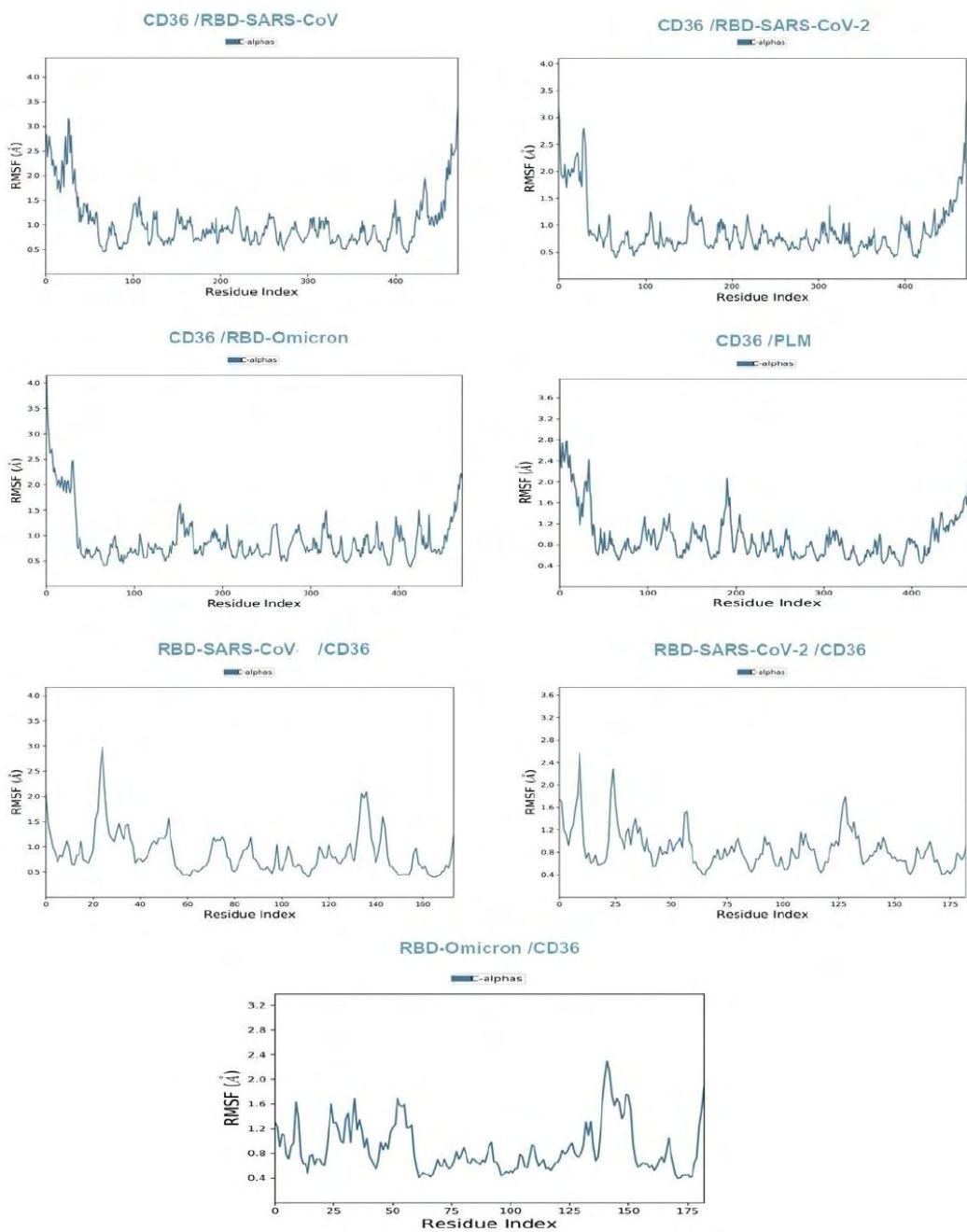
The RBD-CoV-2/CD36 complex rose from 2.4 to 2.8  $\text{\AA}$  around 2 ns and then remained stable where a small rise was observed. Whereas, the RBD-CoV/CD36 complex rose from 1.6 to 3.2  $\text{\AA}$  then was stable during the simulated period and RBD-Omicron CD36 complex rose from 2.0 to 2.8  $\text{\AA}$  compared to PLM complex rose from 2.8 to 2.8  $\text{\AA}$ . Stability in 2 ns simulation with average backbone RMSD of 0.4 and 0.25 nm, respectively. However, the RMSD showed the RBD SARS-CoV complex is less stable, it experienced a gradual increase in the RMSD value with fluctuations, stabilizing at an average of 0.25 nm. To investigate Protein flexibility that helped assessed of residues were involved of the complex, the C  $\alpha$  root mean square fluctuation (C  $\alpha$  RMSF) was calculated. Wherefore, the authors have examined the conformational flexibility of the RBD domains for SARS-CoV-2, SARS-CoV and Omicron with CD36 by analyzing their movement data. Figure 6 shows the C  $\alpha$  RMSF values for the complexes SARS-CoV-2 RBD/CD36, SARS-CoV RBD/CD36 and OMicron over a period of 2 ns simulations have been conducted. As seen in this figure that the RMSF of SARS-CoV-2 RBD have generally increased over time than these of SARS-CoV and Omicron with the CD36 receptor. On the other hand, the SARS-CoV-2 RBD is very extremely flexible than the SARS-CoV RBD. Therefore, most residues with a high C  $\alpha$  RMSF (Figure 6(A)) compared to SARS-CoV, especially higher in the loop near the binding site (470–490) (Figure 6(B)). To more comprehend the impact of this higher flexibility of SARS-CoV-2 compared to SARS-CoV and Omicron for CD36, the correlation between

modification of residues in sequence and secondary structure of proteins motions of RBD domain of SARS-CoV-2 and SARS-CoV was analyzed during different simulation.



**Figure 5:** Analysis of all backbone atom root mean square deviations (RMSD) of the CD36 receptor with SARS-CoV, SARS-CoV2, Omicron



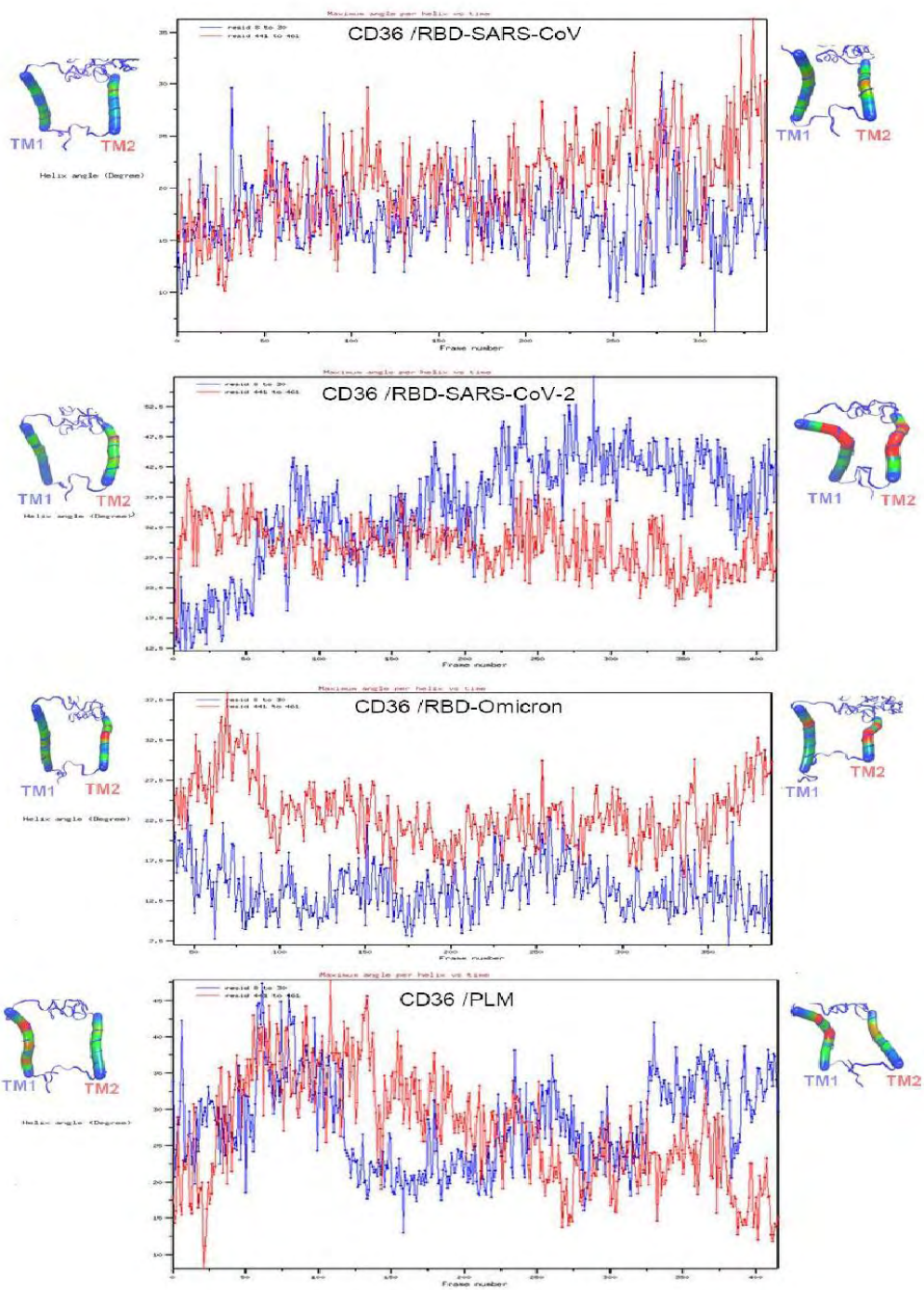


**Figure 6.** Root mean square fluctuations (RMSF) of of the CD36 receptor with SARS-CoV, SARS-CoV2, Omicron



### **Analysis of properties of helix Analysis TM1 and TM2 of CD36 receptor**

To study the evolution of the TM domain (TM1 and TM2) helix conformation of CD36 proteins over time, we concentrated on the structure analysis of upper helix using Bendix analyses it analyzed static peptide's local angles along the length of the helix showed in [Figure 7](#). The result of the analyses of static peptide's local angles along the length of TM1 and TM2 the helical angles were measured with 6.5 curvature degree for SARS-CoV-2 ([Figure 7](#)) in comparison with SARS-CoV ([Figure 7](#)). And the helical angles were measured with 20 curvature degree. On the whole, our investigation of upper helix proposes that the upper helix at of CD36 receptor tends to be wrapped, that influences the movement of the helix, loops, and TMs of opening fatty acid transport site signalisation as revealed by the covariance analysis.



**Figure 7. (a-d)** Analysis of Properties of Helix angle (TM1 and TM2) CD36 during MD simulations in presence of SARS-CoV, SARS-CoV2, Omicron

## **Conclusion**

The computational analysis approach through protein-protein docking and MD simulations appears to be an essential tool to identify specific taste dysfunctions structural mechanism such as for COVID-19 and all other viral. Indeed, the lipid receptor (CD36) receptor represents a powerful target on the RBD spike protein of SARS-CoV-2. Development of synthetic mimic peptides de-rived from CD36 receptor should lead to discovery a novel group of therapeutic and vaccine for SARS-CoV-2.

## **Declarations**

### **Author contribution statement**

J. Akachar: Conceived and designed the experiments; Performed the experiments; Wrote the paper.

E.M. Bouricha, Analyzed and interpreted the data.

R. El Jaoudi and A. Ibrahimi: Conceived and designed the experiments; Analyzed and interpreted the data.

### **Funding statement**

This work was supported by National Funding from the Moroccan Ministry of Higher Education and Scientific Research (Covid-19 Program) to AI and a grant to AI from Institute of Cancer Research and the PPR-1 program to AI.

### **Declaration of interests statement**

The authors declare no conflict of interest.

### **Additional information**

No additional information is available for this paper.

## **Reference**

1. Tracking SARS-CoV-2 variants. World Health Organization. 2022. Accessed March 18, 2022. <https://www.who.int/health-topics/typhoid/tracking-SARS-CoV-2-variants>
2. Parma, V. et al. More than smell—COVID-19 is associated with severe impairment of smell, taste, and chemesthesis. *Chem. Sens.* 45(7), 609–622 (2020).
3. Santos, R. E. A. et al. Onset and duration of symptoms of loss of smell/taste in patients with COVID-19: A systematic review. *Am.J. Otolaryngol.* 102889, 2225 (2021).
4. Spinato G, Fabbris C, Polesel J, et al. Alterations in smell or taste in mildly symptomatic outpatients with SARS-CoV-2 infection. *JAMA.* 2020;323:2089–2090.



5. Lechien JR, Chiesa-Estomba CM, Place S, et al. Clinical and epidemiological characteristics of 1,420 European patients with mild-to-moderate coronavirus disease 2019. *J Intern Med.* 2020;288:335–344.
6. Vaira LA, Deiana G, Fois AG, et al. Objective evaluation of anosmia and ageusia in COVID-19 patients: single-center experience on 72 cases. *Head Neck.* 2020;42:1252–1258.
7. [Press Release N° 08/2022 - Covid-19, flash survey Iss: 95.8% of Omicron positive samples on 17 January]. Istituto Superiore di Sanità. Accessed March 18, 2022. [https://old.iss.it/documents/20126/0/report\\_Flash\\_surveyVarianti\\_17+gennaio+2022.pdf/daa6af2a-3670-3757-bc8a-651042081c4f?t=1643624975322](https://old.iss.it/documents/20126/0/report_Flash_surveyVarianti_17+gennaio+2022.pdf/daa6af2a-3670-3757-bc8a-651042081c4f?t=1643624975322)
8. Greenwalt DE, Lipsky RH, Ockenhouse CF, Ikeda H, Tandon NN, Jamieson G. Membrane glycoprotein CD36: a review of its roles in adherence, signal transduction, and transfusion medicine. *Blood* 1992;80:1105–15.
9. Martin C, Chevrot M, Poirier H, Passilly-Degrace P, Niot I, Besnard P. CD36 as a lipid sensor. *Physiol Behav* 2011;105:36–42.
10. Baillie A, Coburn C, Abumrad N. Reversible binding of long-chain fatty acids to purified FAT, the adipose CD36 homolog. *J Membr Biol* 1996;153:75–81.
11. Febbraio M, Hajjar DP, Silverstein RL. CD36: a class B scavenger receptor involved in angiogenesis, atherosclerosis, inflammation, and lipid metabolism. *J Clin Invest* 2001;108:785–91.
12. Simons PJ, Kummer JA, Luiken JJ, Boon L. Apical CD36 immunolocalization in human and porcine taste buds from circumvallate and foliate papillae. *Acta Histochem* 2011;113:839–43.
13. Waterhouse, A. et al. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46(W1),W296–W303. <https://doi.org/10.1093/nar/gky427> (2018).
14. Lovell, S. C. et al. Structure validation by C $\alpha$  geometry: Phi, psi and C $\beta$  deviation. *Proteins* 50(3), 437–450. <https://doi.org/10.1002/prot.10286> (2003).
15. Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. Stereochemistry of polypeptide chain configurations. *J Mol Biol.* 7,95–99. [https://doi.org/10.1016/s0022-2836\(63\)80023-6](https://doi.org/10.1016/s0022-2836(63)80023-6) (1963).
16. Colovos, C. & Yeates, T. O. Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Sci.* 2(9), 1511–1519. <https://doi.org/10.1002/pro.5560020916> (1993).
17. R. Chen, L. Li, Z. Weng, ZDOCK: an initial-stage protein-docking algorithm, *Proteins: Struct. Funct. Bioinform.* 52 (1) (2003) 80–87.
18. R. Chen, Z. Weng, Docking unbound proteins using shape complementarity, desolvation, and electrostatics, *Proteins: Struct. Funct. Bioinform.* 47 (3) (2002) 281–294.
19. E. Katchalski-Katzir, I. Shariv, M. Eisenstein, A.A. Friesem, C. Aflalo, I.A. Vakser, Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques, *Proc. Natl. Acad. Sci. Unit. States Am.* 89 (6) (1992) 2195–2199.
20. S.Y. Huang, X. Zou, An iterative knowledge-based scoring function for protein–protein recognition, *Proteins: Struct. Funct. Bioinform.* 72 (2) (2008) 557–579.
21. L. Schrodinger, The PyMOL Molecular Graphics System, 2014, Version 1.6.
22. L.C. Xue, J.P. Rodrigues, P.L. Kastiris, A.M. Bonvin, A. Vangone, PRODIGY: a web server for predicting the binding affinity of protein–protein complexes, *Bioinformatics* 32 (23) (2016) 3676–3678.
23. G. Weng, E. Wang, Z. Wang, H. Liu, F. Zhu, D. Li, T. Hou, HawkDock: a web server to predict and analyze the protein–protein complex based on computational docking and MM/GBSA, *Nucleic Acids Res.* 47 (W1) (2019) W322–W330.

24. S. Release, 3: Desmond Molecular Dynamics System, DE Shaw research, New York, NY, Maestro-Desmond Interoperability Tools, Schrodinger, New York, NY, 2017
25. Dahl, A. C., Chavent, M. & Sansom, M. S. Bendix: Intuitive helix geometry analysis and abstraction. *Bioinformatics* 28(16),2193–2194. <https://doi.org/10.1093/bioinformatics/bts357> (2012)



**CHAPITRE V: Identifying epitopes for cluster of differentiation and design of new peptides inhibitors against human SARS-CoV-2 spike RBD by an *in-silico* approach.**

## V.1.Résumé de l'article

### **Identification par approche *in silico* des épitopes pour un cluster de différenciation et conception de nouveaux inhibiteurs de peptides contre le domaine RBD de la protéine "spike" du SARS-CoV-2 humain.**

La maladie à coronavirus 19 (COVID-19) est une infection hautement contagieuse et à propagation rapide causée par le coronavirus 2 du syndrome respiratoire aigu sévère (SRAS-CoV-2). Dans certains cas, la maladie peut être mortelle, entraînant plus d'un million de décès dans le monde selon l'OMS. À l'heure actuelle, il n'existe aucun traitement efficace contre le COVID-19, mais de nombreux inhibiteurs de structure peptidique ont montré une puissante activité antivirale contre le SRAS-CoV-2 et certains d'entre eux sont actuellement en cours d'essais cliniques. Malgré leurs activités prometteuses, le développement de petites molécules à usage clinique peut être limité par de nombreux facteurs tels que l'effet hors cible, la faible stabilité et la faible biodisponibilité. Le peptide T20, utilisé pour les soins du VIH, est un exemple réussi d'inhibiteurs de l'enveloppe virale qui imite le domaine COOH-terminal de la sous-unité gp41, y compris le VIH1, et l'empêche de se lier aux récepteurs des cellules hôtes (Lalezaret *al.*, 2003 ; Champagneet *al.*, 2009). Il a été démontré que les mutations S487T et K479N dans les boucles externes du domaine RBD de la protéine spike du SARS-CoV améliorent la liaison virale au récepteur ACE2 des cellules hôtes et peuvent être responsables du saut du virus des civettes aux humains (Wanet *al.*, 2015 ; Li *et al.*, 2013). Les classes de différenciation CD147, CD209, CD299 ont été identifiées comme des corécepteurs d'entrée essentiels pour la spécificité des espèces du SRAS-CoV-2 chez l'homme, bien que les mécanismes sous-jacents ne soient pas encore entièrement élucidés.

Dans notre étude, le criblage virtuel par amarrage protéine-protéine a été utilisé dans un premier temps pour identifier les épitopes critiques pour trois cibles (CD147, CD209 et CD299), en utilisant deux serveurs (CABS-Dock et ZDOCK). L'énergie libre de liaison calculée par la méthode MM-GBSA a été appliquée par la suite pour déterminer la force de la liaison du domaine RBD de la protéine spike du SARS-CoV-2 vis à vis de ces trois cibles et en utilisant le récepteur ACE2 comme référence. Cette énergie MM-GBSA a été décomposée pour révéler les résidus de points chauds à l'interface des complexes RBD SARS-CoV-2 /CD299. Ce dernier a été utilisé, par la suite, pour produire des dérivés des séquences peptidiques hypothétiques avec des activités de liaison spécifiques au domaine RBD du SARS-CoV-2. En fin, la dynamique moléculaire (MD) a été utilisée pour l'étude du comportement des deux complexes RBD SARS-CoV-2 /CD299 et RBD

SARS-CoV /CD299 et également pour les dérivés des séquences peptidiques. Ces derniers ont fait l'objet d'une analyse des propriétés physico-chimiques en utilisant le serveur PepCalc.

Les résultats des énergies libres de liaison de ces trois cibles ont montré une affinité élevée du domaine RBD du SARS-CoV-2 pour le récepteur CD299. En outre, ces résultats nous ont permis de faire la conception de quatre inhibiteurs peptidiques hypothétiques (P1, P2, P3 et P4). Après une étude d'amarrage et de dynamique moléculaire, ces peptides ont montré des caractéristiques de liaison conduisant à une stabilité favorables avec le domaine RBD du SARS-CoV-2. De plus, l'analyse des propriétés physico-chimiques a révélé que tous ces peptides ont une solubilité satisfaisante. Ces quatre peptides peuvent donc être considérés d'avantage comme des candidats potentiels dans les futures études de découverte de médicaments anti-SARS CoV-2.

Ce chapitre est présenté sous forme d'un article publié dans le Journal Heliyon. 2020 (Akachar *et al.*, 2020). Dans cette publication, on a réalisé tout le travail d'analyse bioinformatique et de rédaction des manuscrits.



Contents lists available at ScienceDirect

Heliyon

journal homepage: [www.cell.com/heliyon](http://www.cell.com/heliyon)

## Research article

Identifying epitopes for cluster of differentiation and design of new peptides inhibitors against human SARS-CoV-2 spike RBD by an *in-silico* approach

Jihane Akachar<sup>a,\*</sup>, El Mehdi Bouricha<sup>a</sup>, Mohammed Hakmi<sup>a</sup>, Lahcen Belyamani<sup>b</sup>, Rachid El Jaoudi<sup>a</sup>, Azeddine Ibrahim<sup>a</sup>

<sup>a</sup> Medical Biotechnology Laboratory (MedBiotech), Biotinova Research Center, Rabat Medical & Pharmacy School, Mohammed Vth University in Rabat, Morocco

<sup>b</sup> Emergency Department, Military Hospital Mohammed V, Rabat Medical & Pharmacy School, Mohammed Vth University in Rabat, Morocco

## ARTICLE INFO

## Keywords:

Coronavirus 19  
Molecular docking  
Molecular dynamics  
Peptide-based drugs  
Spike protein  
Cluster of differentiation  
Computer science  
Engineering  
Physics  
Chemistry  
Biological sciences

## ABSTRACT

The coronavirus disease 19 (COVID-19) is a highly contagious and rapidly spreading infection caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). In some cases, the disease can be fatal which resulted in more than one million deaths worldwide according to the WHO. Currently, there is no effective vaccine or treatment for COVID-19, however many small-molecule inhibitors have shown potent antiviral activity against SARS-CoV-2 and some of them are now under clinical trials. Despite their promising activities, the development of these small molecules for the clinical use can be limited by many factors like the off-target effect, the poor stability, and the low bioavailability. The clusters of differentiation CD147, CD209, CD299 have been identified as essential entry co-receptors for SARS-CoV-2 species specificity to humans, although the underlying mechanisms are yet to be fully elucidated. In this paper, protein-protein docking was utilized for identifying the critical epitopes in CD147, CD209 and CD299 which are involved in the binding with SARS-CoV-2 Spike receptor binding domain (RBD). The results of binding free energies showed a high affinity of SARS-CoV-2 RBD to CD299 receptor which was used as a reference to derive hypothetical peptide sequences with specific binding activities to SARS-CoV-2 RBD. Molecular docking and molecular dynamics simulations of the newly designed peptides showed favorable binding features and stability with SARS-CoV-2 RBD and therefore can be further considered as potential candidates in future anti-SARS CoV-2 drug discovery studies.

## 1. Introduction

In December 2019, an epidemic of pneumonia, named COVID-19 by the World Health Organization (WHO), emerged in Wuhan city of China and rapidly spread worldwide; causing an ongoing pandemic [1, 2]. Whole genome sequencing of the causative virus showed that it belongs to the beta-coronavirus subfamily with most similarity to SARS-like BAT coronaviruses bat-SL-CoVZC45 and bat-SL-CoVZXC21 [3]. This new coronavirus was named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) by the Coronavirus Study Group of the International Committee on Taxonomy of Viruses [4]. The previous coronavirus (SARS-CoV) binds to human angiotensin-converting enzyme 2 (ACE2) receptor via its spike protein receptor-binding domain (RBD). It has been signaled also that SARS-CoV-2 can recognize some CD markers along with ACE2 in order to infect human cells [5, 6, 7]. A wide range of studies have rushed to accelerate the discovery of new drug and vaccine candidates against SARS-CoV-2, specially a new target receptor of the immune

response as a CD marker that binding to S protein. The large liberation of cytokines through the immune system in reaction to the viral infection and/or secondary inflammation can consequence in a cytokine storm and offer of sepsis that are the induce of death in 28% of incurable COVID-19 patient [8].

T cell receptor (TCR) has a central role in adaptive immune responses, their repertoires are considerably decreased in the first stages of acute SARS-CoV-2 infection but start to return to normal in the convalescent stage [9]. Dendritic cells (DC), which are highly abundant in the lung and sensitive to viral diseases, induce a TCR-mediated cell apoptosis in the absence of costimulatory molecules [10]. Several RNA expression profiling studies have demonstrated that human DC also express the HACE2 receptor alongside with the dendritic cell-specific ICAM-3-grabbing non-integrin (DC-SIGN), especially CD209 and CD299 DC-SIGNs of family C-type lectin receptors which are found on the surface of macrophages and dendritic cells. Those receptors are characterized by a high affinity binding to mannose-rich glycans that are present on other

\* Corresponding author.

E-mail address: [jihane.akachar@um5s.net.ma](mailto:jihane.akachar@um5s.net.ma) (J. Akachar).

<https://doi.org/10.1016/j.heliyon.2020.e05739>

Received 28 August 2020; Received in revised form 27 October 2020; Accepted 11 December 2020

2405-8440/© 2020 Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

cellular receptors or pathogens [11, 12, 13]. L-SIGN (CD299) act as a mannose receptor for several viruses and it is associated with viral entry into host cells [14, 15]. C-type lectin receptor was firstly determined as a human immunodeficiency virus (HIV) attachment factor and described after as a receptor for hepatitis C virus and dengue virus. L-SIGN, which is related to ACE2, is present also on human type II alveolar cells and may enhance the ACE2-mediated cellular attachment and entry of SARS-CoV-2, Ebola virus, cytomegalovirus and other viruses as well. CD147 was described as a red blood cell (RBC) receptor for the parasite *Plasmodium falciparum* and was also cited as a target for COVID-19 known as Basigin (BSG) [16, 17, 18, 19, 20, 21]. CD147 is a heavily glycosylated protein that functions as a principal upriver stimulator of the matrix metalloproteinases (MMPs); and its expression has been shown to be up-regulated in some conditions such as asthma and diabetes [22]. CD147 expression is induced by the high glucose concentration (25 mM) in monocytes [23]. Recently, Bao et al., suggested potential correlation between CD147 and diabetes mellitus in clinical complications resulting from SARS-CoV-2 infection [24]. Thus, the design of peptide inhibitors targeting the RBD domain of SARS-CoV-2 spike protein can be a promising strategy to impair the viral entry into host cells. Peptides are reported as powerful drug-like inhibitors at the mesoscale over chemistry and protein therapeutics. Peptides are characterized by their high specificity and potency toward the targets, as well as a low toxicity and limited undesirable effects toward the organism [25]. Furthermore, peptides are central constituents of the immune systems in protozoans, invertebrates, flora, and vertebrates, including mammals, that are able to neutralize various infectious agents like viruses, microorganisms, and mushroom [26]. The mimicking peptides designed [27] to detect the specific epitopes of proteins which are involved in protein-protein disorganization [28]. The T20 peptide, used for HIV care, is a

successful example of viral envelope inhibitors that mimics the COOH-terminus domain of the gp41 subunit including HIV1 and prevent it from binding to the host cells receptors [29, 30]. It has been shown that the S487T and K479N mutations in the external loops of SARS-CoV spike RBD enhances the binding of the viral spike to the host cells receptor ACE2 and may be responsible of the virus jump from civets to humans [31, 32]. In the search of peptide inhibitors targeting the RBD of the SARS-CoV-2 spike, we designed four hypothetical peptide inhibitors (P1, P2, P3, and P4) based on interfaces of the docked complexes of RBD with CD147, CD299 and CD209 using CABS-Dock and ZDOCK servers. The free energy of binding computed by the MM-GBSA method was applied to determine the strength of the binding RBD to CD147, CD299, CD209 and ACE2 which was taken as a reference. The MM-GBSA energy was further decomposed as per-residue contributions to reveal the hotspot residues in the interface of the RBD/CD complexes and molecular dynamics (MD) are a powerful method for improving stability and fluctuation of protein residues [33]. Finally, the spike glycoprotein epitopes in CD299 were identified and four peptides were designed through computational approaches to block the binding interactions between the RBD domain of SARS-CoV-2 and its CD targets.

2. Materials and methods

For identifying a specific Epitopes CD marker and design a new antiviral peptide against COVID- 19, the workflow adopted in this computational study is illustrated in Figure 1 and included screening of three target CD markers CD147, CD209 and CD299 as well as hACE2 which was used as a control in molecular docking and molecular dynamics simulations to predict the binding modes and estimate their binding affinity with SARS-CoV-2 RBD and SARS-CoV RBD.

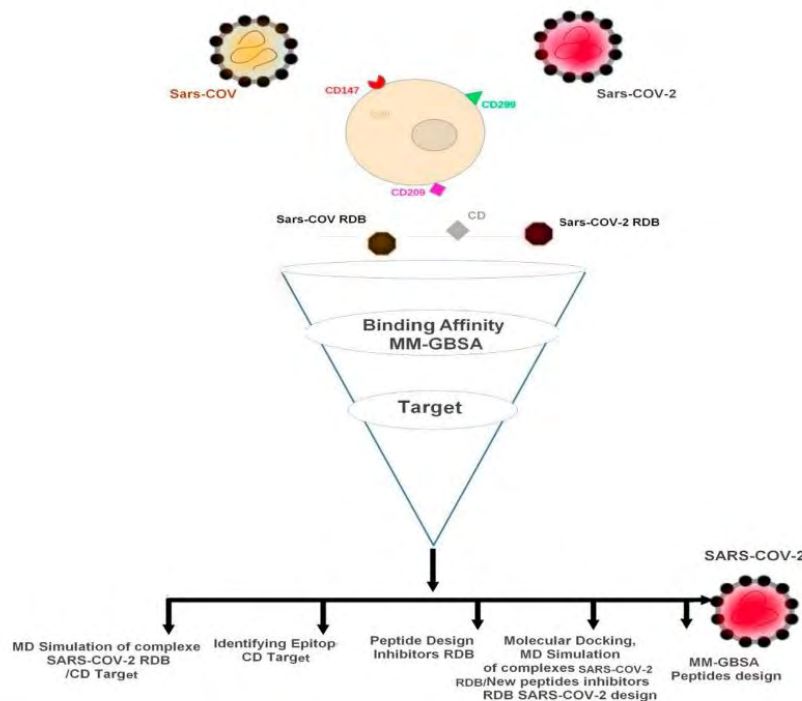


Figure 1. An overview of Integrative workflow adopted in this study. The methods included screening of three CD Markers targets using molecular docking, MM-GBSA binding free energy estimation and molecular dynamics simulations.



**Table 1.** Binding free energy (kcal.mol<sup>-1</sup>), binding affinity and dissociation constant of CD147, CD209, CD299 and ACE2 in complex with SARS-CoV-1 and SARS-CoV-2 RBD.

Targets	RBD-CoV-2			RBD-CoV-1		
	$\Delta G$ (kcal mol <sup>-1</sup> )	Kd	MM-GBSA $\Delta E^{binding}$	$\Delta G$ (kcal mol <sup>-1</sup> )	Kd	MM-GBSA $\Delta E^{binding}$
CD147	-12.2	1.2E-09	-43.62	-13.6	1.0E-10	-63.29
CD299	<b>-19.3</b>	7.3E-15	-69.42	<b>-15.5</b>	4.2E-12	-63.87
CD209	-18.7	2.0E-14	-71.14	-17.6	1.3E-13	-62.22
ACE2	-10.5	2.0E-08	-46.77	-10.8	1.2E-08	-56.26

The bold formatted values presented a higher binding affinity of CD299 models in complex with SARS-CoV-2 RBD and SARS-CoV-1 RBD.

### 2.1. Data set

The crystallized structure files of The SARS-CoV-2 RBD [6M17], SARS-CoV RBD [2AJF], CD147 [5X0T], CD299 [1K9J], CD209 [1S14] and ACE2 [6M17] were downloaded from the RCSB Protein Data Bank (PDB) (<https://www.rcsb.org/>). The anti-SARS-CoV peptide [AVPid: AVP1477] was downloaded from the Database of Antiviral Peptides (AVPdb) (<http://crdd.osdd.net/servers/avpdb/index.php>) to be included as a control. The initial structures were prepared by removing water and non-protein molecules and optimized in Discovery Studio (version 2.5.5) and PyMOL (version 1.6). The prepared structures were used as inputs for the molecular docking experiments.

### 2.2. Construction of RBD/CD marker complexes by protein-protein docking

The ZDOCK (version 2.1) (<http://zdock.umassmed.edu/>) protein-protein rigid body docking program based on the fast Fourier transform (FFT) correlation techniques [34, 35, 36] was used in this study to search globally for all possible binding configurations between the ligands (RBD of CoV-2 and CoV-1) and the receptors (CD markers). During protein-protein docking, the default parameters of ZDOCK were applied. The receptor was fixed and the ligand was sampled with an Euler angle interval of 15 throughout the rotational space, resulting in a total of 3600 rotations. For each rotation of the ligands, the final search was performed on the translational space using a grid spacing of 1.2 Å, and only the top translation with the highest configuration complementarity according to the scoring function in ZDOCK was kept. This yielded a total of 3600 putative binding modes of CD147, CD299 and CD209 with SARS-CoV and SARS-CoV-2 RBD domains. By default, the top 2000 putative CD/CoV-2 RBD and CD/CoV-1 RBD binding configurations were retained

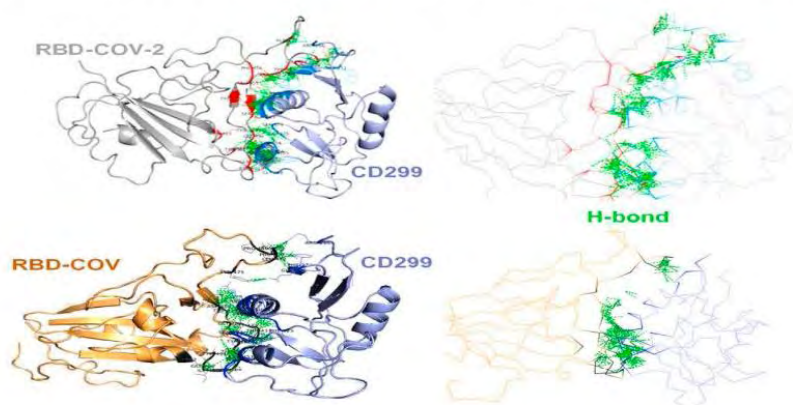
in the final results of ZDOCK. These initial global search configurations from ZDOCK were further refined and rescored using the more accurate ITScore-PP method [37]. The final docking poses were processed and analyzed using the tools of PRODIGY web server (<https://bianca.science.uu.nl/prodigy/>) [38, 39]. The graphical images were generated using PyMOL (version 1.6).

### 2.3. MM-GBSA binding free energy decomposition of the docked RBD/CD markers complexes

Estimated free energies of binding of SARS-CoV-2 RBD/CD and SARS-CoV RBD/CD predicted complexes was calculated using the MM-GBSA decomposition module of the HawkDock server (<http://cadd.zju.edu.cn/hawkdock/>) [40] in order to analyze the individual energy contributions of each residue in the RBD/CD binding interface. All hydrogens and heavy atoms were added to the protein-protein complex using the leap module in Amber16, and the ff02 force field. Afterward, the complex was optimized in vacuo by 2000 cycles of steepest descent and 3000 cycles of conjugate gradient minimizations. Finally, the polar desolvation energy was computed using the altered GB (GBOBC1) models. The outside and inside solute dielectric constants were set to 80 and 1, respectively. Graphpad Prisme program was then used to plot the graphs from the MM-GBSA computations.

### 2.4. Protein-protein molecular dynamics simulations

The most realistic scenario of protein-protein interactions cannot be replicated by docking experiments. To better understand the stability of the complexes between SARS-CoV-2 RBD and SARS-CoV RBD with CD receptors, molecular dynamics simulations (MDS) were performed for 5000 ps using the desmond module of Schrödinger software 2017-4 [41].



**Figure 2.** Docking of SARS-CoV-2 RBD (yellow cartoon) and SARS-CoV-1 RBD (grey cartoon) against CD299 receptor using ZDOCK algorithm. Hydrogen bonds at the interface of two proteins are represented by the green lines.

**Table 2.** Hydrogen bond interactions at the SARS-CoV-2 RBD-CD299 and SARS-CoV-1 RBD-CD299 interfaces.

Complex	H-bond residues CD299
DB-CoV-2/CD299	Arg266, Cys268, Pro269, Asp271, Thr273, Phe274, Phe275, Gln276, Ser285, Ala309, Glu310, Gln312, Asn313, Phe314, Leu317, Gln318, Arg321, Pro349, Ser350, Gln352, Arg353, Tyr354, Trp355, Asn356, Ser357, Ser374, Gly375, Trp376,
RBD-CoV/CD299	Arg266, Phe274, Gln276, Asn313, Leu317, Arg321, Thr319, Ser320, Arg324, Asn323, Phe325, Ser372, Gly373, Gly375, Ser374, Asn356, Ser357, Arg353
Mannose/CD299	Thr319, Ser320, Asn323, Glu359, Asn361, Ser363, Glu366, Ser372, Asn377, Asp378

The OPLS3e force field parameters were used in all simulations. A three-step process was followed which included system builder, energy minimization and molecular dynamics simulation. The structures of SARS-CoV-2 and SARS-CoV RBD docking complexes with CD were solvated in an orthorhombic boundary box using system builder by choosing the TIP3P as water model. The system charge was neutralized by adding Sodium (Na<sup>+</sup>) and chloride (Cl<sup>-</sup>) ions. The SHAKE algorithm was adopted for bond geometry constraints, and the cutoff radius parameter of 10 Å was chosen in Coulomb interactions. The electrostatic interactions were produced by on the particle mesh Ewald method (PME). The prepared systems in each run were subjected to energy minimization and equilibrated via an NPT ensemble using Berendsen thermostat at a temperature of 300K and 1 bar pressure. Finally, a total of 5000 ps of production MD was performed for each complex and the trajectories were saved every 4.8ps. The stability of the complexes throughout the simulations was assessed by the root mean square deviation (RMSD) and the root mean square fluctuation (RMSF) metrics. Multiple sequence alignment and secondary structure estimation (SSE) were performed using ClustalW [42] and the simulation interaction diagram (SID) implemented in the Desmond MD package.

### 2.5. Docking of the designed peptides at RBD domain of SARS-CoV-2

The sequences of new specific peptide inhibitors were predicted from the binding interface between the RBD of SARS-CoV-2 and CD299 receptor. These peptides were selected based on their predicted binding affinity, H-bonds count and MM-GBSA binding free energy. A total of 4 peptide sequences were selected and built using PyMOL (version 1.6). These peptides were derived from CD299 receptor as the starting structures of a molecular system used for the docking. Protein-peptide docking was carried out using the CABS-dock server (<http://biocomp.chem.uw.edu.pl/CABSdock>) [43]. This tool requires two inputs: the amino acid sequences of the peptides P1, P2, P3, P4, and the antiviral SARS-CoV as a control, and the 3D structure of the receptor (Spike SARS-CoV-2 RBD, PDB ID: 6M17) after elimination unnecessary co-crystallized water and ions. Protein-docking was carried out using the default settings of CABS-dock server.

### 2.6. Calculation of the free energy of binding of the designed peptides at RBD domain of SARS-CoV-2

The binding free energy calculation was carried out to predict the conformational stability of the complexes between SARS-CoV-2 RBD and the newly designed peptides. The MM-GBSA method implemented in the HawkDock server (<http://cadd.zju.edu.cn/hawkdock/>) was used for the estimation of the binding free energy of the studied complexes. The MM-GBSA decomposition plots were produced using GraphPad Prism 5 software (GraphPad Software, La Jolla, CA, USA).

### 2.7. Molecular dynamics simulations of the designed peptides in complex with RBD domain SARS-CoV-2

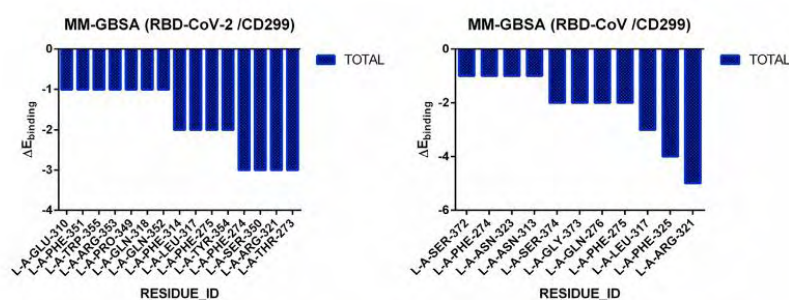
Molecular dynamics (MD) simulations of the docking complexes of RBD with the 4 candidate peptides were performed for 5000 ps using the desmond simulation package included in Schrödinger suite 2017-4 as described in section 2.4.

### 2.8. Analysis of the physicochemical properties of the candidate peptide designs

The physicochemical properties of new peptides were evaluated by online peptide calculators Pepcalc server including (theoretical molecular weight, iso-electric point, peptide net charge, extinction coefficient and estimated solubility) (<https://pepcalc.com/>) [44].

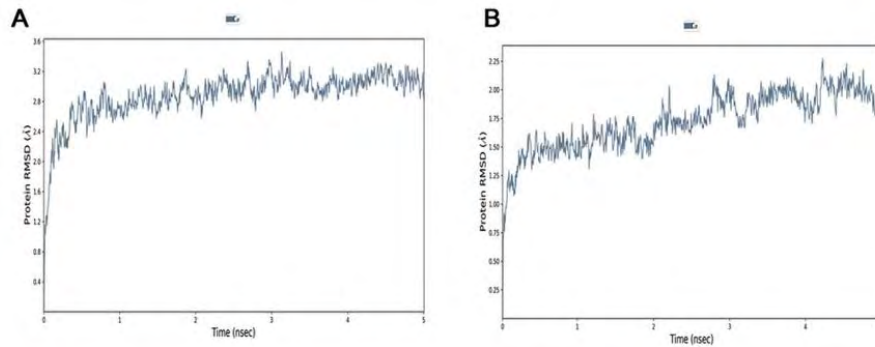
## 3. Results

The designed peptides (20 amino acids), showed a great impact on the binding of the human CD Markers to SARS-CoV-2 RBD epitopes. The combination of the important residues binding to specific epitopes, which formed scaffold pairs. This last was subsequently remodeled and refined into continuous peptides and one peptide was selected for further biochemical validation. The detail information methods provide for these selected peptides workflow is shown in Figure 1.

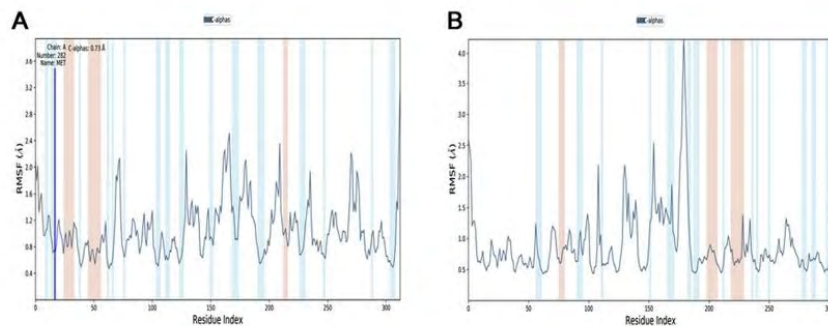


**Figure 3.** Comparisons of binding free energies of Hydrogen bond residues of the CD299 receptor inside SARS-CoV-2 RBD and SARS-CoV RBD throughout the Docking using MM-GBSA HawkDock server.





**Figure 4.** Analysis of all backbone atom root mean square deviations (RMSD) values of RBD-CoV-2/CD299 and RBD-CoV/CD299 complex (A and B) respectively over 5 ns simulation time.



**Figure 5.** Root mean square fluctuations (RMSF) of CD299/SARS-CoV-2 RBD and CD299/SARS-CoV-2 RBD complex (A and B) respectively, were computed and plotted of each residue for the RBD domain (residues 130–300) and the CD299 (residues 0–129).

### 3.1. Molecular dockings studies of CD147, CD209 and CD299 against the RBD of SARS-CoV-2 and SARS-CoV-1

The molecular docking of SARS-CoV-1 and SARS-CoV-2 RBD against CD147, CD209, CD299 and ACE2 was performed using ZDOCK server. The top ranked 500 models were clustered into six groups with slightly different conformations. The binding affinity and MM-GBSA binding free energy of the docked complexes were calculated using PRODIGY and HawkDock webservers, respectively (Table 1).

The complex of SARS-CoV-2 RBD with CD299 showed  $\Delta G$  value of  $-19.3$  kcal mol $^{-1}$  and a dissociation constant of  $7.3E-15$  which are higher than those observed in the complexes with CD147, CD209 and ACE2. Moreover, the CD299 models presented a higher number of H-bonds in complex with SARS-CoV-2 RBD compared to SARS-CoV-1 RBD and mannose, suggests that this is a weak binding (Table 1, Figure 2).

SARS-CoV-2 interacting with human CD299 were analyzed. The binding interaction between the RBD viral spikes is completely similar. The main interaction region of SARS-CoV-2 RBD with the CD299 receptor comprised 28 residues (Table 2), while only 18 residues were found in the interface of the SARS-CoV RBD with CD299. In addition, as mentioned above, the existence of the two loops in the binding domain of CD299 is likely to induce a stabilizing effect on its interaction with the RBD of SARS-CoV-2.

Figure 2 strongly suggest that these loops may enhance the electrostatic interactions between the SARS-CoV-2 RBD and the CD299 receptor. The residues present in these loops and showing direct interaction with the SARS-CoV-2 RBD are Arg266, Cys268, Pro269, Asp271, Thr273,

Phe274, Phe275, Gln276, Ser285, Ala309, Glu310, Gln312, Asn313, Phe314, Leu317, Gln318, Arg321, Pro349, Ser350, Gln352, Arg353, Tyr354, Trp355, Asn356, Ser357, Ser374, Gly375, Trp376. As can be seen in Table 2, the higher number of protein-protein contacts and the longer loops resulted in an increase in the binding affinities in SARS-CoV-2 RBD in comparison with SARS-CoV (Table 1). Accordingly, these loops could have a main role in determining the host receptor specificity for the SARS-CoV-2 spike protein. These key residues of the CD299 receptor were used as reference to design new peptide inhibitors that could impair the viral attachment to the host cells.

### 3.2. Comparison of the binding free energies hydrogen bond residues using MM-GBSA of the best human CD target between SARS-CoV-2 and SARS-CoV RBD

Table 1 shows a comparison between the free energies of SARS-CoV-2 RBD and SARS-CoV with CD marker CD299, CD147, CD209, and ACE2 using the MM-GBSA program in HawkDock server. The free energies value confirms the results of binding energy calculated with the PRODIGY web server. The complex RBD of SARS-CoV-2/CD299 shows a importantly low binding free energy. than that of SARS-CoV  $-69.49$  and  $-63.89$  kcal/mol, also compared to other CD marker and ACE2. The total of Hydrogen bond residues binding free energy that are  $< -1$  kcal/mol are 15 residues in complex SARS-CoV-2/CD299, compared with residues present in the SARS-CoV RBD/CD299 with 11 residues (Figure 3). These residues are worth considering as critical residues for the binding of RBD to a CD299.





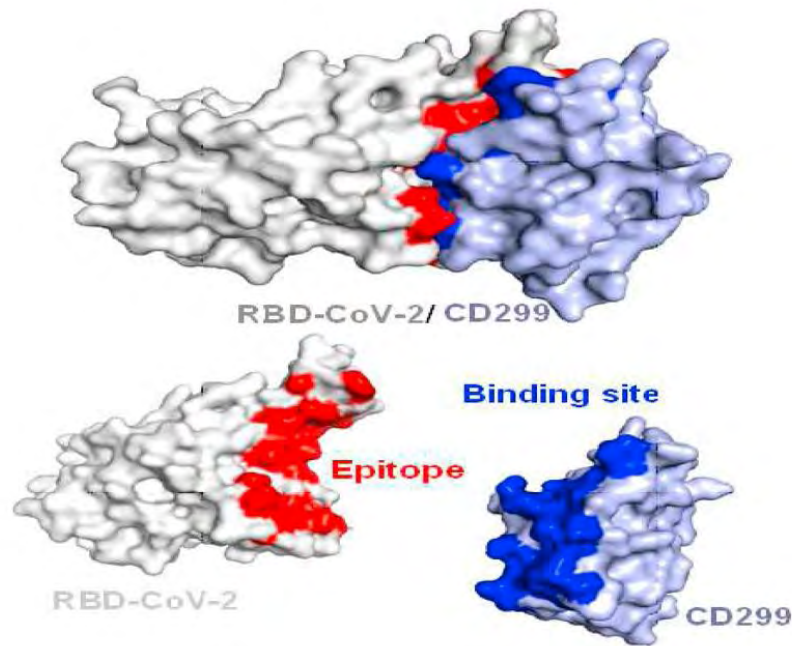


Figure 7. Docking models of SARS-CoV-2 RBD epitopes/CD299, the binding site epitopes SARS-CoV-2 RBD surface indicated by red color, and the binding site in CD299 indicated by blue color.

Plot of root-mean-square deviation (RMSD) were produced by molecular dynamics (MD) of the complexes. The backbone RMSD value of the two complexes were stable at  $< 3 \text{ \AA}$  after a gradual rise around of 5 ns (Figure 4).

The RBD-CoV-2/CD299 complex rose from 2.4 to 2.8  $\text{\AA}$  around 5 ns and then remained stable where a small rise was observed. Whereas, the RBD-CoV/CD299 complex rose from 1.5 to 1.75  $\text{\AA}$  then was stable during the simulated period. Stability in 5 ns simulation with average backbone RMSD of 0.4 and 0.25 nm, respectively. However, the RMSD showed the RBD SARS-CoV complex is less stable, it experienced a gradual increase in the RMSD value with fluctuations, stabilizing at an average of 0.25 nm. To investigate Protein flexibility that helped assessed of residues were involved of the complex, the  $C\alpha$  root mean square fluctuation ( $C\alpha$  RMSF) was calculated. Wherefore, the authors have examined the conformational flexibility of the RBD domains for SARS-CoV-2 and SARS-CoV with CD299 by analyzing their movement data. Figure 5 shows the  $C\alpha$  RMSF values for the complexes SARS-CoV-2 RBD/CD299 and SARS-CoV RBD/

CD299 over a period of 5 ns simulations have been conducted. As seen in this figure that the RMSF of SARS-CoV-2 RBD have generally increased over time than these of SARS-CoV with the CD299 receptor. On the other hand, the SARS-CoV-2 RBD is very extremely flexible than the SARS-CoV RBD. Therefore, most residues with a high  $C\alpha$  RMSF (Figure 5(A)) compared to SARS-CoV, especially higher in the loop near the binding site (470–490) (Figure 5(B)).

To more comprehend the impact of this higher flexibility of SARS-CoV-2 compared to SARS-CoV for CD299, the correlation between modification of residues in sequence and secondary structure of proteins motions of RBD domain of SARS-CoV-2 and SARS-CoV was analyzed during different simulation. Figure 6 provides a description about the Multiple Sequence Alignment and the secondary structure throughout the simulation. As shown in the results, the RBD domain shares most residues with 71.11% identity between SARS-CoV-2 and SARS-CoV. Detailed sequence alignment showed that SARS-CoV-2 present significant sequence variation at various regions with the human SARS

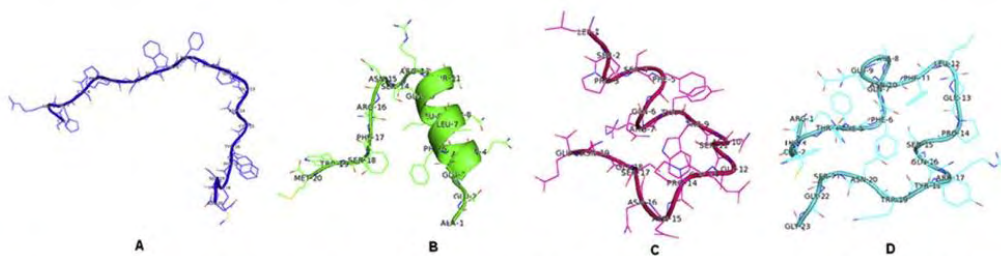
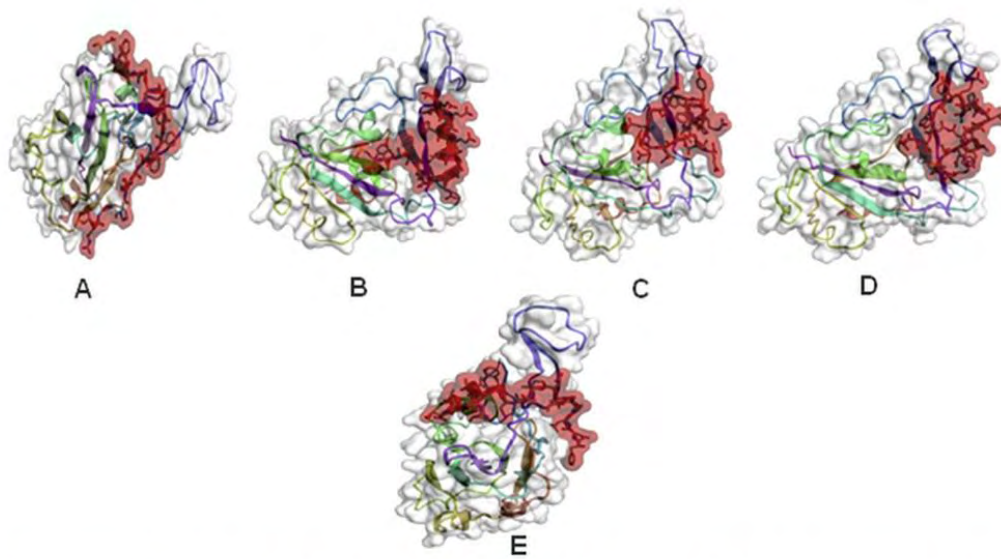


Figure 8. Tertiary structure of designed peptides. The cartoon and line representation of the peptidomimetics structures (A: peptide P1, B: peptide P2, C: peptide P3, D: peptide P4).

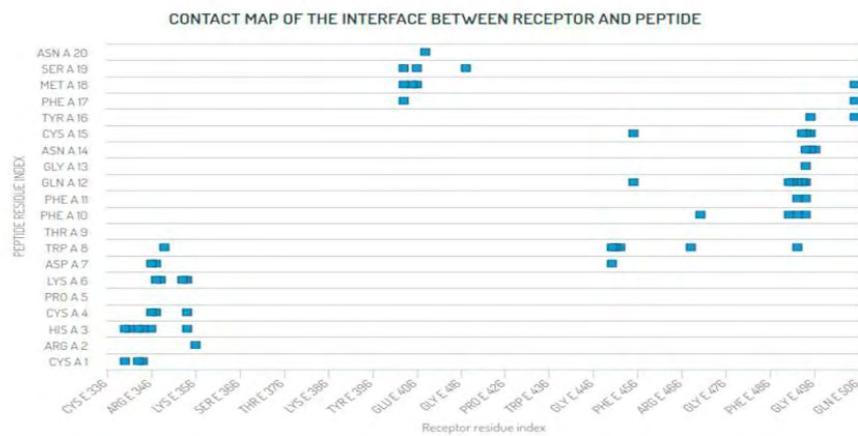




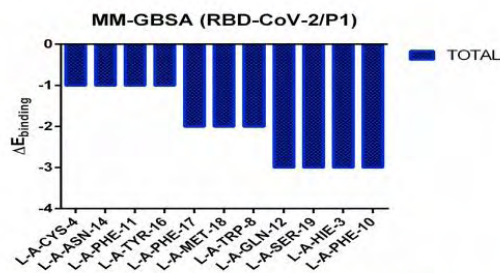
**Figure 9.** The binding modes of complex SARS-CoV-2 RBD/New designed peptides P1,P2, P3, P4 and anti-SARS-CoV. The New designed peptides binding site of the domain RBD is shaded in red and RBD SARS-CoV-2 are shaded in gray using CABS-dock. (A:P1/CD299, B: P2/CD299, C: P3/CD299, D: P4/CD299 E: Inhibitor SARS-CoV/CD299).

**Table 3.** The binding affinity and dissociation constant of complexes RBD-CoV-2/New Peptides design (20 amino acids) (P1, P2, P3 and P4) and peptide inhibitor SARS-CoV as control. Relative binding free energies of complexes estimated using MM-GBSA module in HawkDock and PRODIGY servers.

Peptides	Sequences	SARS-CoV-2		
		$\Delta G$ (kcal mol <sup>-1</sup> )	Kd (M) at 25.0 °C	MM-GBSA $\Delta E_{binding}$
P1(20)	CRHCPKDWTFFQGNCFMNSN	-12.9	3.8E-10	-62.83
P2(20)	AEEQNFLQLQTSRSNRFSWM	-10.1	4.2E-08	-27.92
P3(20)	LSPSFQRYWNSGEPNNSGNE	-9.0	2.4E-07	-34.25
P4 (23)	RCWTFQAEFLQPSQRYWNSGG	-11.2	5.9E-09	-47.07
Inhibitor SARS-CoV (20)	FKLPLGINITNFRAILTAFS	-10.4	2.2E-08	-41.95



**Figure 10.** Contact map of the interface between SARS-CoV-2 RBD and the best Peptide score affinity throughout the Docking using CABS-dock server.



**Figure 11.** Per-residue binding free energy decomposition (kcal/mol) of complexes RBD-CoV-2/CD299 and RBD-CoV-2/CD299 throughout the Docking using MM-GBSA module in HawkDock server.

coronavirus including the N-terminal region and the middle part of N protein (Figure 6(A)). Moreover, the authors suggest that this variation can change the structural conformation and secondary structure by increasing the flexibility of the loops. As a result, the secondary structure analysis of all residues in both SARS-CoV-2 and SARS-CoV is necessary (Figure 6(B)(C)). The residues located in (353–356) and (403–408) in RBD SARS-CoV, showed various poses during the simulation in different systems. In these residues, the alpha-helices conformation change to beta-strands in RBD SARS-CoV continuously (Figure 6(C)). However, this confirmation seemed to increase the flexibility of the residues located in the loops that may respond to high binding of Covid-19 in the active site of host cells, as showed with CD299 specific target protein receptor.

**3.4. Docking of new designed specific peptides inhibitors at domain RBD of spike SARS-CoV-2**

Four peptides namely, P1, P2, P3, and P4 were designed based on the selection of some residues of hydrogen bond of CD299 receptor domain binding site with RBD SARS-CoV-2 (Figure 7).

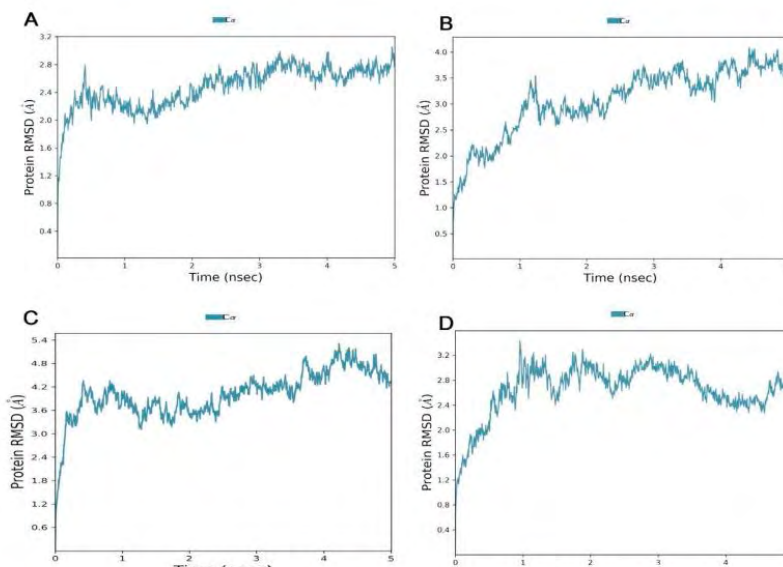
Each protein amino acid resulted in the generation of 4 peptidomimetics among the four designed peptides. Docking analysis was dedicated to predict the binding modes and affinity of these peptidomimetics on the RBD SARS-CoV-2 envelope protein. The all peptidomimetics tertiary structures predictions (Figure 8) were generated using CABS server (Figure 9) and were further minimized.

All peptidomimetics are bound specifically to the surface of the binding site CD299, the peptidomimetic P1 is located near the RBD SARS-CoV-2 binding site. The peptide-protein binding energy was determined in PRODIGY web server, for P1, P2, P3, P4 and using anti viral SARS-CoV as a control, range from –12.9 kcal/mol to –10.1 kcal/mol, –9.0 kcal/mol, –11.2 kcal/mol and –10.4 kcal/mol, respectively (Figure 9(A), (B), (C), (D), (E)) and (Table 3).

The lowest binding affinity related that the H-bonds was stronger between peptide P1 and the RBD CoV-2 protein (Table 3). Figure 10 shows the contact map of the interface between SARS-CoV-2 RBD and the best Peptide P1 score affinity throughout the Docking using the CABS-dock server. However, to confirm these results obtained from the docking studies, the free energy calculation of binding was determined using MM-GBSA program, HawkDock server and molecular dynamics simulations.

**3.5. Binding free energies and hydrogen bond residues analysis using MM-GBSA of new designed peptides inhibitors with RBD spike SARS-CoV-2**

The MM-PBSA method estimates the binding free energy indicate that P1 binds more tightly to the protein than P2, P3, P4 and anti-SARS-CoV with values of –62.83, –27.92, –34.25, –47.95 and –19.51 kcal/mol respectively (Table 3). This is in accordance with the binding affinity interaction findings where the P1 is elevated than other peptides (Table 3). To find the essential amino acids that help to the binding affinity, decomposition energy was computed for the P1-bound complex. Figure 11 illustrates the decomposed energies per residue basis for the peptide P1. Phenylalanine contribution binding in P1 is more pronounced followed by Histidine and Serine. All of the



**Figure 12.** Analysis of all backbone atom root mean square deviations (RMSD) of all new peptides inhibitors designs over 5 ns simulation time with RBD domain SARS-CoV-2 over 5 ns. A: peptide P1, B: peptide P2, C: peptide P3 D: peptide P4.

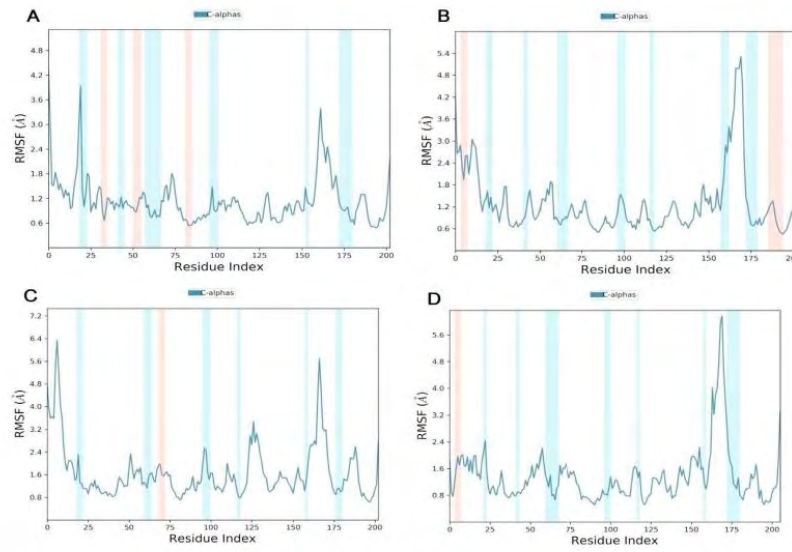


Figure 13. Root mean square fluctuations (RMSF) of all new peptides inhibitors designs calculated for every residue with RBD domain SARS-CoV-2 over 5 ns. Calculated and plotted for each residue of the RBD domain (residue 20–200) and the peptides (residue 0–19). A: peptide P1, B: peptide P2, C: peptide P3 D: peptide P4.

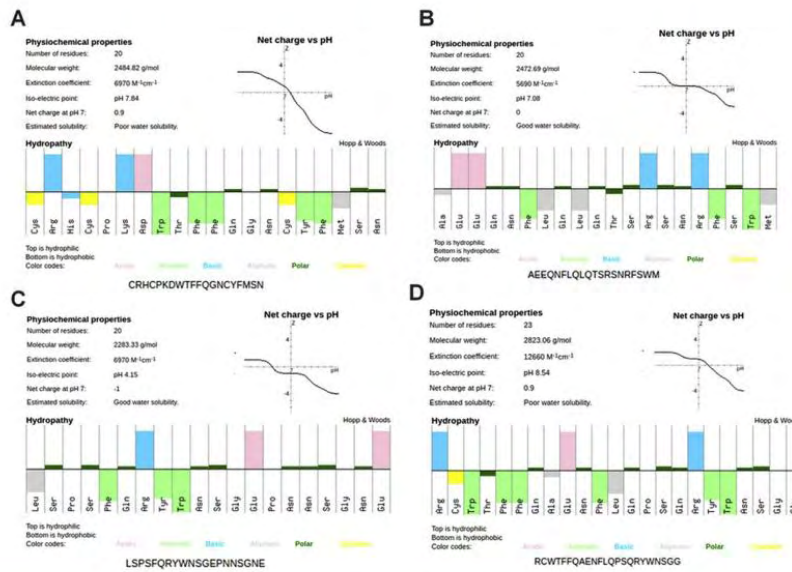


Figure 14. Analysis of molecular characteristics of the new mimetic peptides design. The PepCalc program was used to analyse the physico-chemical characteristics of the peptide sequences, showing differences in amphipathic properties (top: hydrophilic; bottom: hydrophobic). (A: peptide P1, B: peptide P2, C: peptide P3, D: peptide P4).

aromatic amino acids in P1 participated in binding to specific epitope of RBD SARS-CoV-2 and provided favorable interaction energy. This implicating that P1 has most residues which contribute in maintaining the peptide at the binding site by favorable peptide-protein interactions.

3.6. Molecular dynamics simulations of new designed peptides inhibitor with RBD SARS-CoV-2: system stability and flexibility

Root mean square deviations RMSD is commonly used to determine the structural stability of all C $\alpha$ -atoms of a system. Equilibrium vibration



and low fluctuation of RMSD were noticed in the P1/SARS-CoV-2 complex compared to other peptides complex that than rose from P1 (2.4–2.8), showing that the previous complex was greater stable at  $< 3 \text{ \AA}$  and induced smaller conformational changes longer time scales (Figure 12).

In addition, RMSF or root mean square fluctuation was also calculated and a two-dimensional plot was generated (Figure 13). RMSF is important to define the flexible areas as of a system. Overall RMSF shows values were higher for P1/SARS-CoV-2 RBD complex structures (Figure 13(A)) compared to other Peptides/SARS-CoV-2 RBD complex, indicating that SARS-CoV-2 RBD is stable overall, but fluctuate highly in residues 476 to 478, as compared to other peptides. It is obviously appeared that P1 is near to the protein binding site and remain at the site during the MD, as compared to other peptides. This is also in accordance with the RMSD calculation, where little change deviations was observed for domain RBD in complex with P1. The residues PHE, THR and ARG of all peptides show elevated fluctuations. The binding site of the domain RBD (residue 476 to 490) shows an important greatly RMSF.

### 3.7. Properties of new peptidomimetics

To study the structure of new peptidomimetics, the PepCalc program was used to analyze their theoretical physico-chemical characteristics of each amino acid residue. The data including the key properties of these peptides, which showed that all peptides were expected to have a satisfactory solubility (Figure 14). All peptides positively and negatively charged and hydrophobic amino acids and two peptides P2 and P3 are better dissolved and P1 had more hydrophobic groups this last exhibited a positive charge compared with the other peptides (Figure 14(A)).

## 4. Discussion

The receptor of immune response like CD marker as a target is an important point to understand the uncontrolled inflammation phenomenon deliver to multi-organ harm induce a deficiency in major of organs particularly of the cardiac, renal systems, and hepatic induced by Covid-19. In this study, we were able to screen *in silico* CD299, CD209, CD147, and ACE2 which are associated with various Severe acute respiratory distress syndrome (SARS) and analyze the mode of binding with RBD SARS-CoV-2 compared to RBD SARS-CoV to look for key structural effects offered by the mutation, and several amino acid substitutions in RBD were identified in the SARS-CoV-2 based on computational analysis. On the results obtained from the *in-silico* predictors when comparing the binding affinity using molecular docking between CD299, CD209, CD147, ACE2 receptors and RBD SARS-CoV-2, SARS-CoV the results showed the CD299 more highly than other CD makers, in binding affinity and with SARS-CoV-2 and confirmed by binding free energy calculation MM-GBSA and H-bond residues. In the present study, the last results already were assured in previous research with Severe acute respiratory distress syndrome (SARS). This is followed by a late stimulation of respiratory DCs to the lymph nodes chain with an insufficient virus-specific T cell response [45]. In patients infected with SARS-CoV-2, there is a start of lymphopenia, and the lymphocyte count is an estimate of disease danger and death-rate [46]. Lymphocyte levels recover with viral removal and the control of the disease, along with adaptive immune cells (CD3+ T cells) exist particularly significant [47]. Immune insufficiency was elucidated by a viral infection and T-cell interaction with the respiratory DCs. However, research studies have demonstrated that the current pandemic can be explained in part by the interaction between SARS-CoV-2 spike protein S, the ACE2/L-SIGN/CD209 receptor on type II alveolar cell of the lung, and the DC-SIGN receptor on the respiratory DCs and related endothelial cells [48]. The infection of DC by SARS-CoV-2 can certainly explain the abundant distal immunopathology observed in COVID-19. The MD simulation was thus conducted to reveal the dynamic properties of mutant protein and substitution in SARS-CoV-2

were the initial characterization by RMSD of complexes SARS-CoV-2/CD299 and SARS-CoV/CD299 showed that mutations caused various pathways during simulation. Consistent with the RMSD profile, also confirmed the mutation-induced a raise conformational flexibility and stability in RBD SARS-CoV-2 compared to RBD SARS-CoV in binding to CD299 receptor. In dynamics status, protein exposes several conformational changes for a specific function, where correlative motions provide a pivotal role in recognizing and binding of several biological macromolecules and this motion is usually perturbed by the mutation [49]. In the present case, the consequences of mutation were confirmed by RMSF analysis, where the effect was revealed in the regions of the loop of CoV-2. So, the variation can modify the structural conformation and the secondary structure which increased the flexibility of the loops, therefore the secondary structure of the RBD SARS-CoV-2 complex was analyzed compared to the RBD SARS-CoV. As a result, we observed the alpha-helices conformation change to beta-strands in RBD SARS-CoV-2 over time. However, this confirmation seemed to increase the flexibility of the residues located of the loops that may respond to the high binding of Covid-19 of the active site of hot cells as CD299 specific target protein receptor. Various aromatic amino acids modulate SARS-CoV-2 highly, which creates an entry portal. These new connection sequences provided the possibility for the design of multifunctional and specific peptides inhibitors. The development of mimicking peptides inhibitors from this key amino acid involved in the interaction of CD299 receptor and RBD SARS-CoV-2 is indispensable. Molecular Docking which predicts the binding affinity of new peptidomimetics designs with RBD SARS-CoV-2, the total analysis results of this research showed all peptides have a good has good score affinity but one peptide P1 bind to RBD SARS-CoV-2 better than other peptides and anti-CoV as control this confirmed by simulation analysis (RMSD, RMSF) and free energy MM-GBSA. Besides, the increase of binding in RBD SARS-CoV-2/Peptides might affect RBD binding properties as well as selectivity. Overall, the P1 is characterized by the hydrophobic residues binding site exhibits a wide range of specificity. Accordingly, the phenolic compounds are preferred by RBD SARS-CoV-2 (Figure 14,(A)), where the mutation L486F in the loop of residues 470–490 near the binding site than in other regions of RBD SARS-CoV-2, disorders the plasticity of target binding loops and increased the attached at binding the region, affecting the immune process and thus contributes to antiviral progression. Unexpectedly, the P1 peptide was highly anti-SARS-CoV-2. Its sequence analysis showed that small fragments of 20 aa in length of the (CRHCPKDWTFQGNCFMNS) has homologies only to specific C-type lectin domain (CTILD) when using query by BLAST-p in NCBI database this result are taking advantage of this potential. The peptide itself can also be used as an agent to inhibit virus infection if it can provide a spatial block between S protein and dendritic cells (DC) remains of special interest. With likely effect of aromatic rings of flight mutations by the coronavirus, a question is whether the impact of phenylalanine (F) is important; the hypothesis of this work is that mimic peptide P1 (CRHCPKDWTFQGNCFMNS) has 3 phenylalanine residues (F) in sequence and all interacted with RBD domain in CoV-2 that is confirmed by Hydrogen bond residues calculation by MM-GBSA (Figure 11). There appear to be a requirement to have one aromatic ring existing in the element proposes a powerful hydrophobic feature of the pharmacophore at that point.

Besides, many drugs include at least one aromatic ring and this is nearly surely for they can compose especially strong stacking interactions in the binding site [50]. In this paper analysis of protein structure to characterize binding and determine the strengths of protein-protein interactions is common in computational approach and that aromatic interactions are both strong and need special attention. However, though all new peptides design presents good results, the efficacy observed for these hydrophobic residues can play a key role in the RBD SARS-CoV-2 disruption activity of increasingly, the proof is preceding that amino acid composition and physicochemical properties have a better impact on antiviral activity.



## 5. Conclusion

The computational analysis approach through protein-protein docking and MD simulations appears to be an essential tool to identify specific CD marker epitopes such as for SARS-CoV-2. Indeed, the C-type lectin family DC-SIGN2 (CD299) receptor represents a powerful target on the RBD spike protein of SARS-CoV-2. Development of synthetic mimic peptides de-ri-ved from CD299 receptor should lead to discovery a novel group of therapeutic and vaccine for SARS-CoV-2.

## Declarations

### Author contribution statement

J. Akachar: Conceived and designed the experiments; Performed the experiments; Wrote the paper.

E.M. Bouricha, M. Hakmi and L. Belyamani: Analyzed and interpreted the data.

R. El Jaoudi and A. Ibrahim: Conceived and designed the experiments; Analyzed and interpreted the data.

### Funding statement

This work was supported by National Funding from the Moroccan Ministry of Higher Education and Scientific Research (Covid-19 Program) to AI and a grant to AI from Institute of Cancer Research and the PPR-1 program to AI.

### Declaration of interests statement

The authors declare no conflict of interest.

### Additional information

No additional information is available for this paper.

## References

- Zhou, X.L., Yang, X.G., Wang, et al., A pneumonia outbreak associated with a new coronavirus of probable bat origin, *Nature* 579 (7798) (2020) 270–273.
- Munster, M., Koopmans, N., van Doremalen, D., van Riel, E., de Wit, A novel coronavirus emerging in China—key questions for impact assessment, *N. Engl. J. Med.* 382 (8) (2020) 692–694.
- Shiva Raësi D, Seyed Ali HZ, Zahra G, Sepideh A. 2020. Covid-19 (SARS-CoV-2) VS Sars-CoV; Summary of all things that healthcare providers should know. *Am. J. Biomed. Sci. Res.* 2020 ; 376-382. .
- C.S.G. of the International, The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2, *Nat. Microbiol.* 5 (4) (2020) 536.
- Moore, T., Dorfman, W. Li, et al., Retroviruses pseudotyped with the severe acute respiratory syndrome coronavirus Spike protein efficiently infect cells expressing angiotensin-converting enzyme 2, *J. Virol.* 78 (19) (2004) 10628–10635.
- Hoffmann, H., Kleine-Weber, S., Schroeder, N., Krüger, T., Herrler, S., Erichsen, M.A., Müller, SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor, *Cell* (2020).
- Grifoni, D., Weiskopf, S.I., Ramirez, J., Mateus, J.M., Dan, C.R., Moderbacher, D., Murrain, Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals, *Cell* (2020).
- Zhang, X., Zhou, Y., Qiu, F., Feng, J., Feng, Y., Jia, S., Wang, Clinical characteristics of 82 death cases with COVID-19, *MedRxiv* (2020).
- Gutiérrez, J., Beckford, H., Alachkar, Deciphering the TCR repertoire to solve the COVID-19 mystery, *Trends Pharmacol. Sci.* (2020).
- Tang-Huau, P., Gueguen, C., Goudot, M., Durand, M., Bohec, S., Baulande, E., Segura, Human in vivo-generated monocyte-derived dendritic cells and macrophages cross-present antigens through a vacuolar pathway, *Nat. Commun.* 9 (1) (2018) 1–12.
- Sungnak, N., Huang, C., Bécavin, M., Berg, R., Queen, M., Litvinukova, K.B., Worlock, SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes, *Nat. Med.* 26 (5) (2020) 681–687.
- Marzi, T., Gramberg, G., Simmons, P., Moller, A.J., Rennekamp, M., Krumbiegel, M., Geier, J., Eiseemann, N., Turza, B., Saunier, A., Steinkasserer, S., Becker, P., Bates, H., Hofmann, S., Pohlmann, DC-SIGN and DC-SIGNR interact with the glycoprotein of Marburg virus and the S protein of severe acute respiratory syndrome coronavirus, *J. Virol.* 78 (2004) 12090–12095.
- Z.Y. Yang, Y. Huang, L. Ganesh, K. Leung, W.P. Kong, O. Schwartz, K. Subbarao, G.J. Nabel, pH-dependent entry of severe acute respiratory syndrome coronavirus is mediated by the spike glycoprotein and enhanced by dendritic cell transfer through DC-SIGN, *J. Virol.* 78 (2004) 5642–5650.
- S.A. Jeffers, S.M. Tusell, L. Gillim-Ross, E.M. Hemmila, J.E. Achenbach, G.J. Babcock, W.D. Thomas Jr., L.B. Thackray, M.D. Young, R.J. Mason, D.M. Ambrosino, D.E. Wentworth, J.C. Demartini, K.V. Holmes, CD209L (L-SIGN) is a receptor for severe acute respiratory syndrome coronavirus, *J. Virol.* 101 (2004) 15748–15753.
- S.A. Jeffers, S.M. Tusell, L. Gillim-Ross, E.M. Hemmila, J.E. Achenbach, G.J. Babcock, D.M. Ambrosino, CD209L (L-SIGN) is a receptor for severe acute respiratory syndrome coronavirus, *Proc. Natl. Acad. Sci. Unit. States Am.* 101 (44) (2004) 15748–15753.
- R.C. da Silva, L. Segat, S. Crovella, Role of DC-SIGN and L-SIGN receptors in HIV-1 vertical transmission, *Hum. Immunol.* 72 (4) (2011) 305–311.
- S. Pohlmann, J. Zhang, F. Baribaud, Z. Chen, G.J. Leslie, G. Lin, et al., Hepatitis C virus glycoproteins interact with DC-SIGN and DC-SIGNR, *J. Virol.* 77 (2003) 4070–4080.
- J.L. Miller, B.J. Dewet, L. Martinez-Pomares, C.M. Radcliffe, R.A. Dwek, P.M. Rudd, S. Gordon, The mannose receptor mediates dengue virus infection of macrophages, *PLoS Pathog.* 4 (2) (2008) e17.
- Alvarez, F., Lasala, J., Carrillo, O., Muniz, A.L., Corbi, R., Delgado, C-type lectins DC-SIGN and L-SIGN mediate cellular entry by Ebola virus in cis and in trans, *J. Virol.* 76 (13) (2002) 6841–6844.
- Halary, A., Amara, H., Lortat-Jacob, M., Messerle, T., Delaunay, C., Houles, F., Fieschi, F., Arenzana-Seisdedos, J.F., Moreau, J., Dechanet-Merville, Human cytomegalovirus binding to DC-SIGN is required for dendritic cell infection and target cell trans-infection, *Immunity* 17 (2002) 653–664.
- Crosnier, L.Y., Bustamante, S.J., Bartholdson, et al., Basigin is a receptor essential for erythrocyte invasion by *Plasmodium falciparum*, *Nature* 480 (7378) (2011) 534–537.
- Moheimani, J., Koops, T., Williams, A.T., Reid, P.M., Hansbro, P.A., Wark, D.A., Knight, Influenza A virus infection dysregulates the expression of microRNA-22 and its targets; CD147 and HDAC4, in epithelium of asthmatics, *Respir. Res.* 19 (1) (2018) 145.
- W. Bao, D. Min, S.M. Twigg, N.A. Shackel, F.J. Warner, D.K. Yue, S.V. McLennan, Monocyte CD147 is induced by advanced glycation end products and high glucose concentration: possible role in diabetic complications, *Am. J. Physiol. Cell Physiol.* 299 (5) (2010) C1212–C1219.
- Raony, C.S. de Figueiredo, Retinal outcomes of COVID-19: possible role of CD147 and cytokine storm in infected patients with diabetes mellitus, *Diabetes Res. Clin. Pract.* (2020).
- G. Castel, M. Chtouy, B. Heyd, et al., Phage display of combinatorial peptide libraries: application to antiviral research, *Molecules* 16 (5) (2011) 3499–3518.
- R.E. Hancock, R. Lehrer, Cationic peptides: a new source of antibiotics, *Trends Biotechnol.* 16 (2) (1998) 82–88.
- Y.M. Hrobowski, R.F. Garry, S.F. Michael, Peptide inhibitors of dengue virus and West Nile virus infectivity, *Virol. J.* 2 (1) (2005) 49.
- W.C. Wimley, S.H. White, Experimentally determined hydrophobicity scale for proteins at membrane interfaces, *Nat. Struct. Biol.* 3 (10) (1996) 842–848.
- Lalezari, K., Henry, M., O’Hearn, J.S., Montaner, P.J., Pillero, B., Trotter, J., Chung, Enfuvirtide, an HIV-1 fusion inhibitor, for drug-resistant HIV infection in North and South America, *N. Engl. J. Med.* 348 (22) (2003) 2175–2185.
- Champagne, A., Shishido, M.J., Root, Interactions of HIV-1 inhibitory peptide T20 with the gp41 N-HR coiled coil, *J. Biol. Chem.* 284 (6) (2009) 3619–3627.
- Lu, Q., Wang, G.F., Gao, Bat-to-human: spike features determining ‘host jump’ of coronaviruses SARS-CoV, MERS-CoV, and beyond, *Trends Microbiol.* 23 (2015) 468–478.
- Li, Receptor recognition and cross-species infections of SARS coronavirus, *Antivir. Res.* 100 (1) (2013) 246–254.
- P.K. Panda, M.N. Arul, P. Patel, S.K. Verma, W. Luo, H.G. Rubahn, R. Ahuja, Structure-based drug designing and immunoinformatics approach for SARS-CoV-2, *Sci. Adv.* 6 (28) (2020), eabb8097.
- Chen, L., Li, Z., Weng, ZDOCK: an initial-stage protein-docking algorithm, *Proteins: Struct. Funct. Bioinform.* 52 (1) (2003) 80–87.
- Chen, Z., Weng, Docking unbound proteins using shape complementarity, desolvation, and electrostatics, *Proteins: Struct. Funct. Bioinform.* 47 (3) (2002) 281–294.
- Katchalski-Katzir, I., Shariv, M., Eisenstein, A.A., Friesem, C., Aflalo, I.A., Vakser, Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques, *Proc. Natl. Acad. Sci. Unit. States Am.* 89 (6) (1992) 2195–2199.
- Huang, X., Zou, An iterative knowledge-based scoring function for protein-protein recognition, *Proteins: Struct. Funct. Bioinform.* 72 (2) (2008) 557–579.
- S. Schrodinger, The PyMOL Molecular Graphics System, 2014, Version 1.6.
- L.C. Xue, J.P. Rodrigues, P.L. Kastiris, A.M. Bonvin, A. Vangone, PRODIGY: a web server for predicting the binding affinity of protein-protein complexes, *Bioinformatics* 32 (23) (2016) 3676–3678.
- Weng, E., Wang, Z., Wang, H., Liu, F., Zhu, D., Li, T., Hou, HawkDock: a web server to predict and analyze the protein-protein complex based on computational docking and MM/GBSA, *Nucleic Acids Res.* 47 (W1) (2019) W322–W330.
- S. Release, 3: Desmond Molecular Dynamics System, DE Shaw research, New York, NY, Maestro-Desmond Interoperability Tools, Schrödinger, New York, NY, 2017, 2017.

- [42] J.D. Thompson, D.G. Higgins, T.J. Gibson, W. Clustal, Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Res.* 22 (22) (1994) 4673–4680.
- [43] M. Kurcinski, M. Jamroz, M. Blaszczyk, A. Kolinski, S. Kmiecik, CABS-dock web server for the flexible docking of peptides to proteins without prior knowledge of the binding site, *Nucleic Acids Res.* 43 (W1) (2015) W419–W424.
- [44] S. Lear, S.L. Cobb, Pep-Calc. com: a set of web utilities for the calculation of peptide and peptoid properties and automatic mass spectral peak assignment, *J. Comput. Aided Mol. Des.* 30 (3) (2016) 271–277.
- [45] H.K. Law, C.Y. Cheung, H.Y. Ng, S.F. Sia, Y.O. Chan, W. Luk, Y.L. Lau, Chemokine up-regulation in sars-coronavirus-infected, monocyte-derived human dendritic cells, *Blood* 106 (7) (2005) 2366–2374.
- [46] O. Tanriverdi, Lymphopenia that may develop in patients treated with temozolomide and immune control check-point inhibitor may be a high risk for mortality during the COVID-19 outbreak, *Med. Oncol.* (Northwood, London, England) 37 (6) (2020).
- [47] R. He, Z. Lu, L. Zhang, T. Fan, R. Xiong, X. Shen, Q. Geng, The clinical course and its correlated immune status in COVID-19 pneumonia, *J. Clin. Virol.* (2020) 104361.
- [48] A. Brufsky, M.T. Lotze, DC/L-SIGNs of hope in the COVID-19 pandemic, *J. Med. Virol.* (2020).
- [49] B.K. Ho, D.A. Agard, Probing the flexibility of large conformational changes in protein structures through local perturbations, *PLoS Comput. Biol.* 5 (4) (2009), e1000343.
- [50] M.D. Poléto, V.H. Rusu, B.I. Grisci, M. Dorn, R.D. Lins, H. Verli, Aromatic rings commonly used in medicinal chemistry: force fields comparison and interactions with water toward the design of new chemical entities, *Front. Pharmacol.* 9 (2018) 395.

# **CONCLUSION GENERALE ET PERSPECTIVES**

Une consommation excessif de lipides en particulier des Acides Gras Saturés (AGS) associé à un développement du tissu adipeux contribue à la prévalence de l'obésité. Une corrélation a d'ailleurs été établie entre la quantité de lipides consommés et la masse corporelle chez l'Homme (Brayet *et al.*, 2004) et chez la souris (Witetal., 2011). Elle est associée à des perturbations métaboliques telles que l'insulino-résistance, l'intolérance au glucose, des dyslipidémies, l'hypertension, qui constituent des facteurs de risque des maladies cardiovasculaires (MCV). Plusieurs données suggèrent l'implication du CD36 dans le transport et le captage des AGLC. Le fatty acid transporter (FAT) ou CD 36 est une glycoprotéine de 88 qui présente une structure en épingle à cheveux avec un large domaine extracellulaire comportant une poche hydrophobe pouvant accueillir jusqu'à 3 AGLC, ainsi que de nombreuses autres molécules comme les LDL oxydés, phospholipides anioniques ou la thrombospondine (Abumrad *et al.*, 1993). Plusieurs travaux ont été réalisés sur la génération de la structure tridimensionnelle de la protéine CD36 pour étudier son interaction avec des molécules antagonistes par les méthodes de la modélisation moléculaire par homologie et l'amarrage moléculaire. Ces publications ont permis d'obtenir seulement des informations nécessaires sur le mode de fixation de certains antagonistes (Tarhda *et al.*, 2015).

Mon sujet de recherche avait pour objectif le développement d'un workflow, par approche *in silico*, par le biais de la modélisation et de dynamique moléculaire pour l'identification de nouvelles molécules bioactives vis à vis du récepteur lipidique CD36 et du domaine RBD du SARS-CoV-2 en activant la voie de signalisation des acides gras et particulièrement ceux spécifiques au goût de gras. La première partie de notre travail a été consacrée pour effectuer des séries de simulations de dynamique moléculaire de la protéine CD36 de type sauvage et de type mutant K164A, en interaction avec un ou trois acides palmitiques (PLM) au maximum, afin de découvrir le mécanisme de fonctionnement du complexe CD36/Acides gras et de déterminer le rôle structural de cette mutation. À cette époque, une seule structure cristallographique du récepteur CD36 (5LGD) étaient disponibles sur le site de la PDB (Hsieh, et al. 2016). Cette structure nous a servi comme base afin de révéler le mode de transport des acides gras en temps réel. D'après Tarhda et al, (2015), l'inhibiteur SSO se fixe sur le couvercle de la CD36, au niveau duquel il y a la Lys-164, ce qui empêche la liaison des AGLCs et la LDLox sur leur site de reconnaissance. Nos résultats ont confirmés et complétés le rôle de la Lys164 du type sauvage, en induisant un changement de conformation pour former un tunnel d'internalisation assurant l'absorption des acides gras d'une part et celui de l'Ala164 de type mutant K164A CD36 influençant d'une manière importante sur le pliage, l'utilité, la solubilité et la forme de stabilité du CD36. Ces résultats ont été également approuvés par des études *in vitro* de Coburn et al, (2003) et Niladri S et al. (2008) qui ont montré que



la substitution de la lys 164 avec l'Ala entraîne une signalisation limitée. En outre, ces simulations ont montré que le résidu Lys164 du type sauvage a fourni la base structurelle de la formation d'une seule entrée au canal principal du CD36 pour la dissociation d'un seul acide palmitique et qu'il pourrait assurer de préférence le transport de deux acides palmitiques au maximum qui est n'est plus à démontrer dans l'ancien travail (Hsieh, et al. 2016) en révélant l'existence de deux entrées putatives de cette cavité centrale, dans lequel les acides gras se déplacent à travers la cavité d'après une coupe à travers une représentation de surface de CD36 montrant la cavité centrale du noyau occupée par deux acides palmitiques. Ce travail a donné lieu à une publication (Akachar et al. 2021). D'après Coburn et al, (2003) et Dramane et al., (2012), la fixation de l'acide gras sur la Lys164 au sommet de la protéine initie une cascade de signalisation parallèlement à l'entrée de AG et qui fait intervenir un flux de calcium qualifié SOC. Tous ces résultats nous suggèrent fortement que le calcium induit par CD36, pourrait l'être par l'intermédiaire d'une dimérisation de domaines transmembranaires TM1 et TM2 de la protéine CD36.

La deuxième partie de notre travail a été consacrée à l'analyse du mécanisme de la régulation allostérique (dimérisation) et l'activation de la signalisation CD36 d'une part, et à la conception et le criblage statique et dynamique d'analogues de composés phénoliques (Faux gras) pour mettre en place une nouvelle approche thérapeutique potentielle contre l'obésité. Les résultats de ce travail ont montré que l'acide palmitique induit une dimérisation contribuant à la communication allostérique entre le site K164 et les domaines TM1 et TM2. Gutmann et al. 2018 ont montré par microscopie électronique monoparticule que la signalisation transmembranaire par le récepteur RI de l'insuline est basée sur la dimérisation induite par le ligand avec un changement structural au sein du dimère récepteur. Ce réaménagement structural de l'ectodomaine se propage aux domaines transmembranaires, qui sont bien séparés dans la conformation inactive, mais viennent se rapprocher lors de la liaison à l'insuline, facilitant l'autophosphorylation des domaines kinases cytoplasmiques. En plus, une équipe de l'Université de Bourgogne a fourni la première preuve que les cellules gustatives CD36-positives jouent un rôle important dans la perception des lipides alimentaires chez la souris. Cette équipe a purifié les cellules de papilles gustatives exprimant la protéine CD36 pour étudier leur activité *in vitro* ainsi et a synthétisé des antagonistes sélectifs des récepteurs du goût des lipides afin de diminuer la sensibilité du goût du gras chez des sujets normaux ou obèses. D'après Shanmugamprema et al (2020), les antagonistes ou modificateurs de goût gras sont des composés synthétisés chimiquement ou purifiés à partir de matériaux biologiques et peuvent produire le principe du goût gras en se liant aux récepteurs du goût gras. Très récemment, Murtaza et al (2020) démontre que l'agoniste TUG891 se lie aux récepteurs linguaux CD36 et déclenche son axe langue-cerveau-intestin pour moduler la préférence gustative du gras. Tous ces résultats nous ont poussé fortement à la conception *in silico*, pour la première fois, de 15 molécules

et par la suite le criblage statique de ces molécules, en utilisant l'amarrage moléculaire. Notre résultat a confirmé un score élevé pour trois analogues (CIN4, COU5, CAF1). L'analogue CIN4 a montré, par la suite, par criblage dynamique, une dimérisation des 2 domaines TM1 et TM2 viennent se rapprochent lors de la liaison à CIN4, qui ressemble à l'acide palmitique. Les premiers résultats obtenus semblent très prometteurs. Un article décrivant ce travail est en cours de soumission.

Dans le cadre de l'étude du goût, récemment d'après Boscolo-Rizzo et al (2022), dans une série de 81 sujets testés positifs pour la variante SARS-CoV-2 Omicron, l'altération de l'odorat et du goût a été auto déclarée par 12 % et 23 % des patients, respectivement. Une analyse *in silico* du mécanisme structurel via le récepteur CD36 sera nécessaire. Le but de notre troisième étude était d'explorer par modélisation moléculaire et simulations dynamiques (MD) le dysfonctionnement gustatif de la chimio réception des acides gras spécifiquement le "goût de gras" des patients testés positifs au COVID-19. Les résultats de cette étude ont montré une affinité élevée et une stabilité du domaine RBD du SARS-CoV-2 non mutant au récepteur CD36 par rapport à celui du SRAS-CoV et du variant Omicron du SARS-CoV-2. L'étude des angles locaux du peptide de TM1 et de TM2 a révélé une intense déformation de ces peptides après fixation de SARS-CoV-2 entraînant un blocage de la signalisation du site de transport des acides gras. Tous ces résultats *in silico* ont confirmé ceux démontré expérimentalement par Boscolo-Rizzo et al (2022), En conclusion, le CD36 représente une cible puissante pour le développement de peptides épitopes dérivés, ce qui devrait conduire à la découverte et conception d'un nouveau groupe thérapeutiques ciblant le dysfonctionnement du goût par le Covid 19 et neutralisant le SRAS-CoV-2. Les résultats obtenus semblent très prometteurs. Un article décrivant ce travail est en cours de soumission.

Des travaux de recherche pour la mise en place de traitement de la dysfonction de goût par le Covid 19 semblent nécessaires. Dans ce contexte, les nouvelles techniques de modélisation moléculaire représentent un moyen indispensable pour identifier des épitopes et la conception de nouvelles molécules qui neutralise le SRAS-CoV-2. Dans notre quatrième étude, le criblage virtuel par amarrage protéine-protéine a été utilisé pour identifier les épitopes critiques pour trois cibles (CD147, CD209 et CD299). Les résultats des énergies libres de liaison de ces trois cibles ont montré une affinité élevée du domaine RBD du SARS-CoV-2 pour le récepteur CD299 qui est bien connu pour se lier à des virus comme le VIH-1 et aussi le SRAS-CoV (Van Breedam et al, 2014). En outre, ces résultats nous ont permis de faire la conception de quatre inhibiteurs peptidiques hypothétiques (P1, P2, P3 et P4). Après une étude d'amarrage et de dynamique moléculaire, ces peptides ont montré des caractéristiques de liaison conduisant à une stabilité favorables avec le domaine RBD du SARS-CoV-2. De plus, l'analyse des propriétés physico-chimiques a révélé que tous ces peptides ont une solubilité satisfaisante. D'après Jagat Narayan Shah et al (2022) un

candidat-médicament, RhACE2-APN01 basé sur le peptide hACE2 recombinant, est déjà entré dans les essais cliniques de phase II. Nos quatre peptides (P1, P2, P3 et P4) peuvent donc être considérées d'avantage comme des candidats potentiels dans les futures travaux de recherche de découverte de médicaments anti-SARS CoV-2. Ce travail a donné lieu à une publication (Akachar et al. 2020) et d'autre publication (Hoffmann et al 2021).

D'autres études sont nécessaires pour réaliser la synthèse chimique sur un grand nombre de molécules issues de notre criblage et par la suite réaliser les tests biologiques *in vitro* et *in vivo*. De plus, des travaux de recherche semblent également intéressants pour procéder à l'automatisation de nos protocoles de calculs utilisés, afin de les rendre facilement utilisables par d'autres chercheurs.

L'ensemble de ces travaux démontre que la modélisation et la dynamique moléculaire un outil qui semble contrôler l'optimisation de la découverte de molécules bioactives dans notre étude "faux gras" via le protéine CD36 qui est un récepteur aux AGLC.

Ces données mettent en lumière l'importance que pourrait avoir CD36 dans la mise en place des maladies liées à des excès de lipides alimentaires puisqu'il contrôle l'hypertriglycéridémie et l'endotoxémie post prandiale. Cette dernière est de plus en plus considérée comme étant à l'origine de l'inflammation bas bruit qui est associée à l'obésité, le diabète de type II et le syndrome métabolique. La langue via CD36 constituerait donc une cible thérapeutique pertinente ainsi ces fonctions semblent d'autant plus importantes que le régime contient une forte proportion d'AGLC insaturés. Une meilleure compréhension des modes de signalisation et de transport spécifiques des AGLC devraient permettre d'envisager des corrections nutritionnelles pour les sujets déficients en CD36 ou souffrant d'hypertriglycéridémie post prandiale, d'endotoxémie et éventuellement d'inflammation intestinale.

Notre expertise nous a mené de développer des protocoles de criblage *in silico* pour d'autre activité anti-bacérien, anti-cancer, anti-inflamatory, anti-diabétique, psychotropique de nouveaux molécules bioactives à prédire l'affinité d'un très grand nombre de molécules (collectées en chimiothèques) pour le site actif ciblé afin d'aboutir à une liste très réduite de molécules présentant l'activité souhaitée.

Dans ce contexte et en collaboration avec le laboratoire de chimie thérapeutique, Laboratoire de pharmacologie, Faculté de médecine et pharmacie, Laboratoire de chimie faculté de science Rabat, Laboratoire de chimie Faculté de médecine Sanaa E Yemen, dans ce travail est un projet développement de méthodes, d'analyse bioinformatique de structure protéiques et de conception de molécules bioactives.

-Un premier travail présente une étude le mécanisme d'action de l'activité cytotoxique par la technique de l'amarrage moléculaire et prédiction de les propriété physico-chimique de dérivés 3-

substitués de la 5,5-diphénylimidazolidine-2,4-dione in silico ciblant le récepteur dihydrofolate réductase humaine (DHFR).

-Un second travail concerne étude de caractérisation structurale et moléculaire, toxicité, activité par l'amarrage moléculaire d'un nouveau dérivé de la phénytoïne: 3-décy-5,5-diphénylimidazolidine-2,4-dione pour l'activité psychotrope ciblant deux récepteurs, le récepteur GABA-A et le canal Na<sup>+</sup>.

-Un troisième travail traite activité antibactérienne de dérivés de 1,5-benzodiazépine-2-one in silico utilisant le docking moléculaire ciblant le récepteur DNA gyrase

-Le travail 4 décrit activité antibactérienne de dérivés de 1,2,3-TRIAZOLYLMETHYL-QUINOXALINE in silico utilisant le docking moléculaire ciblant le récepteur DNA gyrase.

-Le travail 5 illustre l'études de activité analgésique et antiinflammatoire Quinazoline-4-one in silico utilisant le docking moléculaire ciblant isol'enzyme cyclooxygénaseCOX2.

- Le travail 5 L'étude de déplacement de la liaison du 99mTc-DTPA à l'albumine sérique humaine en présence de furosémide et de metformine en utilisant l'amarrage moléculaire.



## **REFERENCES BIBLIOGRAPHIQUES**

Abumrad NA, el-Maghrabi MR, Amri EZ, Lopez E, Grimaldi PA.(1993) Cloning of a rat adipocyte membrane protein implicated in binding or transport of long-chain fatty acids that is induced during preadipocyte differentiation. Homology with human CD36. *J Biol Chem*;268(24): 17665–8.

Armesilla AL, Calvo D, Vega MA.(1996) Structural and functional characterization of the human CD36 gene promoter: identification of a proximal PEBP2/CBF site. *J Biol Chem.*;271:7781–7.

Armesilla AL, Vega MA.(1994) Structural organization of the gene for human CD36 glycoprotein. *J Biol Chem.*;269:18985–91.

Amersilla AL, Calvo D, Vega MA. (1996). Structural and functional characterisation of the human CD36 gene promoter &ndash;Identification of a proximal PEBP2/CBF site. *J Biol Chem*; 271: 7781-7.

Asch A. (1996). To tell the truth : will the real CD36 please stand up ? *J Lab Clin Med*; 127: 321-5.

Atkinson, F., Cole, S., Green, C., & Van de Waterbeemd, H. (2002). Lipophilicity and other parameters affecting brain penetration. *Current Medicinal Chemistry-Central Nervous System Agents*, 2(3), 229-240.

Andricioaei, I., and M. Karplus. (2001). On the calculation of entropy from covariance matrices of the atomic fluctuations. *J. Chem. Phys.* 115:6289–6292.

Aparoy P, Reddy KK, Reddanna P. (2012) Structure and ligand based drug design strategies in the development of novel 5- LOX inhibitors. *Curr Med Chem.*;19(22):3763-78.

B. P. Welford (1962). “Note on a Method for Calculating Corrected Sums of Squares and Products.” *Technometrics* 4(3):419-420.

Brooks, C.; Case, D. (1993) Simulations of peptide conformational dynamics and thermodynamics. *Chemical Reviews*, 93, 2487-2502.

Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. (1983) CHARMM - A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry*, 4, 187-217.

Bazooyar, F.,Taherzadeh, M., Niklasson, C., Bolton, K., (2013). Molecular modeling of cellulose dissolution. *Journal of Computational and Theoretical Nanoscience*, 10(11), 2639-2646.

Baillie AGS, Coburn CT, Abumrad NA. (1996). Reversible binding of long-chain fatty acids to purified FAT, the adipose CD36 homolog. *J Membr Biol.*;153:75–81.

Ben-Shlomo H, Naor Z: (1991). Preferential release of catecholamine from permeabilized PC12 cells by  $\alpha$ - and  $\beta$ -type protein kinase C subspecies. *Biochem J* 280:65–69

Bray, George A, Sahaspor Paeratakul, et Barry M Popkin. (2004). « Dietary Fat and Obesity: A Review of Animal, Clinical and Epidemiological Studies ». *Physiology & Behavior* 83 (4): 549 55. doi:10.1016/j.physbeh.2004.08.039.

Baillie A, Coburn C, Abumrad N. (1996) Reversible binding of long-chain fatty acids to purified FAT, the adipose CD36 homolog. *J Membr Biol*;153:75–81.

C. Audouze, (2003) Vers une parallélisations par bandes en chimie quantique, Laboratoire de Mathématique, UMR CNRS 8628, Université Paris-Sud ,1-8.

Capener CE, Shrivastava IH, Ranatunga KM, Forrest LR, Smith GR, Sansom MS.(2000). Homology modeling and molecular dynamics simulation studies of an inward rectifier potassium channel. *Biophysical journal*;78(6):2929-42.

Cartoni C, Yasumatsu K, Ohkuri T, Shigemura N, Yoshida R, Godinot N, le Coutre J, Ninomiya Y, Damak S. (2010). Taste preference for fatty acids is mediated by GPR40 and GPR120. *J Neurosci.*;30:8376–82.

Cavasotto CN, Phatak SS. (2009) Homology modeling in drug discovery: current trends and applications. *Drug Discov Today*;14(13-14):676-83.

Contreras, J. M., Parrot, I., Sippl, W., Rival, Y. M., &Wermuth, C. G. (2001). Design, synthesis, and structure– activity relationships of a series of 3-[2-(1-benzylpiperidin-4-yl) ethylamino] pyridazine derivatives as acetylcholinesterase inhibitors. *Journal of medicinal chemistry*, 44(17), 2707-2718.

Contreras, J. M., Rival, Y. M., Chayer, S., Bourguignon, J. J., &Wermuth, C. G. (1999). Aminopyridazines as acetylcholinesterase inhibitors. *Journal of medicinal chemistry*, 42(4), 730-741.

Chevrier S., Zurbuchen Y., Cervia C., Adamo S., Raeber M.E., de Souza N., Sivapatham S., Jacobs A., Bachli E., Rudiger A., et al. (2021). A distinct innate immune signature marks progression from mild to severe COVID-19. *Cell Rep. Med.*;2:100166.

Chevrot, Michael, Arnaud Bernard, Déborah Ancel, Marjorie Buttet, Céline Martin, Souleymane Abdoul-Azize, Jean-François Merlin, et al. (2013). « Obesity Alters the Gustatory Perception of Lipids in the Mouse: Plausible Involvement of Lingual CD36 ». *Journal of Lipid Research* 54 (9): 2485-2494.

Chevrot, Michael, Patricia Passilly-Degrace, Déborah Ancel, Arnaud Bernard, Géraldine Enderli, Marlène Gomes, Isabelle Robin, et al.(2014). « Obesity Interferes with the Orosensory Detection of Long-Chain Fatty Acids in Humans ». *The American Journal of Clinical Nutrition* 99 (5): 975-983.

Canton J, Neculai D, Grinstein S. (2013). Scavenger receptors in homeostasis and immunity. *Nat Rev Immunol*;13:621–34.

Coburn C, Abumrad NA, editors. (2003) Structure function of CD36 and evidence for it's role in facilitating fatty acid transport. Wiley-VCH Publishers GMBH & Co.

Chow SL, Hollander D. (1979). Linoleic acid absorption in the unanesthetized rat: mechanism of transport and influence of luminal factors on absorption. *Lipids*;14:378—85.

Cho,S., Park,E.M., Febbraio,M., Anrather,J., Park,L., Racchumi,G., Silverstein,R.L., and Iadecola,C. (2005). The class B scavenger receptor CD36 mediates free radical production and tissue injury in cerebral ischemia. *J. Neurosci.* 25, 2504-2512.

Corpeleijn,E., van der Kallen,C.J., Kruijshoop,M., Magagnin,M.G., de Bruin,T.W., Feskens,E.J., Saris,W.H., and Blaak,E.E. (2006). Direct association of a promoter polymorphism in the CD36/FAT

- fatty acid transporter gene with Type 2 diabetes mellitus and insulin resistance. *Diabet. Med.* 23, 907-911.
- Cechetto JD, Sadacharan SK, Berk PD, Gupta RS. (2002). Immunogold localization of mitochondrial aspartate aminotransferase in mitochondria and on the cell surface in normal rat tissues. *Histol Histopathol*;17(2):353–64.
- Case, D. A.;all. (2012). AMBER 12. University of California, San Francisco.
- Chevrot, M. (2013). Impact de l'obésité sur la détection oro-sensorielle des lipides alimentaires chez la souris et chez l'Homme (Doctoral dissertation, Dijon)
- Degrace-Passilly P, Besnard P. (2012). CD36 and taste of fat. *Curr Opin Clin Nutr Metab Care.* Mar;15(2):107-11. doi: 10.1097/MCO.0b013e32834ff19c. PMID: 22248592.
- Durrant, J. D.; McCammon, J. A. (2011). Molecular dynamics simulations and drug discovery. *BMC Biol*, 9, 71.
- Dunn, N. R., Pearce, G. L., & Shakir, S. A. W. (2000). Adverse effects associated with the use of donepezil in general practice in England. *Journal of Psychopharmacology*, 14(4), 406-408.
- Drover VA, Ajmal M, Nassir F, Davidson NO, Nauli AM, Sahoo D, et al. (2005) CD36 deficiency impairs intestinal lipid secretion and clearance of chylomicrons from the blood. *J Clin Invest*;115(5):1290—7.
- Drover VA, Nguyen DV, Bastie CC, Darlington YF, Abumrad NA, Pessin JE, et al.(2008).CD36 mediates both cellular uptake of very long chain fatty acids and their intestinal absorption in mice. *J Biol Chem*;283(19):13108—15.
- Davletov, B., & Montecucco, C.(2010). Lipid function at synapses. *Current opinion in neurobiology*, 20(5), 543-549.
- Djeziri, F. Z., Belarbi, M., Murtaza, B., Hichami, A., Benammar, C., & Khan, N. A. (2018). Oleonic acid improves diet-induced obesity by modulating fat preference and inflammation in mice. *Biochimie*, 152, 110-120
- de Wit, N. J., Boekschoten, M. V., Bachmair, E. M., Hooiveld, G. J., De Groot, P. J., Rubio-Aliaga, I., et al. (2011). Dose-dependent effects of dietary fat on development of obesity in relation to intestinal differential gene expression in C57BL/6J mice. *PLoS One* 6:e19145.
- Erol A. (2020). Role of oxidized LDL-induced “trained macrophages” in the pathogenesis of COVID-19 and benefits of pioglitazone: A hypothesis. *Diabetes Metab. Syndr*;14:713–714
- El-Yassimi, Abdelghani, Aziz Hichami, Philippe Besnard, et Naim Akhtar Khan. (2008). « Linoleic Acid Induces Calcium Signaling, Src Kinase Phosphorylation, and Neurotransmitter Release in Mouse CD36-Positive Gustatory Cells ». *The Journal of Biological Chemistry* 283 (19): 12949-59
- Ferreiro, E., Resende, R., Costa, R., Oliveira, C. R., & Pereira, C. M. (2006). An endoplasmic-reticulum-specific apoptotic pathway is involved in prion and amyloid-beta peptides neurotoxicity. *Neurobiology of disease*, 23(3), 669-678.



- Fichert, T., Yazdanian, M., & Proudfoot, J. R. (2003). A structure–permeability study of small drug-like molecules. *Bioorganic & medicinal chemistry letters*, 13(4), 719-722.
- Flegal KM, Carroll MD, Ogden CL, Curtin LR (2010). Prevalence and trends in obesity among US adults, 1999-2008. *JAMA* Jan 20;303(3):235-41.
- Floudas C, Fung H, McAllister S, Mönnigmann M, Rajgaria R. (2006). Advances in protein structure prediction and de novo protein design: A review. *Chemical Engineering Science*;61(3):966-88.
- F. Li, (2013). Receptor recognition and cross-species infections of SARS coronavirus, *Antivir. Res.* 100 (1) 246–254.
- Febbraio M, Hajjar DP, Silverstein RL. (2001). CD36: a class B scavenger receptor involved in angiogenesis, atherosclerosis, inflammation, and lipid metabolism. *J Clin Investig*;108:785–91
- Grossfield, A. (2011). Recent progress in the study of G protein-coupled receptors with molecular dynamics computer simulations. *Biochim Biophys Acta*, 1808, 1868-78.
- Guesnet, Philippe, Jean-Marc Alessandri, Pierre Astorg, Fabien Pifferi, and Monique Laviaille. (2005). “Les Rôles Physiologiques Majeurs Exercés Par Les Acides Gras Polyinsaturés (AGPI).” *Oléagineux Corps Gras Lipides* 12: 333–343.
- Ghose, A.K., Herbertz, T., Hudkins, R.L., Dorsey, B.D., Mallamo, J.P., (2011). Knowledge-based, central nervous system (CNS) lead selection and lead optimization for CNS drug discovery, *ACS Chem. Neurosci.* 3(1), 50-68.
- Gualtieri, F., Dei, S., Manetti, D., Romanelli, M. N., Scapecchi, S., & Teodori, E. (1995). The medicinal chemistry of Alzheimer's and Alzheimer-like diseases with emphasis on the cholinergic hypothesis. *Il Farmaco*, 50(8), 489-503.
- Gardinassi L.G., Souza C.O.S., Sales-Campos H., Fonseca S.G. (2020). Immune and Metabolic Signatures of COVID-19 Revealed by Transcriptomics Data Reuse. *Front. Immunol*;11:1636.
- Glatz J.F., Luiken J.J. (2017). From fat to FAT (CD36/SR-B2): Understanding the regulation of cellular fatty acid uptake. *Biochimie*;136:21–26.
- Glatz J.F.C., Luiken J. (2018). Dynamic role of the transmembrane glycoprotein CD36 (SR-B2) in cellular fatty acid uptake and utilization. *J. Lipid Res*;59:1084–1093.
- G. Lu, Q. Wang, G.F. Gao, (2015). Bat-to-human: spike features determining 'host jump' of coronaviruses SARS-CoV, MERS-CoV, and beyond, *Trends Microbiol.* 23 468–478.
- Greenwalt DE, Watt KW, So OY, Jiwani N. PAS IV, (1990). an integral membrane protein of mammary epithelial cells, is related to platelet and endothelial cell CD36 (GP IV). *Biochemistry*.; 29(30): 7054-9.
- Greenwalt DE, Lipsky RH, Ockenhouse CF, Ikeda H, Tandon NN, Jamieson G. (1992). Membrane glycoprotein CD36: a review of its roles in adherence, signal transduction, and transfusion medicine. *Blood*;80:1105–15.

- Huang,H., Campbell,S.C., Bedford,D.F., Nelius,T., Veliceasa,D., Shroff,E.H., Henkin,J., Schneider,A., Bouck,N., and Volpert,O.V. (2004). Peroxisome proliferatoractivated receptor gamma ligands improve the antitumor efficacy of thrombospondin peptide ABT510. *Mol. Cancer Res.* 2, 541-550.
- HyperChem (Molecular Modeling System) Hypercube, Inc.,1115 NW, 4th Street, Gainesville, FL 32601, USA. (2008).
- Hopkins, A. L., Groom, C. R., & Alex, A. (2004). Ligand efficiency: a useful metric for lead selection. *Drug discovery today*, 9(10), 430-431.
- Hansch, C., Leo, A., (1979). *Substituent constants for correlation analysis in chemistry and biology*, Third ed. Wiley, New York.
- Hansson, T.; Oostenbrink, C.; van Gunsteren, W. (2002). Molecular dynamics simulations. *Curr Opin Struct Biol*, 12, 190-6.
- Ibrahimi A, Abumrad NA. (2002). Role of CD36 in membrane transport of longchain fatty acids. *Curr Opin Clin Nutr Metab Care*;5(2):139–45.
- Isabelle HININGER-FAVIER, Université Joseph Fourier, Grenoble. Année universitaire 2010/2011).
- Johnson, T. W., Dress, K. R., & Edwards, M. (2009). Using the Golden Triangle to optimize clearance and oral absorption. *Bioorganic & medicinal chemistry letters*, 19(19), 5560-5564.
- J.P. Lalezari, K. Henry, M. O'Hearn, J.S. Montaner, P.J. Piliero, B. Trottier, J. (2003). Chung,Enfuvirtide, an HIV-1 fusion inhibitor, for drug-resistant HIV infection in North andSouth America, *N. Engl. J. Med.* 348 (22) 2175–2185.
- Kollman, P. (1993). Free-energy calculations : applications to chemical and biochemical phenomena.*Chemical Reviews*, 93, 2395-2417.
- Kitchen D. B., Decornez H., Furr J. R., Bajorath J. (2004): Docking and scoring in virtual discovery: methods and applications. *Drug discovery*. 3: 935-949.
- Karplus, M.; McCammon, J. A. (2002). Molecular dynamics simulations of biomolecules. *Nat Struct Biol*, 9, 646-52.
- K. Champagne, A. Shishido, M.J. Root, (2009). Interactions of HIV-1 inhibitory peptide T20with the gp41 N-HR coiled coil, *J. Biol. Chem.* 284 (6) 3619–3627.
- KHERFIA, B. (2012). *ETUDE DYNAMIQUE ET QUANTIQUE DE LA LIAISONHYDROGENE FAIBLE* (Doctoral dissertation, Université d'Oran).
- Kryger, G., Silman, I., & Sussman, J. L. (1999). Structure of acetylcholinesterase complexed with E2020 (Aricept®): implications for the design of new anti-Alzheimer drugs. *Structure*, 7(3), 297-307.
- Kerns, E.H., Li Di., (2008). *Drug-Like Properties: Concepts, Structure Design and Methods from ADME to Toxicity Optimization*, first ed. Elsevier, United States.

Kar NS, Ashraf MZ, Valiyaveetil M, Podrez EA. (2008). Mapping and characterization of the binding site for specific oxidized phospholipids and oxidized low density lipoprotein of scavenger receptor CD36. *J Biol Chem.*;283:8765–71.

Kliwer SA, Sundseth SS, Jones SA, Brown PJ, Wisely GB, Koble CS, Devchand P, Wahli W, Willson TM, Lenhard JM, Lehmann JM. (1997). Fatty acids and eicosanoids regulate gene expression through direct interactions with peroxisome proliferator-activated receptors alpha and gamma. *Proc Natl Acad Sci U S A.* Apr 29;94(9):4318-23. doi: 10.1073/pnas.94.9.4318. PMID: 9113987; PMCID: PMC20720.

Kuda O, Pietka TA, Demianova Z, Kudova E, Cvacka J, Kopecky J, Abumrad NA. (2013) Sulfo-N succinimidyl Oleate (SSO) inhibits fatty acid uptake and signaling for intracellular calcium via binding CD36 lysine 164. SSO also Inhibits oxLDL uptake by macrophages. *J Biol Chem.*

Kuang, M., Febbraio, M., Wagg, C., Lopaschuk, G.D., and Dyck, J.R. (2004). Fatty acid translocase/CD36 deficiency does not energetically or functionally compromise hearts before or after ischemia. *Circulation* 109, 1550-1557.

Koonen, D.P., Glatz, J.F., Bonen, A., and Luiken, J.J. (2005). Long-chain fatty acid uptake and FAT/CD36 translocation in heart and skeletal muscle. *Biochim. Biophys. Acta* 1736, 163-180.

Kieffer N, Bettaieb A, Legrand C, Coulombel L, Vainchenker W, Edelman L, Breton-Gorius J. (1989) Developmentally regulated expression of 78 kDa erythroblast membrane glycoprotein immunologically related to the platelet thrombospondin receptor. *Biochem J*; 262 : 835-42.

Le Foll, Christelle, Ambrose Dunn-Meynell, Serguei Musatov, Christophe Magnan, et Barry E Levin. (2013). «FAT/CD36: A Major Regulator of Neuronal Fatty Acid Sensing and Energy Homeostasis in Rats and Mice ». *Diabetes* 62 (8): 2709-16.

Laugerette, Fabienne, Patricia Passilly-Degrace, Bruno Patris, Isabelle Niot, Maria Febbraio, Jean-Pierre Montmayeur, et Philippe Besnard. (2005). « CD36 Involvement in Orosensory Detection of Dietary Lipids, Spontaneous Fat Preference, and Digestive Secretions ». *The Journal of Clinical Investigation* 115 (11): 3177-84.

Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced drug delivery reviews*, 23(1-3), 3-25.

Li Y., Yang P., Zhao L., Chen Y., Zhang X., Zeng S., Wei L., Varghese Z., Moorhead J.F., Chen Y., et al. (2019). CD36 plays a negative role in the regulation of lipophagy in hepatocytes through an AMPK-dependent pathway. *J. Lipid Res*;60:844–855.

Lobo MV, Huerta L, Ruiz-Velasco N, Teixeira E, de la Cueva P, Celdran A, et al. (2001). Localization of the lipid receptors CD36 and CLA-1/SR-BI in the human gastrointestinal tract: towards the identification of receptors mediating the intestinal absorption of dietary lipids. *J Histochem Cytochem*;49(10):1253—60.

Lorenza Bordoli 1, 2, Florian Kiefer 1, 2, Konstantin Arnold 1, 2, Pascal Benkert 1, 2, James Battey 1, 2 & Torsten Schwede, (2009). Protein structure homology modeling using SWISS-MODEL workspace ; 30 July.

Lundby A, Lage K, Weinert BT, Bekker-Jensen DB, Secher A, Skovgaard T, Kelstrup CD, Dmytriiev A, Choudhary C, Lundby C, Olsen JV. (2012). Proteomic analysis of lysine acetylation sites in rat tissues reveals organ specificity and subcellular patterns. *Cell Rep*;2:419–31.

Leeson, P. D., & Springthorpe, B. (2007). The influence of drug-like concepts on decision-making in medicinal chemistry. *Nature Reviews Drug Discovery*, 6(11), 881-890.

Lechien JR, Chiesa-Estomba CM, Place S, et al. (2019). Clinical and epidemiological characteristics of 1,420 European patients with mild-to-moderate coronavirus disease. *J Intern Med*. 2020;288:335–344.

Lagarde, N. (2014). Méthodes de criblage virtuel *in silico*: importance de l'évaluation et application à la recherche de nouveaux inhibiteurs de l'interleukine 6 (Doctoral dissertation, Paris, CNAM).

Moullé, V-S, A Picard, C Le Foll, B-E Levin, et C Magnan. (2014). « Lipid Sensing in the Brain and Regulation of Energy Balance ». *Diabetes & Metabolism* 40 (1): 29-33. doi:10.1016/j.diabet.2013.10.001

Murphy, M. P., & LeVine III, H. (2010). Alzheimer's disease and the amyloid- $\beta$  peptide. *Journal of Alzheimer's Disease*, 19(1), 311-323.

Martin, Y. C. (2005). A bioavailability score. *Journal of medicinal chemistry*, 48(9), 3164-3170.

MarvinSketch 17.1.2 (2017), Chemaxon (<http://www.chemaxon.com>).

Marta Yanina Pepino, Ondrej Kuda, Dmitri Samovski and Nada A Abumrad. (2014). Structure-Function of CD36 and Importance of Fatty Acid Signal Transduction in Fat Metabolism, *Annu Rev Nutr.*; 34: 281–303.

Marta Yanina, Latisha Love-Gregory, Samuel Klein, et Nada A Abumrad. (2012). « The Fatty Acid Translocase Gene CD36 and Lingual Lipase Influence Oral Sensitivity to Fat in Obese Subjects ». *Journal of Lipid Research* 53 (3): 561-66.

Mariadason JM, Nicholas C, L'Italien KE, Zhuang M, Smartt HJ, Heerdt BG, et al. (2005). Gene expression profiling of intestinal epithelial cell maturation along the crypt-villus axis. *Gastroenterology*;128(4):1081—8.

Martin C, Chevrot M, Poirier H, Passilly-Degrace P, Niot I, Besnard P. (2011). CD36 as a lipid sensor. *Physiol Behav*;105:36–42.

Masuda D, Hirano K, Oku H, Sandoval JC, Kawase R, Yuasa-Kawase M, et al. (2009). Chylomicron remnants are increased in the postprandial state in CD36 deficiency. *J Lipid Res*;50(5):999—1011.

McRee, D. E. (1999). *Practical protein crystallography*. Elsevier.

Melkemi, N., & Belaidi, S. (2014). Structure-property relationships and quantitative structure-activity relationship modeling of detoxication properties of some 1, 2-dithiole-3-thione derivatives. *Journal of Computational and Theoretical Nanoscience*, 11(3), 801-806.

- Michaud-Agrawal, N., Denning, E. J., Woolf, T. B., & Beckstein, O. (2011). MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *Journal of computational chemistry*, 32(10), 2319-232
- Mokdad AH, Serdula MK, Dietz WH, Bowman BA, Marks JS, Koplan JP. (1999). The spread of the obesity epidemic in the United States, 1991-1998. *JAMA*. Oct 27;282(16):1519-22.
- Moore, K.J., El, K.J., Medeiros, L.A., Terada, K., Geula, C., Luster, A.D., and Freeman, M.W. (2002). A CD36-initiated signaling cascade mediates inflammatory effects of beta-amyloid. *J. Biol. Chem.* 277, 47373-47379.
- Moullé, V-S, A Picard, C Le Foll, B-E Levin, et C Magnan. (2014). « Lipid Sensing in the Brain and Regulation of Energy Balance ». *Diabetes & Metabolism* 40 (1): 29-33.
- Murtaza B, Hichami A, Khan AS, Shimpukade B, Ulven T, Ozdener MH, Khan NA. (2020). Novel GPR120 agonist TUG891 modulates fat taste perception and preference and activates tongue-brain-gut axis in mice. *J Lipid Res.* Feb;61(2):133-142. doi: 10.1194/jlr.RA119000142. Epub 2019 Dec 5. PMID: 31806728; PMCID: PMC6997603.s
- Nauli AM, Nassir F, Zheng S, Yang Q, Lo CM, Vonlehmden SB, et al. (2006). CD36 is important for chylomicron formation and secretion and may mediate cholesterol uptake in the proximal intestine. *Gastroenterology*;131(4):1197—207.
- Nassir F, Wilson B, Han X, Gross RW, Abumrad NA (2007). CD36 is important for fatty acid and cholesterol uptake by the proximal but not distal intestine. *J Biol Chem*;282(27):19493—501.
- Niot I, Poirier H, Tran TT, Besnard P.(2009) Intestinal absorption of long-chain fatty acids: evidence and uncertainties. *Prog Lipid Res*;48(2):101—15.
- Neculai D, Schwake M, Ravichandran M, Zunke F, Collins RF, Peters J, Neculai M, Plumb J, Loppnau P, Pizarro JC, Seitova A, Trimble WS, Saftig P, Grinstein S, Dhe-Paganon S.(2013) Structure of LIMP-2 provides functional insights with implications for SR-BI and CD36. *Nature*.;504:172–6.
- Okumura T, Jamieson GA.(1992 ) Platelet glycoprotein. I. Orientation of glycoproteins of the human platelet surface. *J Biol Chem.* 1976; 251(19): 5944-9.
- human umbilical vein endothelial cells express CD36 *in vivo* and *in vitro*. *J Immunol*; 148
- Pearlman, D. A. (1995) AMBER a package of computer-programs for applying molecular mechanics, normal-mode analysis, molecular-dynamics and free-energy calculations to simulate the structural and energetic properties of molecules. *Computer Physics Communications*, 91, 1-41.
- Palm, K., Stenberg, P., Luthman, K., & Artursson, P. (1997). Polar molecular surface properties predict the intestinal absorption of drugs in humans. *Pharmaceutical research*, 14(5), 568-571
- Patrick, G. N., Zukerberg, L., Nikolic, M., de La Monte, S., Dikkes, P., & Tsai, L. H. (1999). Conversion of p35 to p25 deregulates Cdk5 activity and promotes neurodegeneration. *Nature*, 402(6762), 615.
- Pajouhesh, H., & Lenz, G. R. (2005). Medicinal chemical properties of successful central nervous system drugs. *NeuroRx*, 2(4), 541-553.



Press Release N ° 08/(2022) - Covid-19, flash survey Iss: 95.8% of Omicron positive samples on 17 January]. Istituto Superiore di Sanità. Accessed March 18, 2022.

Parma, V. et al. (2020). More than smell—COVID-19 is associated with severe impairment of smell, taste, and chemesthesis. *Chem. Sens.*45(7), 609–622.

Poirier H, Degrace P, Niot I, Bernard A, Besnard P. (1996) Localization and regulation of the putative membrane fatty acid transporter (FAT) in the small intestine. Comparison with fatty acid-binding proteins (FABP). *Eur J Biochem*;238(2):368—73.

Pepino, Press Release N ° 08/2022 - Covid-19, flash survey Iss: 95.8% of Omicron positive samples on 17 January]. Istituto Superiore di Sanità. Accessed March 18, 2022.

Pepino, M. Y., Kuda, O., Samovski, D., & Abumrad, N. A. (2014). Structure-function of CD36 and importance of fatty acid signal transduction in fat metabolism. *Annual review of nutrition*, 34, 281-303.

Poirier H, Degrace P, Niot I, Bernard A, Besnard P. (1996). Localization and regulation of the putative membrane fatty acid transporter (FAT) in the small intestine. Comparison with fatty acid-binding proteins (FABP). *Eur J Biochem*;238(2):368—73.

Ross AC. (1993). Overview of retinoid metabolism. *J Nutr*;123(2): 346—50.

Stremmel W, Lotz G, Strohmeyer G, Berk PD. (1985). Identification, isolation, and partial characterization of a fatty acid binding protein from rat jejuna microvillous membranes. *J Clin Invest*;75(3):1068–76.

Shiau YF, Fernandez P, Jackson MJ, McMonagle S. (1985). Mechanisms maintaining a low-pH microclimate in the intestine. *Am J Physiol*;248(6 Pt 1):G608–G617.

Simantov,R., Febbraio,M., and Silverstein,R.L. (2005). The antiangiogenic effect of thrombospondin-2 is mediated by CD36 and modulated by histidine-rich glycoprotein. *Matrix Biol.* 24, 27-34.

Sabaouni I, Moussa A, Vannier B, Semlali O, Pietka TA, Abumrad NA, Ibrahim A. (2013). The Whole Genome Expression Analysis using Two Microarray Technologies to Identify Gene Networks That Mediate the Myocardial Phenotype of CD36 Deficiency. *Bioinformatics*.

Silverstein RL, Febbraio M. (2009). CD36, a scavenger receptor involved in immunity, metabolism, angiogenesis, and behavior. *Sci Signal*;2(72):re3.

Swerlick, R. A., Lee, K. H., Wick, T. M., & Lawley, T. J. (1992). Human dermal microvascular endothelial but not human umbilical vein endothelial cells express CD36 *in vivo* and *in vitro*. *The Journal of Immunology*, 148(1), 78-83.

Sukhotnik I, Gork AS, Chen M, Drongowski RA, Coran AG, Harmon CM. (2001). Effect of low fat diet on lipid absorption and fatty-acid transport following bowel resection. *Pediatr Surg Int*;17(4):259—64.

Swerlick RA, LeeKH, WickTM, LawleyTJ. Human dermal microvascular endothelial but not : 78-83.

Sumida C, Graber R, Nunez E. (1993). Role of fatty acids in signal transduction: modulators and messengers. *Prostaglandins Leukot Essent Fatty Acids*. Jan;48(1):117-22. doi: 10.1016/0952-3278(93)90019-s. PMID: 7678706.

Simon, Trang, Victoria R Cook, Anuradha Rao, et Richard B Weinberg. (2011). « Impact of Murine Intestinal Apolipoprotein A-IV Expression on Regional Lipid Absorption, Gene Expression, and Growth ». *Journal of Lipid Research* 52 (11): 1984-1994.

Su, Xiong, et Nada A. Abumrad. (2009). « Cellular Fatty Acid Uptake: A Pathway under Construction ». *Trends in Endocrinology and Metabolism: TEM* 20 (2): 72-77.

Storch J, McDermott L. Structural and functional analysis of fatty acid-binding proteins. *J Lipid Res*. 2009;50(Suppl):S126-31.

Silverstein RL, Febbraio M. (2009). CD36, a scavenger receptor involved in immunity, metabolism, angiogenesis, and behavior. *Sci Signal*;2(72):re3.

Schultes, S., de Graaf, C., Haaksma, E. E., de Esch, I. J., Leurs, R., & Krämer, O. (2010). Ligand efficiency as a guide in fragment hit selection and optimization. *Drug Discovery Today: Technologies*, 7(3), e157-e162.

Salah, T., Belaidi, S., Melkemi, N., & Tchouar, N. (2015). Molecular Geometry, Electronic Properties, MPO Methods and Structure Activity/Property Relationship Studies of 1, 3, 4-Thiadiazole Derivatives by Theoretical Calculations. *Reviews in Theoretical Science*, 3(4), 355-364.

Seltzer, B. (2005). Donepezil: a review. *Expert opinion on drug metabolism & toxicology*, 1(3), 527-536.

Simmons, W. K., Martin, A., & Barsalou, L. W. (2005). Pictures of appetizing foods activate gustatory cortices for taste and reward. *Cerebral Cortex*, 15(10), 1602-1608.

Shah JN, Guo GQ, Krishnan A, Ramesh M, Katari NK, Shahbaaz M, Abdellattif MH, Singh SK, Dua K. (2022). Peptides-based therapeutics: Emerging potential therapeutic agents for COVID-19. *Therapie*.

Santos, R. E. A. et al. (2021). Onset and duration of symptoms of loss of smell/taste in patients with COVID-19: A systematic review. *Am.J. Otolaryngol*. 102889, 2225.

Shanmugamprema D, Muthuswamy K, Subramanian G, Ponnusamy V, Krishnan V, Subramaniam S. 2020. Fat taste signal transduction and its possible negative modulator components. *Prog Lipid Res*. 2020 Jul;79:101035. doi: 10.1016/j.plipres.2020.101035. Epub May 8. PMID: 32437786.

Simons PJ, Kummer JA, Luiken JJ, Boon L. (2011). Apical CD36 immunolocalization in human and porcine taste buds from circumvallate and foliate papillae. *Acta Histochem*;113:839-43.

Spinato G, Fabbris C, Polesel J, et al. (2020). Alterations in smell or taste in mildly symptomatic outpatients with SARS-CoV-2 infection. *JAMA*;323:2089-2090.

Talle MA, Rao PE, Westberg E, Allegar N, Makowski M, Mittler RS, Goldstein G. (1983). Patterns of antigenic expression on human monocytes as defined by monoclonal antibodies. *Cell Immunol*; 78(1): 83-99.

Teichmann SA, Chothia C, Gerstein M. (1999). Advances in structural genomics. *Curr Opin Struct Biol*.;9(3):390-9.

Tracking SARS-CoV-2 variants. World Health Organization (2022). Accessed March 18, 2022.

Tran TT, Poirier H, Clement L, Nassir F, Pelsers MM, Petit V, et al.(2011).Luminal lipid regulates CD36 levels and downstream signaling to stimulate chylomicron synthesis. *J Biol Chem*;286(28):25201—10.

Tran, T. T. T., Buttet, M., Traynard, V., Besnard, P., Poirier, H., & Niot, I. (2012). Mécanisme d'absorption intestinale des acides gras à longue chaîne: rôle émergent du CD36. *Cahiers de Nutrition et de Diététique*, 47(6), 272-279..

Wang X, Song K, Li L, Chen L. (2018).Structure-Based Drug Design Strategies and Challenges. *Curr Top Med Chem*;18(12):998-1006.

Wager, T. T., Hou, X., Verhoest, P. R., & Villalobos, A. (2010). Moving beyond rules: the development of a central nervous system multiparameter optimization (CNS MPO) approach to enable alignment of druglike properties. *ACS chemical neuroscience*, 1(6), 435-449.

Wager, T. T., Hou, X., Verhoest, P. R., & Villalobos, A. (2016). Central nervous system multiparameter optimization desirability: application in drug discovery. *ACS chemical neuroscience*, 7(6), 767-775.

Wager, T. T., Chandrasekaran, R. Y., Hou, X., Troutman, M. D., Verhoest, P. R., Villalobos, A., & Will, Y. (2010). Defining desirable central nervous system drug space through the alignment of molecular properties, *in vitro* ADME, and safety attributes. *ACS chemical neuroscience*, 1(6), 420-434.

Waring, M. J. (2009). Defining optimum lipophilicity and molecular weight ranges for drug candidates—molecular weight dependent lower logD limits based on permeability. *Bioorganic & medicinal chemistry letters*, 19(10), 2844-2851.

Westbrook J, Feng Z, Chen L, Yang H, Berman HM. (2003). The Protein Data Bank and structural genomics. *Nucleic acids research.*;31(1):489-91.

Van de Waterbeemd, H., Camenisch, G., Folkers, G., Chretien, J. R., &Raevsky, O. A. (1998). Estimation of blood-brain barrier crossing of drugs using molecular size and shape, and H-bonding descriptors. *Journal of drug targeting*, 6(2), 151-165.

Veber, D. F., Johnson, S. R., Cheng, H. Y., Smith, B. R., Ward, K. W., &Kopple, K. D. (2002). Molecular properties that influence the oral bioavailability of drug candidates. *Journal of medicinal chemistry*, 45(12), 2615-2623.

Vitkup D, Melamud E, Moulton J, Sander C. (2001). Completeness in structural genomics. *Nature structural biology*;8(6):559-66.

Vaira LA, Deiana G, Fois AG, et al. (2020). Objective evaluation of anosmia and ageusia in COVID-19 patients: single-center experience on 72 cases. *Head Neck*;42:1252–1258.

YesnerLM, HuhHY, PearceSFA, SilversteinRL. (1995) Regulation of monocyte CD36 and thrombospondin-1 expression by soluble mediators. *Arterioscler Thromb Vasc Biol*; 16: 1019-25.

Zerroug, A., Belaidi, S., BenBrahim, I., Sinha, L., & Chtita, S. (2019). Virtual screening in drug-likeness and structure/activity relationship of pyridazine derivatives as Anti-Alzheimer drugs. *Journal of King Saud University-Science*,31(4),595-601.

Zineb Tarhda and Azeddine Ibrahimi. (2015). Insight into the mechanism of lipids binding and uptake by CD36 receptor *Bioinformation*; 11(6): 302–306.

Zuccotti G, Fabiano V, Dilillo D, Picca M, Cravidi C, Brambilla P. (2013). Intakes of nutrients in Italian children with celiac disease and the role of commercially available gluten-free products. *J Hum Nutr Diet*. Oct;26(5):436-44. doi: 10.1111/jhn.12026. Epub 2012 Dec 17. PMID: 23252793.

# **PUBLICATIONS ET COMMUNICATIONS**





## Jihane AKACHAR

Marocaine, célibataire

Adresse : B.P 5, App 9, Ocean  
Rabat  
Téléphone : (+212) 6 68 21 1698  
Email: [jihane.akachar@gmail.com](mailto:jihane.akachar@gmail.com)

### Formation

---

- 2015-2020**      **Doctorat en Sciences de la Vie et de la Santé**, spécialité Biotechnologie Médicale et Bioinformatique. Faculté de Médecine et de Pharmacie, Université Mohamed V de Rabat.  
**Sujet de thèse** : Développement d'un protocole pour la découverte des molécules bioactives par des approches in Silico **Outils** : IA, Big Data, Chémoinformatique, Bioinformatique Structurale, Modélisation Moléculaire, Dynamique Moléculaire, Docking (Amarrage Moléculaire).
- 2008-2010**      **Master Intelligence Artificiel et Bioinformatique**, ENSA Tanger.
- 2005-2008**      **Licence Science de la Vie (SVI)**, Université Mohammed Premier Oujda.
- 2003-2004**      **Baccalauréat en Sciences Expérimentales**, Lycée El Jadida Nador.

### Expériences Professionnels

---

- Aout 2020**  
**Janvier 2021**      **Superviseur méthodologique et scientifique (E-Santé) au sein de l'université internationale Abulcasis (IASS) Rabat.**
- Septembre 2017**  
**Aout 2020**      **Professeur vacataire contrat CDD au sein faculté de médecine et pharmacie de Rabat**  
-Participer par des cours en Modélisation, Intelligence Artificiel, Ontologie
- Juin2017-**  
**Aout2017**      **Stage au sein du Laboratoire Dynamique des Structures et des Interactions des Macromolécules Biologiques, DSIMB, UMR1231 /INSERM Paris**  
**Sujet** : Dynamique moléculaire des macromolécules transmembranaire et des molécules chimiques.
- Avril2017-**  
**juin2017**      **Stage au sein du Laboratoire Physiologie de la Nutrition & Toxicologie, UMR 1231 INSERM/UB/AgroSup, Université de Bourgogne**  
**Sujet** : Etudes et teste de nouvelles molécules au niveau des cellules de bourgeons de goût par analyse de la signalisation calcique in vitro, culture cellulaire et génotypage.
- Janvier2011-**  
**juin2013**      **Stage de doctorat au sein du Laboratoire des Technologies (LTI), ENSAT Tanger**  
**Sujet** : La création d'une Ontologie Informatique dans le domaine de la neuroscience.
- Janvier2010-**  
**juin2010**      **Stage de fin d'étude du Master, Laboratoire de Modélisation des Phénomènes Physiques et Risques Naturelles (MPPRN), Faculté des sciences et techniques (FST) Tanger.**  
**Sujet** : Développement d'algorithme pour la modélisation moléculaire en 3D.

## Compétences

<b>Biologie</b>	Biologie Moléculaire, Biochimie, Enzymologie, Microbiologie, Génétique, Biologie cellulaire
<b>Informatique</b>	Programmation, Base de données, R, IA, Suites Office, Développement web,
<b>Modélisation</b>	Moléculaire Dynamique Moléculaire, Statistique et Mathématique

## Centres d'intérêt

<b>Activités Associatives</b>	Membre de l'association Biotechnologie Médicale Rabat , Participation à l'organisation d'une conférence internationale 'Euro-Mediterranean Medical Informatics and Telemedicine' à la faculté Polydisciplinaire Nador.
<b>Loisirs</b>	Lecture, tennis, Cinéma, Cuisine, Voyage, Music
<b>Webmaster</b>	Intranet du Laboratoire Biotechnologie Médicale (2016)

## Publications

### Articles

- Guerrab, W., **Akachar, J**, El Jemli, M. (2022). *Synthesis, molecular docking, ADMET evaluation and in vitro cytotoxic activity evaluation on RD and L20B cell lines of 3-substituted 5,5-diphenylimidazolidine-2,4-dione derivatives*. Journal of Biomolecular Structure and Dynamics.
- Nadeem Abad, Nada Kheira Sebbar, Ahmed Talbaoui, Youssef Bakri, **Jihane Akachar**. (2022). *SYNTHESIS AND MODELING OF NEW 1,2,3-TRIAZOLYLMETHYL-QUINOXALINE DERIVATIVES OBTAINED VIA 1,3-DIPOLAR CYCLOADDITION REACTIONS*. JMCH.
- Akachar, J**, Etchebest, C, El Jaoudi, R & Ibrahimi, A. (2021). *The computational analyses, molecular dynamics of fatty-acid transport mechanism to the CD36 receptor*. Nature Journal of Scientific Report.
- Guerrab, W., El Jemli, M., **Akachar, J**, Demirtaş, G., Mague, J. T., Taoufik, J., ... & Ramli, Y. (2021). *Design, synthesis, structural and molecular characterization, toxicity, psychotropic activity and molecular docking evaluation of a novel phenytoin derivative: 3-decyl-5,5-diphenylimidazolidine-2,4-dione*. Journal of Biomolecular Structure and Dynamics.
- Akachar, J**, Bouricha, E. M., Hakmi, M., Belyamani, L., El Jaoudi, R., & Ibrahimi, A. (2020). *Identifying epitopes for cluster of differentiation and design of new peptides inhibitors against human SARS-CoV-2 spike RBD by an in-silico approach*. Heliyon,
- Hakmi, M., Bouricha, E. M., **Akachar, J**, Lmimouni, B., El Harti, J., Belyamani, L., & Ibrahimi, A. (2020). *In silico exploration of small-molecule  $\alpha$ -helix mimetics as inhibitors of SARS-COV-2 attachment to ACE2*. Journal of Biomolecular Structure and Dynamics.
- Bouricha, E. M., Hakmi, M., **Akachar, J**, Zouaidia, F., & Ibrahimi, A. (2020). *In-silico identification of potential inhibitors targeting the DNA binding domain of estrogen receptor  $\alpha$  for the treatment of hormone therapy-resistant breast cancer*. Journal of Biomolecular Structure and Dynamics.
- Bouricha, E. M., Hakmi, M., **Akachar, J**, Belyamani, L., & Ibrahimi, A. (2020). *In silico analysis of ACE2 orthologues to predict animal host range with high susceptibility to SARS-CoV-2*. 3 Biotech.
- Al-kaf, A. G. (2020). *Introductory Chapter: The Newest Research in Quinazolinone and Quinazoline Derivatives*. Quinazolinone and Quinazoline Derivatives.
- Chemlal, L., **Akachar, J**, Makram, S., Zoubir, B., Cherrah, Y., Eljaoudi, R., ... & Faouzi, M. A. (2019). *The displacement study of  $^{99m}\text{Tc}$ -DTPA—Human serum albumin binding in presence of furosemide and metformin by using equilibrium dialysis and molecular docking*. IUBMB life.
- Akachar, J**, Abudunia, A, El Jaoudi, R, Bennani, FA, Ansar, M, Ibrahimi, A, Al-kaf, A. (2018). *Synthesis, Anti-hypoxic Activity and Molecular Docking Studies of New Quinazolinone-4 Derivatives*. ISESCO Journal of Science and Technology.
- Allam, L., Lakhlii, W., Tarhda, Z., **Akachar, J**, Ghriji, F., Amri, H. E., & Ibrahimi, A. (2017). *Three-dimensional structure prediction of the human LMTK3 catalytic domain in DYG-in conformation*. J Biomol Res Ther.
- Al-kaf, A, **Akachar, J**. (2017). *SYNTHESIS, ANALGESIC, ANTI-INFLAMMATORY AND MOLECULAR DOCKING STUDIES OF NEW QUINAZOLIN-4-ONE DERIVATIVES*. Moroccan Journal of Heterocyclic Chemistry.

## RESUME

Une consommation excessive de lipides en particulier d'Acides Gras Saturés associé à un déséquilibre qualitatif contribue à la prévalence de l'obésité. Le récepteurs lipidiques CD36 est un candidat plausible en tant que capteur lipidique gustatif, également induisant une cascade de signalisation cellulaire. Le but de cette thèse a été la conception, le criblage et l'identification de nouvelles molécules bioactives par le développement de workflow basés sur l'approche in silico. Nous avons utilisé des séries de simulations de dynamique moléculaire de la protéine CD36, en interaction avec un à trois acides palmitiques (PLM), afin d'établir un premier workflow pour découvrir le mécanisme du fonctionnement du complexe CD36/Acides gras et de déterminer le rôle structural de la Lys 164. Les résultats ont montré que la Lys164 influence d'une manière importante sur la forme de stabilité du CD36 et. La conception grâce aux résultats du premier workflow de 15 molécules activant la voie de signalisation des acides gras et plus particulièrement ceux qui sont spécifiques au goût de gras. Uniquement, l'analogue CIN4 a montré, par la suite, par criblage dynamique, une dimérisation des deux domaines TM1 et TM2 qui ressemble à l'acide palmitique. En outre, les étapes de ce premier workflow ont permis d'explorer le dysfonctionnement gustatif de la chimio-réception des acides gras spécifiquement le "goût de gras" des patients testés positifs au COVID-19. Les résultats de cette étude ont montré une affinité élevée et une stabilité du domaine RBD du SARS-CoV-2 non mutant au récepteur CD36, un deuxième workflow a été établi. Le criblage virtuel a été utilisé dans un premier temps pour identifier les épitopes critiques pour trois cibles (CD147, CD209 et CD299). Les résultats ont montré une affinité élevée du domaine RBD du SARS-CoV-2 pour le récepteur CD299. Le CD36 et CD299 représente, donc, une cible puissante pour la découverte et la conception d'un nouveau groupe thérapeutiques ciblant le dysfonctionnement du goût par le Covid 19 et neutralisant le SRAS-CoV-2.

**Mot-clés:** Acide gras, CD36, Dynamique moléculaire, Modélisation moléculaire, Bioinformatique, Conception de molécule, COVID19.

## ABSTRACT

Excessive consumption of lipids, in particular Saturated Fatty Acids, associated with a qualitative imbalance contributes to the prevalence of obesity. The CD36 lipid receptor is a plausible candidate as a taste lipid sensor, also inducing a cellular signaling cascade. The aim of this thesis was the design, screening and identification of new bioactive molecules through the development of workflows based on the in silico approach. We used series of molecular dynamics simulations of the CD36 protein, in interaction with one to three palmitic acids (PLM), in order to establish a first workflow to discover the mechanism of the functioning of the CD36/fatty acid complex and to determine the structural role of Lys 164. The results showed that Lys164 has an important influence on the form of stability of CD36 and. The design thanks to the results of the first workflow of 15 molecules activating the signaling pathway of fatty acids and more particularly those which are specific to the taste of fat. Only the CIN4 analog subsequently showed, by dynamic screening, a dimerization of the two TM1 and TM2 domains which resembles palmitic acid. In addition, the steps of this first workflow made it possible to explore the gustatory dysfunction of fatty acid chemoreception specifically the "fat taste" of patients who tested positive for COVID-19. The results of this study showed high affinity and stability of the non-mutant SARS-CoV-2 RBD domain at the CD36 receptor, a second workflow was established. Virtual screening was first used to identify critical epitopes for three targets (CD147, CD209 and CD299). The results showed high affinity of the SARS-CoV-2 RBD domain for the CD299 receptor. CD36 and CD299 therefore represent a powerful target for the discovery and design of a new therapeutic group targeting taste dysfunction by Covid 19 and neutralizing SARS-CoV-2.

**Keywords:** Fatty acid, CD36, Molecular dynamics, Molecular modeling, Bioinformatics, Molecule design, COVID19.

**Structure de recherche:** Laboratoire de Biotechnologie Médicale (MedBiotech), Faculté de Médecine et Pharmacie de Rabat.